TOWARDS SMART EARTHWORK SITES USING LOCATION-BASED GUIDANCE AND MULTI-AGENT SYSTEMS

Faridaddin Vahdatikhaki

A Thesis
in the Department
of
Building, Civil and Environmental Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

July 2015

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Faridaddin Vahdatikhaki

Entitled: **Towards Smart Earthwork Sites Using Location-based Guidance and Multi-agent Systems**

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Building Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Nematollaah Shiri	Chair
Dr. Carl Haas	External Examiner
Dr. Javad Dargahi	External to Program
Dr. Tarek Zayed	Examiner
Dr. Zhenhua Zhu	Examiner
Dr. Amin Hammad	Supervisor

Approved by _____

Chair of Department or Graduate Program Director

_____ _____

Dean of Faculty

ABSTRACT

Towards Smart Earthwork Sites Using Location-based Guidance and Multi-agent Systems
Faridaddin Vahdatikhaki, Ph.D.
Concordia University, 2015

The growing complexity and scope of construction projects is making the coordination and safety of earthwork of a great concern for project and site managers. The difficulty of safeguarding the construction workers is mainly commensurate with the type, scale, and location of the project. In construction operations, where heavy machines are used, various safety and risk issues put the timely completion of a project at stake. Additionally, the construction working environment is heavily susceptible to unforeseen changes and circumstances that could impact the project, both cost and schedule wise. As a response to the looming safety threats or unforeseen changes of working conditions, re-planning is almost always required, in both proactive (preemptive) or reactive (corrective) fashion. In order for re-planning to yield the optimum results, real-time information gathering and processing is a must. Global Positioning System (GPS) and other Real-time Location Systems (RTLSs) have been used for the purpose of real-time data gathering and decision-making in recent years. Similarly, Location-based Guidance Systems (LGSs), e.g., Automated Machine Control/Guidance (AMC/G), have been recently introduced and employed, mainly for the purpose of high-precision earthwork operations. However, currently the application of available *LGSs* (i.e., *AMC/G*) is restricted to the machine-level task control and improvement. Also, the high cost of procuring available *LGSs*, which cost approximately $80,000 for every new piece of equipment, limits the availability of *LGSs* for small and medium size contractors. Furthermore, the valuable real-time data gathered from various pieces of equipment on site are not effectively utilized to

continuously update the simulation models developed at the design phase so that a more realistic view of project progress is available in the execution phase. Finally, despite the growing availability of *LGSs*, their application for safety is limited to real-time proximity-based object detection and warnings. In view of the ability to control the finest motion of *LGS*-enabled earthwork equipment, there is a great potential to boost their level of application to the project level, where decisions about the equipment control are made based on the global consideration of a fleet rather that a local view of one single equipment. To the best of the author's knowledge, a generic methodology that combines real-time data-gathering technologies, *LGS* and intelligent decision making tools, particularly Multi-agent Systems (*MASs*), and addresses the safety-sensitive re-planning, is missing.

On this premise, this research pursues a methodology which addresses the issue of coordination and safety improvement through the integration of *LGSs* and *MASs*. In a nutshell, this research is dedicated to the pursuit of the following objectives: (1) to enable the project-level coordination, monitoring and control through the integration of a *MAS* architecture and a *LGS* to help better resolve operational and managerial conflicts; (2) to provide a method for improving the performance of pose estimation based on affordable *RTLS*s so that *LGS*s can be applied to a wider scope of older earthwork equipment; (3) to devise a generic framework for Near Real-Time Simulation (*NRTS*) based on data from *LGSs*; and (4) to develop a mechanism for improving the safety of earthwork operations using the capabilities of the *LGS*, *NRTS,* and *MAS*.

In the proposed framework, every staff member of the project is represented by an exclusive agent in the *MASs*. More affordable positioning technologies, such as Ultra-Wideband (UWB), are utilized to provide accurate real-time data about the location of machines and workers. An optimization-based method is proposed to consider a set of geometric and operational constraints

that govern the behavior of the Data Collectors (DCs) attached to the equipment to improve the equipment pose estimation accuracy. *NRTS* is used to keep track of the progress of the project and fine-tune the schedule based on the data captured from the site. The agents observe the progress of work executed by their associated equipment, and if any anomalies are detected, viable corrective measures are devised and executed. The inputs to this system are: (a) a stream of real-time data, e.g., location data, flowing from the site, (b) the project design data, and (c) the project progress data and the schedule. Furthermore, a two-layer safety mechanism monitors the safe operation of different pieces of equipment. The first layer of this mechanism enables the equipment to plan a collision-free path considering the predicted movement of all other pieces of equipment. The second layer is acting as a last line of defense in view of possible discrepancies between the predicted paths and actual paths undertaken by the operators. Several prototypes and case studies are developed to demonstrate and verify the feasibility of the proposed framework.

It is found that the proposed optimization-based method has a very strong potential to improve the pose estimation using redundancy of more affordable *RTLS DC*s. Also, the proposed overarching *NRTS* approach provides a tracking-technology-independent method for processing, analyzing, filtering and visualizing the equipment states that can work with various types of *RTLS* technologies and under the availability of different levels of sensory data. The proposed safety system is found to provide a balance between economic use of space and the ability to warn against potential collisions in an effective manner using the pose, state, geometry, and speed characteristics of the equipment. Additionally, the safety system demonstrates the ability to provide a reliable basis for the generation of the risk maps of earthwork equipment, using the expected pose and state, and considering the proximity-based and visibility-based risks.

The *MAS*-based framework helps expand the effective domain of *LGS*s from machine-level guidance to fleet-level coordination. In the view of the presented case studies, the *MAS* structure is found to be effective in assigning different operations and tasks of a project to the specific agents that will be responsible for their realization. Using a combination of strategic and tactical planning methods, the *MAS* is able to effectively provide readily executable guidance/control for equipment operators considering a variety of safety issues.

Dedicated to my parents and my wife for their interminable love and unwavering support

ACKNOWLEDGEMENTS

First and foremost, I feel obliged to extend my heartfelt appreciations to Dr. Amin Hammad for his omnipresent support. In the course of the past few years, he epitomized the true meaning of mentoring by being always available, resourceful and inspiring. This thesis is owed, more than anything else, to his intellectual inputs, moral encouragement and selfless commitment to conveying his knowledge to a fare-thee-well. The legacy of his excellent work as an advisor will shape my professional career for years to come.

I would, also, like to thank the member of my committee, Dr. Carl Haas, Dr. Javad Dargahi, Dr. Tarek Zayed, and Dr. Zhenhua Zhu for their valuable inputs and precious time.

This research has been supported by grants from the National Science and Engineering Research Council of Canada and WorkSafeBC. This support is greatly appreciated.

I have benefited from the tremendous intellectual and moral support of several intelligent individuals with whom I have collaborated and interacted in the course of my Ph.D. research. Hassan Siddiqui and Shayan Setayeshgar contributed to this research by helping me conduct multiple lab tests and case studies. Their dedication and support are deeply appreciated. The invaluable contributions of Homam AlBahnassi and Seyed Mohammad Langari, Khaled El Ammari, Bruno Paes, and Henrique Machado Gasparotto to the implementation of the prototype and case studies for Look-ahead Equipment Workspace are greatly acknowledged. Alhussain Taher made a greatly appreciated contribution to this research by implementing the structure of the multi-agent system in the specialized agent design tool (*JADE*). Mohammed Mawlana has been a fantastic friend who never refrained from giving me constructive and valuable comments. I would like to thank Dr. Cheng Zhang and all my other colleagues who have helped me during

my studies at Concordia. I would also like to acknowledge the logistical support of Mr. Pierre Vilgrain from PCL Constructors Westcoast in conducting the outdoor test performed in Vancouver.

My research would have not come to fruition if it was not for the endless and loving support of my parents. Their helping hands were my only recourse in years fraught with tension and frustration. My wife has infused me with love and care throughout these years. I cannot thank her enough for the patience and understanding with all the irregular working hours and absenteeism.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ABBREVIATIONS

ABS	Agent-based Simulation
ACD	Activity Cycle Diagram
AI	Artificial Intelligence
AMC	Automated Machine Control
AMC/G	Automated Machine Control and Guidance
AMG	Automated Machine Guidance
API	Application Programming Interface
AR	Augmented Reality
BLS	Bureau of Labor Statistics
DC	Data Collector
DDA	Design Document Agent
DGPS	Deferential Global Positioning System
DIRT	Damage Information Reporting Tool
CA	Coordinator Agent
CGA	Common Ground Alliance
CPI	Cost Performance Index
CPM	Critical Path Method
CPGPS	Carreir Phase Enhancement Global Positioning System
CPU	Central Processing Unit
CR	Center of Rotation
CRCP	Continuously Reinforced Concrete Pavement
DC	Data Collector
DES	Discrete-Event Simulation
DEW	Dynamic Equipment Workspace
DOF	Degrees of Freedom
DOT	Department of Transportation
DTM	Digital Terrain Model
EF	Environmental Factor
GC	Geometric Constraints

GCA	General Coordinator Agent
GIS	Geographic Information System
GPS	Global Positioning System
HIVACS	Hazama Intelligent Vehicle Automatic Control System
IDEF	Integration Definition
IMU	Inertial Measuring Unit
INDOT	Indiana Department of Transportation
INS	Inertial Navigation System
LAEW	Look-ahead Equipment Workspace
LGS	Location-based Guidance System
LIDAR	Light Detection and Ranging
MAS	Multi-Agent System
NRTS	Near Real-Time Simulation
OA	Operator Agent
OBI	On-Board Instrumentation
OC	Operational Constraint
OO	Objected-Oriented
PDA	Project Document Agent
PDF	Probability Density Function
PERT	Program Evaluation and Review Technique
PPI	Project Performance Indicators
QC/QA	Quality Control and Quality Assurance
RC	Radio-Controlled
RFID	Radio Frequency Identification
RTK	Real-time Kinematic
RTLS	Real-time Location System
SD	System Dynamics
SSA	Site State Agent
TCA	Team Coordinator Agent
TDMA	Time Division Multiple Access
TIN	Triangulated Irregular Network

UML	Unified Modeling Language
UWB	Ultra-Wideband
VBA	Visual Basic for Applications
WA	Worker Agent
WISDOT	Wisconsin Department of Transportation
WSN	Wireless Sensor Network

CHAPTER 1: INTRODUCTION

1.1 GENERAL BACKGROUND

The construction industry is concerned with improving the productivity and safety of construction projects (Beavers et al. 2006). In earthwork operations, where heavy machines are used, various safety and risk issues put the timely completion of a project at stake. Additionally, the construction working environment is heavily susceptible to unforeseen changes and circumstances that could impact the project, both cost and schedule wise. As a response to the looming safety threats or unforeseen changes of working conditions, re-planning is almost always required. In order for re-planning to yield the optimum results, real-time information gathering and processing is a must. The Global Positioning System (GPS) and other Real-time Location Systems (*RTLS*s) have been used for the purpose of real-time data gathering and decision-making in recent years (Perkinson et al. 2010). Similarly, Location-based Guidance Systems (LGSs), e.g., Automated Machine Control/Guidance (AMC/G), are introduced and have been employed mainly for the purpose of high-precision earthwork operations. *LGS* integrates geo-positioning technologies with 3D design models and Digital Terrain Models (DTMs) to either (1) support the machine operator through the provision of continuous guidance on a digital screen mounted in the cabin of the machine, or (2) control the position and movements of the equipment. While *GPS* and total stations are the main tracking technologies used in *AMC/G*, other types of emerging *RTLSs*, e.g., Ultra-Wideband (UWB), can be integrated with similar monitoring mechanisms to provide monitoring and guidance capabilities for earthwork equipment.

1.2 PROBLEM STATEMENT

In spite of the benefits of *LGS*s, their current application is limited to the improvement of machine-level tasks, which is not sufficient to address the project-level monitoring and decision-making needs. In the context of complex earthwork operations, where several teams are concurrently working towards different ends, it is conceivable that the globally coordinated operation dictates an arrangement other than the machine-level optimized work. In other words, merely optimizing the operation of every single piece of equipment may not result in the best arrangement for the entire project.

The problem of providing near real-time guidance or control support for the operators of earthwork equipment based on the consideration of the entire fleet can become complex, in line with the fleet size and equipment interactions. For such complex problems, the conventional approach of central problem solving becomes far-fetched, because it is difficult or impractical to globally grasp and analyze the multi-dimensionality and dynamisms of such problems. Distributed intelligent systems are designed to address such complex problems in terms of several collaborating intelligent agents, who try to solve the overall problem by synthesizing limited views of individual agents (Ferber 1999). Such systems are referred to as Multi-Agent Systems (*MAS*s), which consist of several intelligent agents capable of interaction.

Also, the high cost of procuring available *AMC/G*, which costs approximately $80,000 for every new piece of equipment (Guy 2011), limits the availability of *AMC/G* for small and medium size contractors. While this cost pertains to the full AMC/G systems including the control mechanisms, even the cost related to the equipment required for providing only the guidance component (which includes two high accuracy GPS receivers) is still exorbitant. On the other hand, the more affordable variations of tracking technologies, e.g. *UWB* and low-cost *GPS*

receivers, do not generate accurate enough data that can be readily used for the construction equipment pose estimation and tracking (Zhang and Hammad 2012; Pradhananga and Teizer 2013). As a result, in order to enable the efficient application of the more affordable *RTLS*s for equipment pose estimation, it is of a paramount importance to process and refine the captured data. It is required to ensure that the refined data satisfy a set of geometric and operational constraints, which are induced by the equipment-specific and job-and-site-specific characteristics. Although several researchers have previously worked in this area (Vahdatikhaki and Hammad 2014; Zhang et al. 2012; Rezazadeh Azar et al. 2015), the area of *RTLS* for construction equipment pose estimation is noticeably under-addressed.

As part of the global coordination of the fleet, detailed planning is always a requirement. Simulation techniques are used to effectively leverage the existing information about projects to develop such detailed plans. Traditionally, these simulation techniques rely mainly upon the statistical data gathered from previous projects of the like nature. However, given the volatility and dynamism of a construction system and the high sensitivity of short-term schedules to variations of constituent parameters, very often it is the case that the initially developed simulation model fails to remain valid and representative of the actual work. To compensate for the possible discrepancies and deviations from the estimated values, i.e. the time and cost of the project, and also in order to make simulation results more realistic, the assumptions made at the time of the model design need to be constantly modified and attuned based on the actual state of the operation. Consequently, some research has been conducted on a mechanism to dynamically monitor the trend of changes in the model parameters and to adapt the simulation model accordingly in near real-time (Lu et al. 2007; Hammad and Zhang 2011; Akhavian and Behzadan 2012; Song and Eldin 2012; Szczesny et al. 2013). However, to the best of the author's

knowledge, the previously proposed methods are not tailored for the integration with *LGSs*. Additionally, both the scope and level of data capturing is limited to one type of equipment, i.e. trucks, and high-level equipment states, e.g. hauling and dumping.

Finally, despite the growing availability of *LGSs*, their application for safety is limited to real-time proximity-based object detection and warnings. In the existing systems, the increasingly affordable advanced sensing and location systems are used to mitigate the collision risks by warning the operators against the potential dangerous proximities in real time (Burns 2002; Carbonari et al. 2011; Zhang and Hammad 2012; Guenther and Salow 2012; Wu et al. 2013; Zolynski et al. 2014). However, the existing methods do not take a full advantage of pose, state, equipment geometry, and speed characteristics of the equipment to improve the efficiency of the collision avoidance. Additionally, there is a need for a solution that is able to reliably predict the operation of the equipment for a long-enough time window to enable different pieces of equipment to adjust their planned paths to avoid collisions in near-real time.

1.3 RESEARCH OBJECTIVES

The objectives of this research are: (1) to enable the project-level coordination, monitoring and control through the integration of a *MAS* architecture and a *LGS* to help better resolve operational and managerial conflicts in earthwork projects; (2) to provide a method for improving the performance of pose estimation based on affordable *RTLS*s so that *LGS*s can be applied to a wider scope of older earthwork equipment; (3) to devise a generic approach for Near Real-Time Simulation (*NRTS*) based on data from *LGSs*; and (4) to develop a mechanism for improving the safety of earthwork operations using the capabilities of the *LGS*, *NRTS* and *MAS*.

1.4 RESEARCH SCOPE

The present research pursues the above-mentioned objectives within the following scope:

(1) The context, within which the proposed method of this research is presented, is mainly automated guidance rather than control. Nevertheless, all the methods that will be presented in this research can further be input for the development of control mechanisms for automated machine control.

(2) Improved *LGS*-based coordination is addressed in the monitoring phase of the projects and mainly at the operational-level (i.e., not at the managerial level);

(3) The presented pose estimation method is intended to improve the applicability of more affordable *RTLSs* for the application in *LGC*s. While in principle it can be used for improving the pose estimation by any tracking technologies, the method has limited application for high-accuracy *GPS*, where the quality of pose estimation is accurate enough without any processing.

(4) The pose estimation method is framed as a complementary process where the final tracking outputs of the *RTLS*s are subject to corrections based on the consideration of equipment-specific characteristics and leveraging the redundancy of *RTLS*s data collectors (DCs) attached to the equipment. This method does not pursue the improvement of the *RTLS*s localization accuracy for single *DC*s.

(5) While it is indispensable for the full safety management of earthmoving sites to also account for workers on-foot, this research focuses primarily on the earthmoving equipment.

(6) The emphasis of the presented safety management system is placed on providing robust inputs for collision avoidance and path re-planning. This research does not address methods

for near real-time path re-planning that can be applied once the potential collisions are detected.

1.5 RESEARCH SIGNIFICANCE

The present research aims to enhance the intelligent decision-making at different operational and managerial levels of earthwork projects using real-time data captured from the site. The present research exploits the increasingly available and affordable *LGS*s to surpass the limits of *LGSs* and transcend them to a project-level coordination tool committed to enhancing the safety of the entire project. This research will furnish contractors with stronger incentives for transiting to *LGS*-based project execution as a technology that can endow them with a greater control over the project and make the project schedule less susceptible to delays and overruns. On the other hand, the manufacturers of *AMC/G* equipment can benefit from this research to broaden their perspective of how a group of equipment can be orchestrated into a coherent fleet. This research would also contribute to improving the safety of earthwork sites by accounting for a wider spectrum of uncertainties and human factors. Accordingly, safety managers and operators can benefit from this research through receiving timely safety-related warnings that can be used to take appropriate actions.

1.6 THESIS LAYOUT

The structure of the thesis is as follows: Chapter 2 is dedicated to the review of the literature on a wide spectrum of subjects pertinent to the research. Chapter 3 presents the overall proposed framework. Chapter 4 discusses the pose estimation method that can be used to improve the quality of *RTLS*-based pose estimation data for the application in *LGS*. Chapter 5 addresses the *NRTS* approach that can be used for enhanced planning of the earthwork projects. Chapters 6 and

7 present the two-layer safety mechanism that can be used for real-time collision avoidance and near real-time collision-free path re-planning, respectively. Chapter 8 covers the components and functionalities of different types of agents in the *MAS* and demonstrates the integration of multiple functions in the proposed *MAS*. Finally, the research conclusions, contributions, and future work are discussed in Chapter 9.

CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

This chapter presents the review of literature in the areas that pertains to the scope of this research. These areas include the simulation and its application in construction industry, the state-of-the-art in automation and *LGS* applications for earthwork equipment, monitoring and control of project performance; safety of earthwork sites, and use of *MAS* for distributed problem solving. The purpose is to identify the shortcomings of current planning, monitoring and coordination of earthwork fleets and to determine how *LGS* can help develop a comprehensive solution that can perform coordination of equipment at a fleet level.

Earthwork refers to a set of operations leading to reshaping the natural surface of earth and encompasses such activities as cutting, hauling, unloading, grading, humidifying and compaction (de Athayde Prata et al. 2008). This operation is a common part of projects such as building foundations work, dam construction, airport construction, road construction, strip-mining, etc. (Shi and AbouRizk S. 1998). Earthwork operations are typically equipment-intensive and thus fraught with financial and safety risks (Peurifoy et al. 2011). The cost of earthwork operations is considered an integral element of the total cost of civil engineering projects (Lambropoulos et al. 1996), so much so that more than 20% of the total cost of the road building projects is estimated to be dedicated to the earthwork operations (Smith et al. 1996). On the other hand, based on the data published by The Bureau of Labor Statistics (BLS 2012), only in 2012, 74 out of the total 775 fatalities (nearly 10%), have been reported as primarily or secondarily caused by major earthwork equipment, e.g., excavators, loaders, graders, scrapers, compactors, or dump trucks.

According to Hinze and Teizer (2011), one-fourth of construction fatalities are due to equipment-related incidents.

Given these statistics, the proper planning of earthwork equipment is indispensable for the safe and efficient execution of earthwork projects. Simulation techniques are used to capture the complexity of interactions between earthwork equipment in computer models. These models can be used to make accurate planning of earthwork operations considering the uncertainties revolve around the execution of such complex operations. Different simulation techniques that are applied for earthwork operations and their shortcomings are introduced and discussed in Section 2.2.

The automation and *LGS*s are introduced and applied to earthwork equipment in recent years to improve the efficiently of equipment operation and ensure accurate pursuit of design plans and schedules. These systems synthesize varied types of data from multiple sensory sources and provide directions for the operator in form of guidance or (partial or full) control. The constituent components and various types of *LGS*s are introduced in Sections 2.3 and 2.3.1, respectively.

While enhanced planning and execution is important for successful competition of earthwork projects, there is a need for continuous progress monitoring and control of projects. These practices help to ensure that project is on the right track and to identify potential causes of discrepancies between the current state of the project and plans. Several automated solutions are introduced and implemented based on the application of various types of *RTLS*s. However, if *LGS*s are properly utilized, their robust tracking capabilities can be well harnessed for the purpose of project monitoring and control. Section 2.4.1 provides an overview of earthwork project monitoring and control practices in literature.

Another important aspect of fleet coordination is the safety management. As much as it is of significance to control that the equipment are executing their tasks according to the plans with minimal deviations, the conformity with the planned operation should not be at the cost of compromise the safety of workers and operators. The valuable pose data coming from *LGS* can be used to develop a safety mechanism that can monitor and improve the safety of earthwork sites. Section 2.5 reviews a wide spectrum of literature on safety management of earthwork equipment and sites.

Finally, the effective coordination of a fleet of *LGS*-enabled equipment requires analysis and communication of a large amount of data and complex near real-time decision making. *MAS* provides the basis for solving such complex problems using the agent concepts and distributed computing. Section 2.6 provides an extensive review of *MAS* and its application in construction industry.

2.2 SIMULATION OF EARTHWORK OPERATIONS

Simulation, in etymological sense of the word, means creating conditions similar to the real world in order to establish a framework for investigation and analysis (Halpin and Riggs 1992). For this purpose, computer simulation tools are used to create a virtual environment similar to the real context of an activity, mainly to identify the shortcomings and deficiencies of a system before the actual implementation. In the case of the construction industry, simulation tools are used to represent the overall logic of various activities required to construct a facility, the required resources and the project's overall environment (AbouRizk 2010).

Although traditional scheduling techniques, e.g., Critical Path Method (CPM) (Kelley and Walker 1959) and Program Evaluation and Review Technique (PERT) (Boulanger 1961), are

still the predominant methods used for the scheduling and planning of construction projects, they are not effective in capturing the complexity and dynamism of construction processes (Halpin and Riggs 1992). This is attributable to the fact that these methods are not able to: (1) consider the active role of resources, their availability and their simultaneous engagement in multiple activities, (2) capture the intricate resource-activity interactions, and (3) account for the uncertainties and randomness that reign over many of the construction activities (Shi 1999).

Simulation, on the other hand, endows considerable power to construction managers through enabling them not only to evaluate a construction process beforehand and optimize its productivity through resource management and work task modifications, but also to develop several alternative methods for the realization of a single goal and to objectively compare different alternatives. The necessity of computer simulation for construction operations emanates from the complexity of the interaction among units and the inability of classic mathematical models to address this complexity effectively (Zayed and Halpin 2001).

According to Halpin and Martinez (1999), simulation can improve the productivity by 30% to 200%, which can correspond to a saving of $2,000 for every hour of analyst time. The general benefits of applying simulation in construction include: (1) development of better project plans, (2) optimization of resource usage, (3) minimization of project cost and duration, and (4) improvement of overall construction project management (AbouRizk 2010).

Despite the effectiveness of computer simulation, when taking the effort and investment required for the development of a computer simulation, which consists of model development, data gathering and synthesis, verification, validation, analysis and presentation, not all projects can be effectively addressed by simulation tools. AbouRizk (2010) and Martinez (2010) specify the main characteristics of projects for which the application of simulation is proven effective as

follows: (1) high degree of uncertainty, (2) technical, logistical and methodical complexity, (3) cyclic and repetitive processes, (4) need flexibility in the modeling logic (need for extensibility and scalability), (5) need integrated solution, and (6) need high detail and accuracy.

There are various types of computer simulation techniques that are employed to emulate construction processes, such as Monte Carlo simulation, Discrete-Event Simulation (DES), System Dynamics (SD), and Agent-based Simulation (ABS). The selection of the type of simulation model is based on the purpose of the analysis.

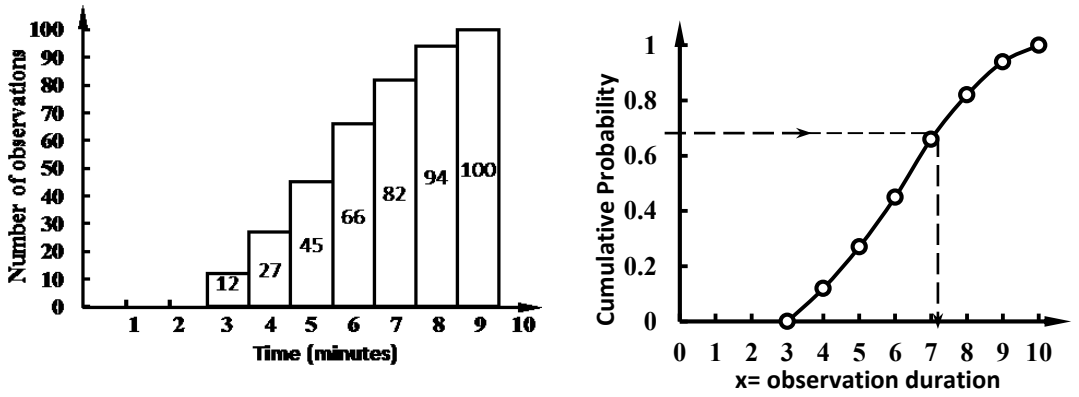
2.2.1 DIFFERENT TYPES OF SIMULATION IN CONSTRUCTION

Monte Carlo simulation is built on random number generation and its integration with the Probability Density Function (PDF) of random variables. This simulation method is used to model stochastic processes by capturing the uncertainties associated with the value of a variable of interest. Assuming that the value of a variable is represented by a PDF, Monte Carlo simulation can be used to sample the random distributions to generate a random value for that variable (Halpin and Riggs 1992). Figure 2-1 depicts the histogram used as the basis for the PDF and the Monte Carlo method for finding the value of a variable associated with a random probability value. In construction industry, this method is used in conjunction with CMP to calculate the total duration of a project in different levels of confidence.

2.2.1.1 DISCRETE EVENT SIMULATION (DES)

DES is a simulation model that captures the real-world process as a coherent set of precedence and mathematical relationships among the elements of the system (Fishman 2001). In other words, *DES* represents a process as a set of time-sequenced activities that interact with a series of resources that will be employed to transform the process's inputs to the outputs. Given that activities are defined in terms of their starting and end events, during which period the engaging

resources in an activity are captured, *DES* is only concerned with the discrete points in time when resources are captured in, and released from, different activities (Halpin and Riggs 1992). The simulation runs through the process and calculates the total duration of the process, considering also the situation where resources are retained in a queue. At the core of *DES* modeling is the Monte Carlo simulation, which is used to calculate the stochastic durations of activities. However, it is also possible to have deterministic durations for activities. Figure 2-2 shows an example of *DES* model for a simple earthwork operation.



(a) Cumulative histogram plot *(b) Cumulative probability density function*

Figure 2-1: Example of Monte Carlo Simulation (Halpin and Riggs 1992)

Halpin and Martinez (1999) investigated the application of *DES* model to simulate the construction of a variety of construction projects, including tunnels, maritime projects, dams, highways, etc., using a CYCLONE-based system called PROSIDYC. The implementation of this system on over 100 processes within 30 different projects has resulted in reduced project delivery time and cost savings of over $10 million.

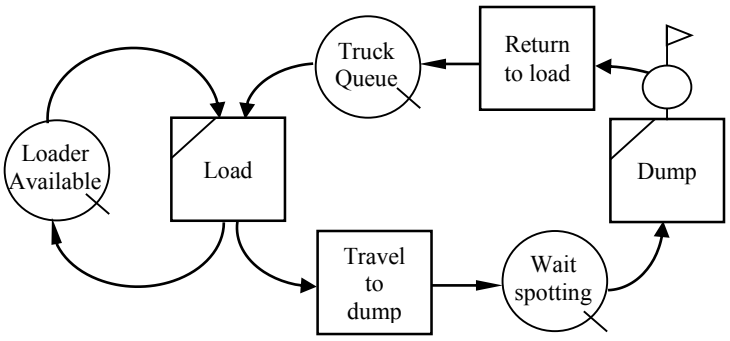


Figure 2-2: Example of DES Model (Halpin and Riggs 1992)

Zayed and Halpin (2001) implemented *DES* for a batch-plant transit mix delivery operation using MicroCYCLONE in pursuit of the following objectives: (1) calculating the optimum number of trucks corresponding to the various distances; (2) determining the optimum supply areas around the batch plant; and (3) developing decision-making tools for concrete batch plant management. Also, the cost of resources was considered to identify the optimum resource combination that minimizes cost per unit. In another effort, Hassan and Guber (2008) used *DES* modeling to optimize the placement of Continuously Reinforced Concrete Pavement (CRCP) in the reconstruction of Interstate-74, using the simplified and graphical version of STROBOSCOPE called EZStrobe (Martinez 2001). Factors considered in the decision-making to ensure that a comprehensive and holistic view of the operation is safeguarded include total operation time, productivity, costs of the operation, average truck delay, and idle times for the paver and the spreader.

2.2.1.2 SYSTEM DYNAMICS (SD)

SD is another simulation technique that is designed to assist in the long-term strategic decision-making through capturing the complex chain of causality between different elements that determine the behavior of a system (Sweetser 1999; Georgiadis et al. 2005). In the context of a *SD* model, a system is represented by a set of stocks (e.g., material, knowledge, people, etc.)

flows between the stocks, and causal loops that bridge the interaction between the flow rates and the factors that affect the behavior of the system (Borshchev and Filippov 2004). This configuration enables to capture the dynamic and complex interaction between a system and the endogenous and exogenous factors that affect the system. Figure 2-3 shows a typical *SD* model of a system in which the adoption of a new technology is simulated in terms of transition from potential adopters to adopters.

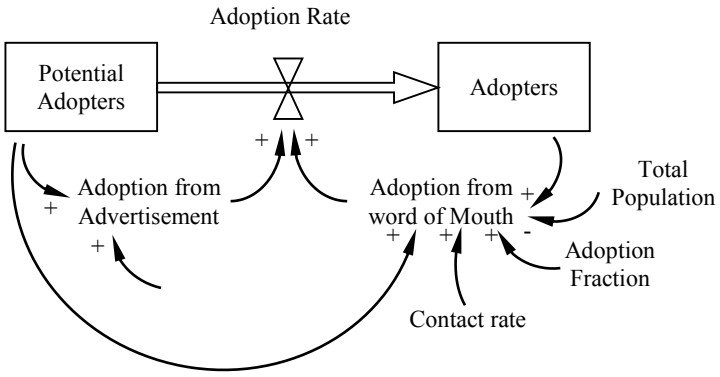


Figure 2-3: Example of SD Model (Sterman 2000)

Regarding the application of *SD* in the construction industry, Ogunlana et al. (2003) investigated the major strategic decisions and policies that a construction organization in a developing country can adopt to improve performance and prosper. The *SD* simulation was used to capture the impact of financial resources, resource management and quality of human resources on the performance of a construction company in terms of productivity, motivation, projects' scope, and direct/indirect costs. Hassan and Al-Hussein (2010) used *SD* simulation to draw a comparison between the operations of single-jib and double-jib cranes in a pre-cast residential building. Alzraiee et al. (2012) developed a simulation model for earthwork operations using *SD*. A model for a simple loading-hauling-dumping operation was developed, considering the scope change

during excavation and the required rework when the work is rejected in the quality assurance phase.

2.2.1.3 AGENT-BASED SIMULATION (ABS)

A more recent method of simulation has emerged based on the concept of agent-based modeling. In this concept, agents are defined as user-assisting software modules with cognitive abilities, e.g., (quasi-)autonomy, perception, reasoning, assessing, understanding, learning, goal processing, and goal-directed knowledge processing, that allow them to observe, sense and affect their environment (Ören et al. 2000). The most conspicuous characteristic of Agent-Based Simulation (ABS) is the decentralized and disaggregative modeling approach. Unlike *SD*, and to a lesser extent *DES*, *ABS* is a bottom-up approach of modeling in which the behavior of the real world is defined at the individual level instead of global/process level, making it a suitable choice when there is an absence of knowledge about the global interdependencies (Borshchev and Filippov 2004). In other words, *ABS* starts with the modeling of the behavioral pattern of individual elements of a system in isolation from the environment, and proceeds with the creation of an environment where different agents interact to replicate the real world (Sawhney et al. 2003). Agents are modeled as autonomous decision-making entities transiting between several states in their lifetime. Figure 2-4, shows an instance of the statechart used to graphically represent the agent's states, the transitions between states, events leading to the transitions, timing, and agent's actions.

Sawhney et al. (2003) reviewed the evolution of simulation in the construction industry and concluded that *ABS* has a great potential for applications in research concerning the emergent behavior of a construction project and the implications of human factors on issues such as safety and construction supply chain. As a proof of concept, a case study was conducted in which

different probabilities for the mal-performance of different crews, which will result in repair work, were considered to investigate the importance of the quality of work executed by different crews in reducing the overall cost and time in a house building project.

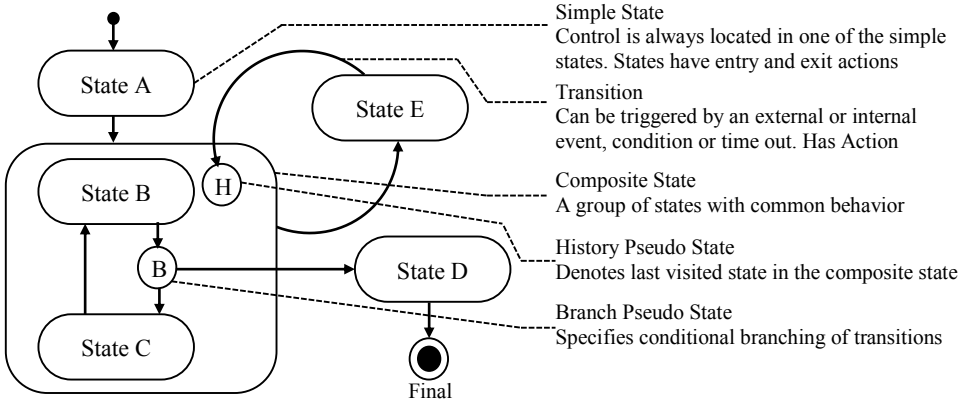


Figure 2-4: Example of Statechart in an ABS Model (Adapted from Borshchev and Filippov 2004)

Rojas and Mukherjee (2006) conceptualized a general purpose multi-agent framework in order to establish a situational simulation for the educational purposes. This framework allows students to interact with a set of agents that can autonomously change the environment based on the combined contents of: (1) knowledgebases, i.e., domain-specific rules, (2) databases, i.e., project data, and (3) user feedbacks. This composition enables agents to trigger events that are either the results of the user decisions, e.g., the need for rework as a result of the lower quality performance caused by the user's decision to impose overtime work policy on the labor, or the environmental factors, e.g., bad weather or a labor strike. This framework creates an educational platform for students to observe the consequences of their decisions in a construction project.

Although not used principally as a simulation platform, an agent-based framework was developed by Van Tol and AbouRizk (2006) to evaluate the simulation model. The main function of the developed framework is to partially relieve the modeler of the burden of utilizing

a large personal knowledge base for the analysis and verification of the simulation model. In this framework, agents use the belief network, which encapsulates a set of correlations between simulation variables and potential inferences/actions therefrom, to continuously observe the simulation model and adjust the model so that desired conditions are met.

Marzouk and Ali (2013) developed a multi-agent simulation framework for a piling operations, which considers a variety of spatial, operational and safety constraints to better replicate the actual site condition. Constraints include the minimum radius around a freshly bored pile within which no new boring can be scheduled until certain duration is passed and the avoidance of the spatial overlap between cranes and riggers.

Among the reviewed simulation techniques, *DES* better suits the scope of this research for the following reasons: (1) With the focus of the research being on earthwork projects, the analysis is performed at the process level and thus the analysis at the higher degrees of abstraction or strategic level does not provide the level of detail required for the operational planning and optimization; (2) While the integration of Monte Carlo simulation and *CPM* allows the probabilistic calculation of the project duration, the method is dismissive of the role of resource availabilities in the analysis, and thus does not provide any information about the performances of different types of equipment; (3) construction operations are highly process-oriented and thus the simulation model needs to give special attention to the sequences of the activities rather than the behavior of the equipment/operators, as is the case in *ABS*; and (4) there is a relatively rich know-how and knowledge in the literature on how to accurately model construction operations in *DES* models, insofar as some tailor-made *DES* platforms are developed for exclusive application in the construction industry, e.g., CYCLONE (Halpin 1977) and STROBOSCOPE (Martinez

1996), an asset which is ostensibly lacking for the case of *SD* and ABS. This makes the application of *DES* very convenient and effective.

2.2.2 CONSTRUCTION-SPECIFIC DES PLATFORMS

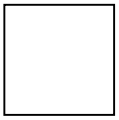
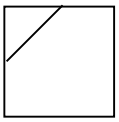
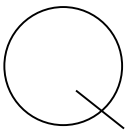
According to Martinez and Ioannou (1999), several modeling paradigms exist in *DES* developments, most notable of which are Process Interaction (PI) and Activity Scanning (AS). The major distinction between the two methods is that while *PI* models focus on the flow of entities through the system from the viewpoint of entities, *AS* models focus on identifying different activities and their starting conditions. In spite of the fact that all construction related operations can be simulated with both models, *AS* models are more suitable on account of being better able to handle operations where several complex logical conditions and intricate resource interactions exist. An advanced variation of *AS* modeling called three-phase *AS* combines *AS* paradigm with Event-Scheduling (EV) method.

Several construction-specific simulation platforms have been developed, including CYCLONE (Halpin 1977), INSIGHT (Kalk 1978), MicroCYCLONE (Lluch and Halpin 1982), RESQUE (Chang 1987), STROBOSCOPE (Martinez 1996), SIMPHONY (Hajjar and AbouRizk 1999), etc., mostly based on *AS* or Three-phase *AS* paradigm. A comprehensive elaboration of different tools falls outside the scope of this research, and only CYCLONE architecture, as the basis for most of the subsequent simulation platforms, will be discussed.

Halpin and Riggs (1992) describe CYCLONE (CYCLic Operations NEtwork) as a modeling system that uses Activity Cycle Diagram (ACD) (Martinez and Ioannou 1999) to represent a cyclic operation. *ACD* in CYCLONE comprises four main elements namely, Normal, Combi, Queue, and Arrows, which indicate the flow direction of resources amid the network, as shown in Table 2-2.

For the modeling purposes, a cyclic operation needs to be initially decomposed into (1) distinctly-discretized time-sequenced activities or tasks, (2) resource units that perform those activities, and (3) resource unit flow routes. The flow of resources is represented by the transition of resource units between two possible states, namely, active and idle. Active state implies the engagement of a resource unit in an activity. On the other hand, when resource units are retained behind an activity because the starting condition of that activity is not satisfied, they are in idle state.

Table 2-1: CYCLONE Elements Used to Build the Simulation Models (Halpin and Riggs 1992)

Symbol	Name	Definition
	NORMAL	The normal work task modeling element, which is unconstrained in its starting logic and indicates active processing of (or by) resource entities.
	COMBI	The constrained work task modeling element, which is logically constrained in its starting logic, otherwise similar to the normal work task modeling element.
	QUEUE	The idle state of a resource entity symbolically representing a queuing up or waiting for use of passive state of resources.
	ARROW	The resource entity directional flow modeling element.

Activities are time-consuming tasks that are represented by NORMAL or COMBI elements, with the difference being that the earlier denotes unconstrained work task that does not have a starting condition, e.g., combination of multiple resources, while the latter indicates a constrained work task that requires a starting condition to be satisfied. The duration of NORMAL or COMBI elements can be defined deterministically or stochastically depending on the nature of tasks they

represent. Resource units are modeled by QUEUE symbols, which basically act as storage for resource units that are in idle state. With this formalism, a QUEUE can only be used where there is a chance of resource units transiting to idle state due to the still-to-be-satisfied starting condition of the succeeding activity; in other words, a QUEUE can only precedes a COMBI.

In addition to the above mentioned principal elements, CYCLONE supports an additional element of FUNCTION, which can be diversified into multipurpose FUNCTIONs and special-purpose COUNTERs. The latter can be used to count the number of time a resource unit passes a designated point in the network. This information can then be employed to calculate the productivity based on the quantity of the counter divided by the unit of time. FUNCTIONs, on the other hand, are used to emulate CONSOLIDATE and GENERATE functions. CONSOLIDATE is used when a multiple number of a resource unit need to be accumulated before the succeeding activity can take place. On the contrary, when a resource unit needs to be disintegrated or divided into smaller portions that fits one run of the succeeding activity, the GENERATE function should be used.

2.2.3 NEAR REAL-TIME SIMULATION IN CONSTRUCTION

As stated in Section 1.2, conventional simulation approach uses the historical data as the main input for the analysis of future scenarios. While this approach is efficient for the representation of fairly stable and recurrent phenomena, it can be suboptimal when applied to the analysis of construction activities, where site-and-project specific parameters introduce a degree of uniqueness to each activity. *NRTS* approach emerged from the effort to address the unreliability of historic data in capturing unstructured and highly volatile phenomena. Many researchers studied the applicability of *NRTS* for construction industry (Lu et al. 2007; Akhavian and Behzadan 2012; Song and Eldin 2012; Pradhananga and Teizer 2013; Akhavian and Behzadan

2013). The effort was made to capture real-time data from the ongoing projects and use this data to continuously fine-tune the simulation model and thus improve the accuracy and realism of the simulation-based analysis. The term "near" in *NRTS* implies the time lag required for the simulation to capture enough data from the site so that the captured data can be considered as a generalizable pattern based on which the simulation can be updated. The window of time needed for this purpose depends on the purpose for which the simulation is being used, e.g., safety or planning and control, and project-specific characteristics, e.g., the length of one full cycle of the operation, etc.

Central to *NRTS* methodology is the automated site data acquisition that uses a continuum of *RTLS* and other data capturing tools along with data communication and transmission methods to capture the project state and compare it with the project plan to generate the progress report of the project.

A *NRTS* system called *HKCONSIM*-Real-time was developed by Lu et al. (2007) in which the vehicle tracking system, discrete-event simulation and evolutionary optimization are integrated to optimize the productivity of the ready-mixed concrete operations. In this system, a set of control and location data are integrated to identify the state of the truck mixer and optimize a set of variables including time interval for dispatching concrete deliveries for each site and the configuration of the truck mixer fleet.

Song and Eldin (2012) described a method of *NRTS* for heavy construction operations. Real-time data and process knowledge are coupled to enable a self-adaptive modeling process that validates and refines a process simulation model, as shown Figure 2-5. This method is presented at a high level of abstraction without providing a detailed description of the method. Furthermore, although a good basis for the identification of different states of the equipment is provided, the

fact that different pieces of equipment are isolated as independent units renders the state identification prone to errors and inefficient in capturing more detailed states of the equipment.

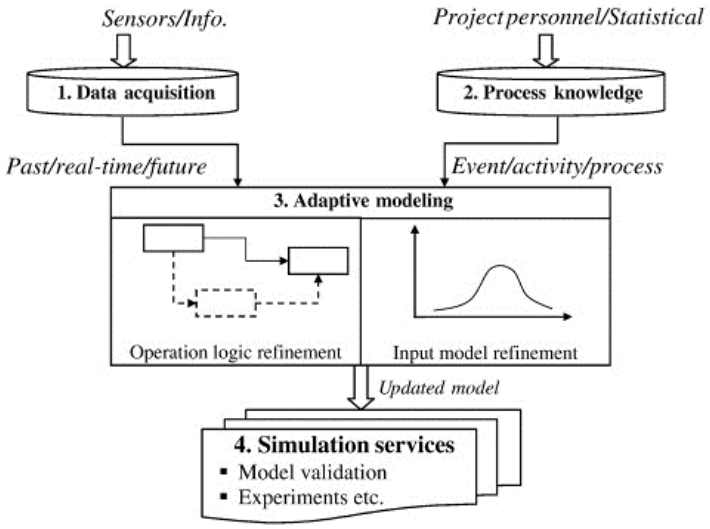


Figure 2-5: Adaptive Real-Time Tracking and Simulation (Song and Eldin 2012)

Another *NRTS* method is developed based on the collection of 3D orientation tracker data by Akhavian and Behzadan (2012). This method pursues a dual objective of updating the simulation model and creating the live 3D animated scene of the operations. This method was further improved through the addition of location and weight data for the state identification of a truck. Also, K-means clustering is used to automatically identify the dynamically changing dumping and loading zones (Akhavian and Behzadan 2013). Nevertheless, although this method concurrently considers several pieces of equipment for the state-identification, only the trucks have been the subject of investigation. Additionally, the assumption that every discrepancy with the estimates requires model update does not consider the environmental factors that might transiently affect the operation cycle time.

Pradhananga and Teizer (2013) suggested the application of *GPS* and the concept of work zones for the capture and analysis of the cycle time information. They also proposed a speed-based

method for the automatic zone detection for trucks. However, this method is inefficient in congested roads and also cannot be applied to other equipment such as excavators.

Szczesny et al. (2013) proposed a high-level conceptual method for the integration of real-time logistic data into a *DES* model using alpha-cut technique in order to capture uncertainties in the scheduling. The new real-time data are injected to the schedule as a newly defined constrain in a constraint-based modeling approach. This method is applicable to the high-level of integration where the possibility of delay in a project is calculated through defining new scheduling constraints based on real-time data. On this ground, this study does not provide an insight on how to capture the operation-level data to study the delay of operations based on the productivity of different pieces of equipment.

As mentioned in Section 1.2, to the best of the author's knowledge, the previously proposed methods are tightly-coupled with the availability of certain types of information or the application of certain types of technologies. Additionally, both the scope and level of data capturing is limited to one type of equipment, i.e., trucks, and high-level equipment states, e.g., hauling and dumping. Similarly, those proposed methods are dependent on smooth and uninterrupted flow of operational events and, therefore, do not consider the breaks in the expected chain of events that may be caused, for instance, by accidents. In the wake of this assumption, in most cases, different pieces of construction equipment are treated in isolation from the reset of the fleet, even if they are working conjunctively, resulting in the oversimplification of the construction dynamism. Another notion, common amid the reviewed methods, is that any discrepancies between the parameters of the initial model and the observed operation require the model update. This assumption tends to ignore the transient nature of some

of the discrepancies that may have been caused by temporary environmental factors, e.g., bad weather.

2.3 PRODUCTIVITY IMPROVEMENT USING AUTOMATION AND LGS

Equipment automation has been researched and investigated by several researchers as a stepping stone to improved productivity of construction operations (Russell and Kim 2003). Automation development in the construction is not solely fueled by the need for the increased productivity but is also approached as a response to poor safety and shortage of workforce (Bonchis et al. 2011). This line of research is founded on the application of emerging technologies in tracking, surveying, robotics, and machine control.

Tatum and Funke (1988) drew on the development of laser technology and its applications in the construction industry and delineated a laser-guided grading mechanism. Such mechanism is claimed to result in an expedited grading, greater quality control, efficient material usage, shorter weather-dependent time window, less equipment hours and all together increased productivity and reduced costs. In virtue of applying this mechanism, a 170,000 ft^2 increase in the average daily production and a \$0.06/ft^2 reduction in cost were reported. They also enumerated the shortcomings of the existing laser-guided grading systems to outline the future areas of development.

A system called Hazama Intelligent Vehicle Automatic Control System (HIVACS) was developed by Saito et al. (1995) for a semi-autonomous dump truck with the adaptability to the changing surroundings, such as long distance driving, high-speed driving, driving within road width, change of route, etc. The *HIVACS* uses a combination of fiber optic gyroscope, rotary encoder sensors, and laser transmitters/receivers to drive the position and direction of the truck.

Additionally, a hybrid system incorporating laser radar and image processing is used for the detection of obstacles. This system was reported to have resulted in 17% labor saving in total workers at the dam construction where the case study was carried out.

Several studies have been conducted on the autonomous excavation in Carnegie Mellon University (Singh and Cannon 1998; Stentz et al. 1999) and Lancaster University (Bradley and Seward 1998). The autonomous excavation requires several layers of analysis including: (1) strategic planning of the sequence of dig regions which includes the definition of gross vehicle movement, vehicle positioning and the excavation time based on the geometry of the site and the goal configuration of the terrain, as shown in Figure 2-6(a); (2) tactical planning of the best configuration in every dig region considering the soil-tool interactions, as shown in Figure 2-6 (b); and (3) the execution of the dig (Singh 1995). In other words, the strategic planning determines the position of the excavator and the tactical planning ascertains the pose of the bucket.

Singh and Cannon (1998) proposed a methodology for the multi-resolution planning of excavators that performs strategic and tactical planning and executes the dig using two onboard range sensors that can generate terrain map. The criteria for the strategic planning are defined based on a set of expert rules, which for instance favor starting from upper surfaces, as shown in Figure 2-6(a). Tactical planning is done by the determination of the distance along the radial line from the excavator (d) and, the angle of attack (α), as shown in Figure 2-6(b), considering three parameters, namely volume swept, time required for the dig and the required energy. The execution is done using a closed loop control scheme. Stentz et al. (1999) extended this work by adding a mechanism to detect the location of the truck and plan the dumping point of the bed of truck, as shown in Figure 2-6(c).

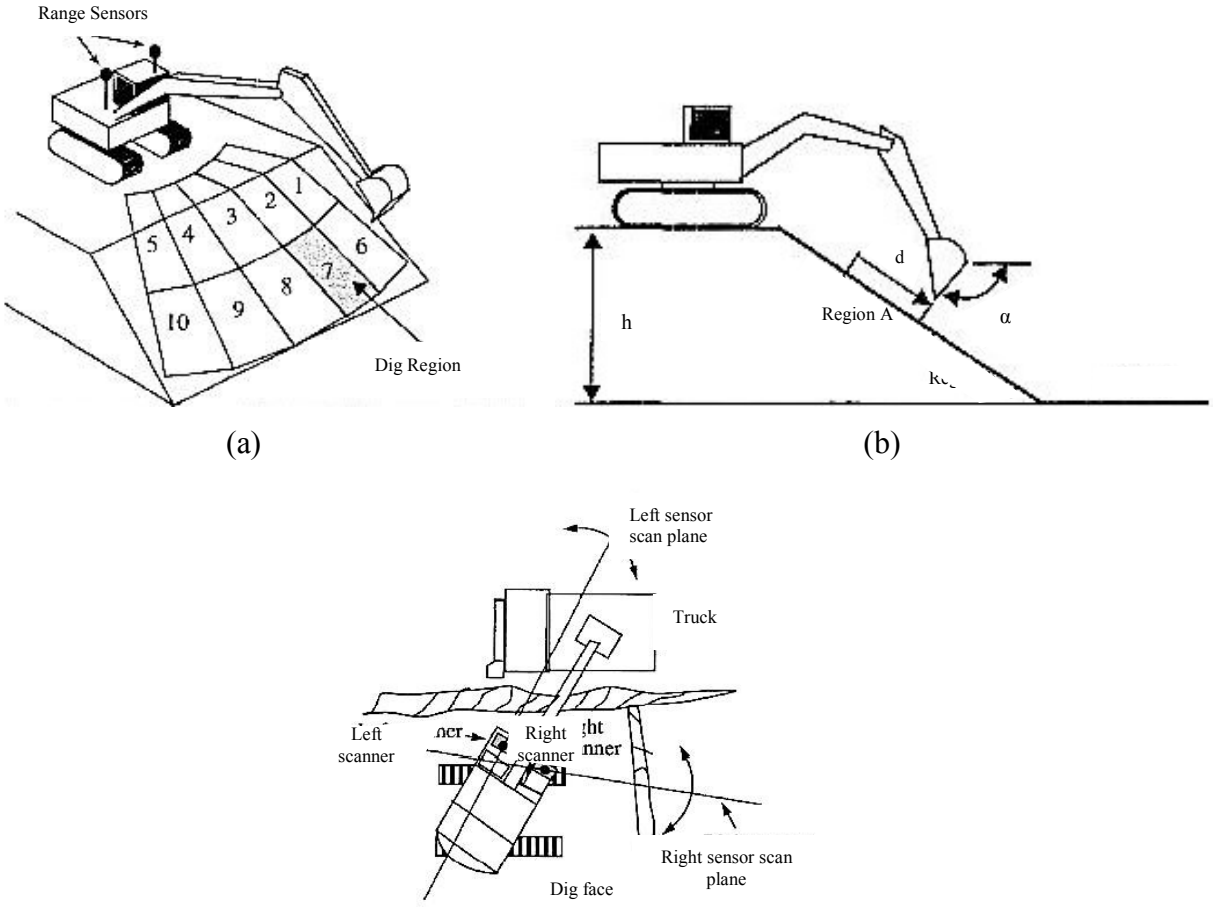


Figure 2-6: (a) Strategic, (b) Tactical, and (c) Dump Planning for an Autonomous Excavator (Adapted from Singh and Cannon 1998; Stentz et al. 1999)

However, given that most of the above-mentioned research has been conducted at the time when *GPS* and other real-time location technologies have not been as popular and available as today, these technologies have not been incorporated in the proposed systems. Furthermore, the lack of supporting tool for 3D models leads to the application of real-time scanning tools that reduced the accuracy of the results as well as adding to the computation efforts required for site scanning. The wide acceptance and application of *GPS* technology along with the emerging 3D models at the industrial level have changed the landscape of automatous equipment since the beginning of 21st century towards more practical systems that incorporate more reliable and available *GPS* technologies and use 3D models for real-time analysis.

For example, Makkonen et al. (2006) developed a methodology for the excavator control based on the integration of 3D models and two *GPS* receivers and an inclinometer. The major distinction of this work from the previous efforts on autonomous excavation is the consideration of 6 Degrees of Freedom (DOFs) for the excavator instead of the conventional 4 *DOF*s, based on the addition of two *DOF*s for Rototilt.

Also, Kaufmann and Anderegg (2008) described a methodology for an automated control of vibratory compaction machinery that analyzes the non-linear dynamics of the soil-machine system to adjust the amplitude, frequency and roller speed in the machine and its vibration-inducing mechanism. The real-time measurement of the attained stiffness and temperature is integrated with the *GPS* location data to visualize the compaction process for the operator. Figure 2-7 shows the visualization of the compaction.

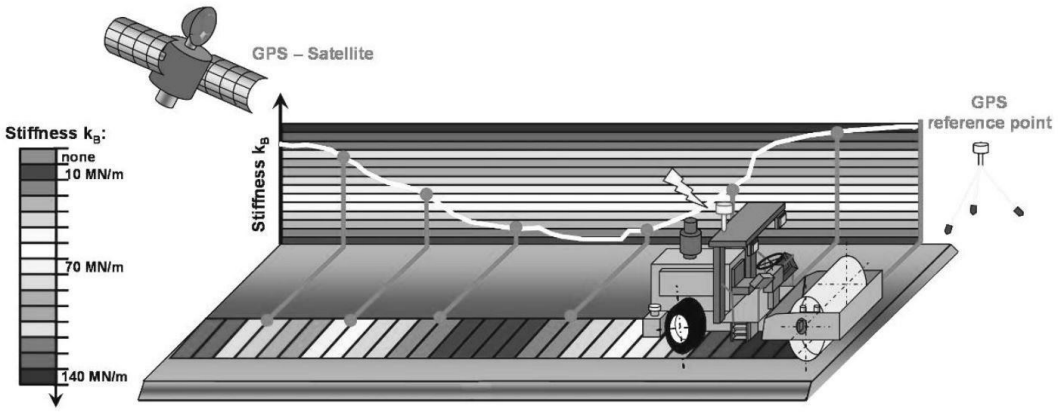


Figure 2-7: Compaction Visualization Integrating GPS Data and Real-Time Compaction Measurement (Kaufmann and Anderegg 2008)

The advent of *GPS* and the industrial adoption of the research on the provision of different levels of autonomy to the construction and mining equipment have culminated in the growing types of equipment using the *AMC/G* technology. While *AMC/G*, as a special type of *LGS* that is being

offered by major heavy equipment manufacturers, uses mainly *GPS* and total stations, *LGS* extends the range of *RTLS* that can be used to obtain the equipment location data.

2.3.1 AUTOMATED MACHINE CONTROL/GUIDANCE (AMC/G)

AMC/G integrates 3D computer models of the design, DTMs and tracking technologies (e.g., *GPS*, robotic total stations, laser systems) to increase the precision of machine operations and, in turn, improve productivity and safety (Kiongoli 2010). This integration can be used to either guide the operator for more accurate machine handling and adjustment, while granting the operator the full autonomy in controlling the machine, or take over some, e.g., blade, or all of the control units in the equipment (Jackson 2008). Viljamaa and Peltomaa (2014) developed a framework for an ontology-based information management system that builds up on the existing *AMC/G* infrastructure to establish a standard for data communication among different parties involved in earthmoving operations.

The application of *AMC/G* carries implications not only on how projects are executed in terms of the technical aspects but also on the state-of-the-practice in design, data communication, design approval, planning, scheduling and control of the project. This is due to the input requirements of *AMC/G*, i.e., 3D design models and *DTM*s, and the significant departure from the traditional design convention that the administrative regulations and organizational workflow need to be adjusted to accommodate *AMC/G*. At the administrative level, for *AMC/G* to be effectively utilized, a streamlined workflow needs to be established among the main stakeholders including the contractors, clients and designers (Vonderohe and Hintz 2010).

2.3.1.1 AMC/G COMPONENTS

AMC/G comprises three main components, namely: (1) the input data, (2) hardware, and (3) tracking technologies (Singh 2010).

Input Data

The two main inputs of *AMC/G* are the *DTM* and 3D computer models of the design. DTM represents the natural surface on which the construction is to be performed and can be used to generate topographical contour maps, surface modeling, volume computation, and engineering design work (Acharya et al. 2000). One of the common forms of the computerized representation of a surface is the Triangulated Irregular Network (TIN) which is a network of space-filling non-overlapping triangles constituted of points with known coordinates (Kumler 1994). *TIN* requires a large amount of input data that consist of the coordinates of spot elevations and breaklines. While spot elevations are representing the locations of significance such as peaks and pits, the breaklines are a set of line segments that represent significant changes in the slope. Several methods can be used to collect the input data depending on the magnitude of the project, ranging from the conventional surveying techniques to Light Detection and Ranging (LIDAR) (Vonderohe and Hintz 2010). In *LIDAR*-based mapping, laser pulses from a low-altitude aircraft, or a ground based station, are used to transmit and receive electromagnetic radiations. The analysis of the reflected light based on the time of flight and the speed of light allows the calculation of the vertical distance from the station to the ground (Uddin 2002). This information in conjunction with the *GPS* data and the orientation data, which are obtained from an Inertial Navigation System (INS) installed on the station, help generate the point cloud of the area that can be in turn used for the development of the *TIN* (Veneziano et al. 2002). Figure 2-8 illustrates the aerial *LIDAR* scanning method. Also, an example of *DTM* is shown in Figure 2-9.

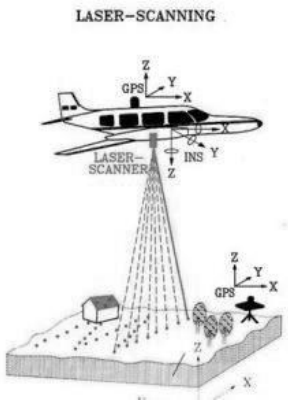


Figure 2-8: Lidar Scanning (Veneziano et al. 2002)

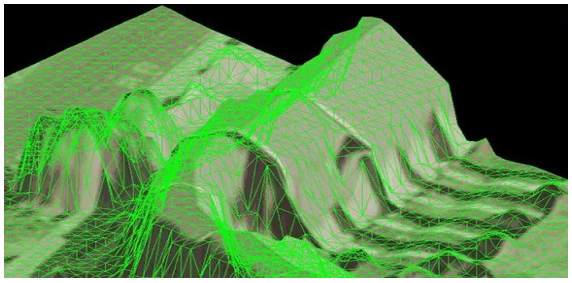


Figure 2-9: Example of DTM (Stanfords Business Mapping 2015)

A 3D model is a graphical file that, at a minimum, contains a representation of the design surface and an alignment tied to the project coordinate system (Vonderohe et al. 2009). Other types of information that a 3D model can contain include existing surface, centerlines, shoulder breaks, ditch line, etc. (Vonderohe 2009).

Hardware

The number and types of hardware used on an equipment to make it *AMC/G* compatible depends on the type of tracking technology, e.g., *GPS*, total station, etc., the type of equipment, e.g., dozer, excavator, etc., and the level of automation offered, i.e., guidance or control. However, typically, *AMC/G* requires the following basic hardware: (1) sensors, (2) data transport devices (cables or wireless transmitters), (3) control devices, e.g., on-board Central Processing Unit

(CPU), (4) displays, (5) actuators, and (6) working tool (Petschko 2008). Figure 2-10 depicts a set of *AMC/G* hardware.

(a) GPS Antnenna

(b) Total station Prism

(c) Cabin display

(d) Control device

Figure 2-10: Different AMC/G Hardware (Leica Geosystems 2015)

Sensors, essentially, comprises *GPS* antenna(s) or total station prism(s), depending on the type of tracking technology. Additional sensors can be deployed to avoid adding more expensive *GPS* antennas or total station prisms for the capturing of all the *DOF*s. Tilt sensors and Inertial Measuring Units (IMUs), for instance, are two types of sensors that can be used to capture the equipment *DOF*s.

Data transport devices are used to transfer measurement data from sensors to the control device. Bus systems, e.g., Serial, Content Addressable Network (CAN) and Ethernet, are usually used for this purpose. The control device compares the sensory data with the design data and finds the right machine action. Depending on the level of automation offered, i.e., guidance or control, either a display is used to provide the operator with the relevant information or commands are sent to the actuators, e.g., motors, valves, etc., for actions (Petschko 2008).

Furthermore, in more sophisticated systems that offer fleet management capabilities, communication tools can be used to enable Machine-to-Machine (M2M), Machine-to-Office (M2O) and Office-to-Machine (O2M) communications and allow sharing of different types of data, such as machine condition, location, productivity, work quality, and 3D design model updates. Mesh radio networks in a standard Wi-Fi environment is an example of methods used for the data transport and sharing (Sturm and Vos 2008), as shown in Figure 2-11.

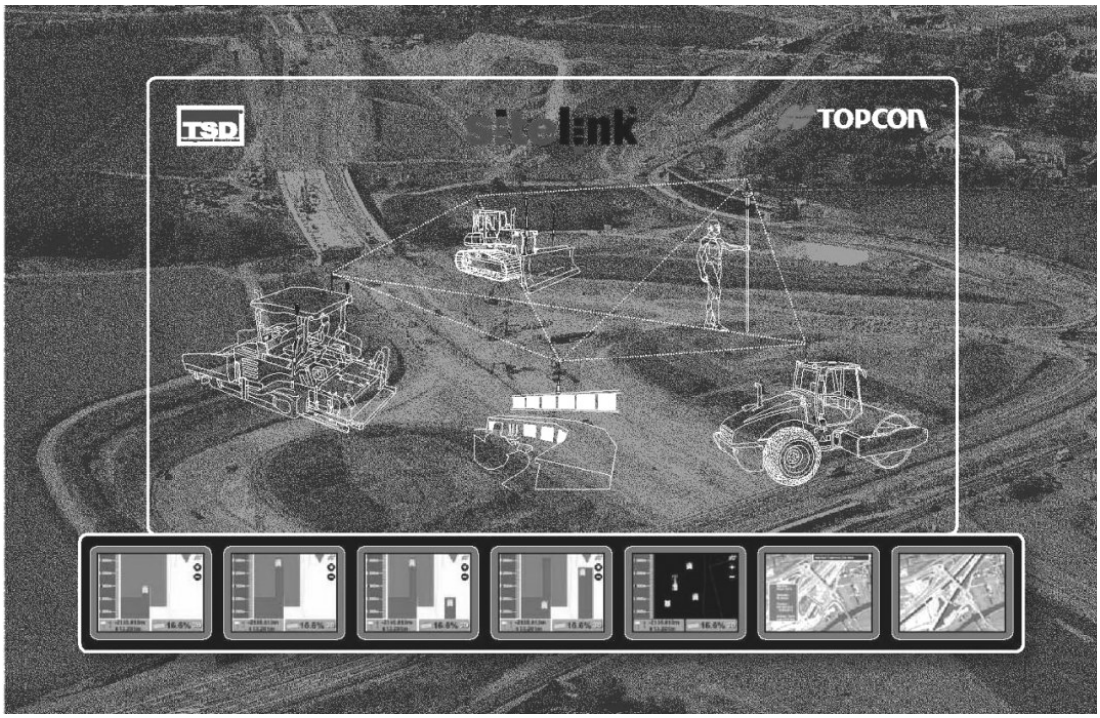


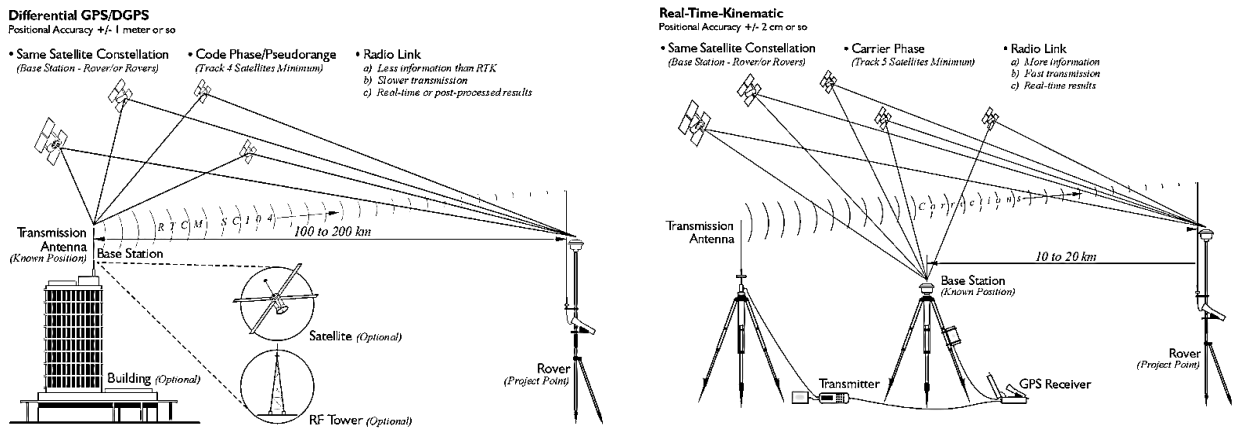
Figure 2-11 Example of Communication Network between Multiple Pieces of Equipment (Sturm and Vos 2008)

Tracking Technologies

Several types of technologies can be used as the tracking component of *AMC/G*. When selecting the type of tracking for *AMC/G*, multiple factors have to be considered, including the required level of accuracy, site layout, visibility, compatibility with the fleet, etc. Most common types of tracking technologies used for *AMC/G* include different variations of GPS and robotic total station.

(i) GPS

Given that un-augmented *GPS* has the accuracy of around 10-30 meters without corrections and 2-5 meters with some corrections, it is required to use enhanced *GPS* that can provide sub-meter accuracy for *AMC/G* (Vonderohe et al. 2009). As shown in Figure 2-12(a), Differential *GPS* (DGPS) uses the ground-based reference stations, whose coordinates are known, to apply corrections to the coordinates of *GPS* rovers and can improve the accuracy to below 1 m (Gan-Mor et al. 2007). Real-time Kinematic (RTK), also referred to as Carrier Phase Enhancement *GPS* (CPGPS), uses also reference receiver with a known location and a transmission antenna, which bridges communication between rovers and reference receiver, to correct the coordinate of *GPS* rovers, as shown in Figure 2-12(b). The main difference between DGPS and RTK *GPS* is that the earlier bases the calculation on the analysis of the pseudorange while the latter uses carrier phase for this purpose.



(a) (b)

Figure 2-12: (a) DGPS and (b) RTK GPS Configurations (Van Sickle 2008)

As a part of RTK *GPS* utilization for *AMC/G*, a process called site calibration or localization needs to be carried out. This process involves the rovers visiting a set of geodetic control points whose coordinates have been previously measured during the initial surveying. The juxtaposition

of the surveying coordinates and the coordinates measured by the *GPS* allows defining a set of parameters that transform *GPS* reference frame into local reference frame. Once determined, these parameters will be embedded in all *GPS* rovers, allowing them to perform the coordinate transformation in real-time (Vonderohe et al. 2009). RTK can provide an accuracy of about 1 cm (Gan-Mor et al. 2007) which makes it a suitable choice for the application in *AMC/G*.

Where a high accuracy of vertical alignment is required, e.g., fine grading, laser augmented RTK *GPS* can be used to achieve a vertical accuracy of 1.5-3 mm (Bryant 2006), as shown in Figure 2-13.

(ii) Robotic Total Station

Safeguarding the millimeter level accuracy, robotic total stations can substitute *GPS* in the congested urban areas or densely wooded areas where the lack of visibility hinders the application of *GPS*. In the total station-based *AMC/G*, a set of dedicated total stations are tracking the location of prisms that are mounted on the equipment with a one-to-one relationship, as shown in Figure 2-14. The real-time location data is then sent back to the equipment for the guidance or control purposes (Singh 2010).

Figure 2-13: Laser Augmented GPS for Improved Vertical Precision (Singh 2010)

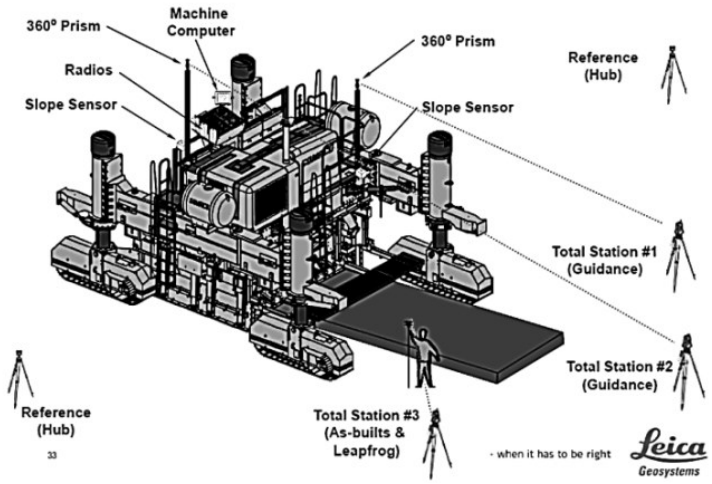


Figure 2-14: Total Station Based AMC/G Used for Paving (Jahren 2013)

Figure 2-15 illustrates a set of different *AMC/G* equipment that can be used in different operations. Several manufactures offer their *AMC/G* solutions for multiple construction and mining operations, such as Caterpillar's MineStar (Caterpillar 2015), Leica's Jigsaw (Leica 2015), Topcon (Topcon 2015) and Trimble (Trimble 2015). These solutions provide *AMC/G* capabilities at different levels from mere operator guidance to fully autonomous machine control. For instance, Caterpillar's MineStar offers a full-fledged overarching fleet management system that integrates machine health monitoring, proximity detection, M2M/M2O/O2M communications, productivity management and optimization to ensure safety and optimized productivity (Caterpillar 2015). However, due to the proprietary nature of these solutions, only very limited information is available on the structure and the underlying algorithms used in these solutions.

(iii) Other types of Sensory Data Used in AMC/G

AMC/G requires other type of sensory data, in addition to the accurate location data in order to perform control mechanisms. Subsequently, a set of dedicated sensors, e.g., engine speed sensors, engine temperature sensors, transmission shift sensors, transmission speed sensors,

pump and motor displacement sensors, steering angle sensors, angle sensors associated with the pivot joints, etc., dynamically communicate with the on-board control device to gather such data as the condition of equipment and status of various hydraulic valves (Halder and Vitale 2010; Mintah et al. 2011).

(a) Excavation (b) Grading (c) Milling

(d) Paving (f) Curb installation (e) Barrier installation

Figure 2-15: Different AMC/G Equipment in Operation (Singh 2010)

To the best of the author's knowledge, despite the wide range of functionalities, the existing *AMC/G* solutions are limited to the machine-level productivity improvement, or fleet coordination based on simple heuristics following linear and deterministic cycle-time calculation of the hauling unit, which is not sufficient to address the project-level monitoring and decision-making needed in complex projects. As mentioned in Section 1.2, in the complex setting of earthwork operations, mere optimization of machine-level work may not necessarily result in a well-coordinated management of a fleet. Another limiting fact about *AMC/G* is the cost. The cost of new *AMC/G*-enabled equipment can be as high as $80,000, which may not be very inviting for contractors. The use of other types of more affordable *RTLS* for *LGS* has not been investigated

yet. Additionally, the capabilities of *AMC/G* are solely employed to improve the productivity and are not utilized effectively to enhance the safety of the project, and the existing safety measures are limited to proximity warnings.

Therefore, a mechanism is required to (1) deploy more cost effective tracking technologies for *LGS*, and (2) harness the control and tracking power of *LGS* for designing an intelligent system that allows project-level coordination of fleet vis-à-vis all the complex interactions of different teams of equipment in a project.

2.3.1.2ADMINISTRATIVE INFRASTRUCTURE

Some governmental bodies and construction companies have started to appreciate the value of employing *AMC/G* in heavy earthwork and road construction projects. For example, Wisconsin Department of Transportation (WISDOT) has conducted research on adopting regulations for the application of *AMC/G* in highway construction projects (Vonderohe 2007; Vonderohe and Hintz 2010). Their efforts are mainly focused on the implementation strategies for regularizing 3D design models and *DTM*s in their projects.

Dunston and Montey (2009), in collaboration with Indiana Department of Transportation (INDOT), have also investigated the issues for streamlining the application of *AMC/G* systems focusing on: (1) Analyzing the data needs and provision, (2) Investigating standards, codes and permission concerns, (3) Analyzing the liability distribution and project information security, and (4) Reviewing implementation costs.

The benefits of employing *AMC/G* are manifold, although some obstacles caused mainly by the current state of the practice can be also identified. Table 2-2 summarizes the *AMC/G* perceived benefits and obstacles.

Table 2-2: Perceived Benefits and Obstacles of AMC/G Implementation (Adapted from Vonderohe and Hintz 2010; Jahren 2013)

Advantages	Obstacles
• More accurate grading and smoother ride • Avoidance of re-work • Faster operations • Significant cost reductions • Improved safety of operations • Environmental-Fuel savings • Ease of constructability review • Less dependency on the operator's expertise • Reduced surveying and staking time/effort • Reduction in re-engineering from design to construction process owing to using 3D models • Reduced traffic interruptions • Safety of the traveling public	• Lack of agency specifications • Lack of equipment • Lack of knowledge concerning benefits • Budgetary limitations • Conversion of paper plans to 3D models • File preparation to achievement of the appropriate model • Lack of competent personnel for implementation • Dependency on third-party consultants for DTM creation

Most noticeably, significant cost savings, faster operations and improved precision are identified as chief advantages of the *AMC/G* implementation. Similarly, the project could benefit from *AMC/G* through reduced dependency upon the expertise of operators, less time and effort required for staking and surveying, and less need for the design-to-construction process re-engineering. It is also reported that by virtue of faster operation, less interruption is made to the traffic. Despite these potentials, the implementation of *AMC/G* is hindered by the lack of administrative regulations and lack of equipment as well as budgetary limitations. In addition, given that the technology is still under development, there is little awareness of the benefits of *AMC/G* amid stakeholders. The difficulty of converting the plans and preparing the model for *AMC/G*, lack of competent personnel and dependency on a consultant for *DTM* creation are also reported to forestall the implementation of *AMC/G*.

2.4 PROJECT MONITORING AND CONTROL

Delays and variations are an inextricable part of construction projects (Wambeke et al. 2010; Rebolj et al. 2008), and often can be attributed to the poor estimation at the planning phase or/and unforeseen circumstances (Navon and Sacks 2007). Subsequently, regardless of the soundness of the initial plan in terms of the optimized time and cost, it is of crucial importance to constantly monitor and control the project as to ensure the smooth progress, early detection of anomalies and prompt corrective measures. Navon (2007) presented a framework for the performance control cycle which encompasses steps for measuring the project performance using Project Performance Indicators (PPI), identifying deviations based on the comparison of actual and planned performance, and devising and executing corrective measures, as shown in Figure 2-16.

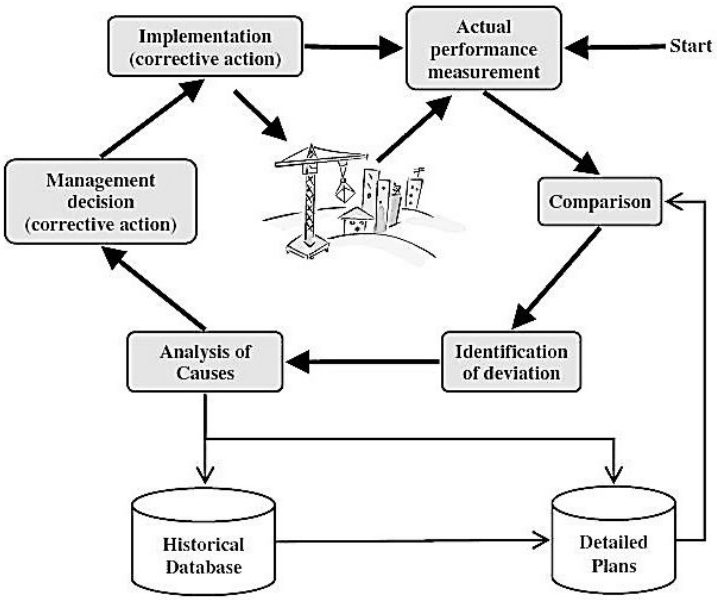


Figure 2-16: Performance Control Cycle (Navon 2007)

Conventionally, the monitoring of construction projects and the measurement of PPIs are performed using manual methods of data collection, which are error-prone, time-consuming and

costly (Navon and Sacks 2007; Azimi et al. 2011). The inadequacy of the traditional methods has led to the adoption of modern automated data capturing technologies, e.g., Radio Frequency Identification (RFID), *GPS*, Laser Detection and Ranging (LADAR), *UWB*, inertial based systems, video/audio capturing, etc., for the monitoring of construction sites (Peyret et al. 2000; Navon et al. 2004; Navon 2007; Navon and Sacks 2007; Alshibani and Moselhi 2007; Rebolj et al. 2008; Perkinson et al. 2010; Azimi et al. 2011; Golparvar-Fard et al. 2013).

RFID is an automatic identification technology to capture and transmit data using radio frequencies (Motamedi and Hammad 2009). Using three main components, namely, antenna, transceiver and transponders, *RFID* emits radio signals to read/write data from/to tags.

UWB is a wireless technology for transmitting large amounts of digital data over a wide spectrum of frequency bands at very low power (less than 0.5 milliwatts) (Ghavami et al. 2004). Owing to its wide bandwidth, *UWB* provides a high accuracy and overcome the multipath problem. The multipath problem refers to the phenomenon of a radio signals being received by more than one path, which is mainly due to the reflection of signals off various surfaces. The combination of the high accuracy, permeability and immunity to the multipath issue renders *UWB* a robust solution for localization problems (Kolodziej and Hjelm 2006).

In a *UWB* system, a set of several sensors working as a single operating unit tracks objects of interest. Of the network of sensors, one is assigned as a master sensor and is responsible to synchronize the timing of sensors. The synchronization of sensors is done using timing signals from each sensor to the master sensor. Location data are collected from *UWB* tags that are attached to objects of interest. The tag data recording by the sensors is done using Time Division Multiple Access (TDMA) method that splits the signal into several time slots and allocates each slot to one tag (Ubisense 2015). Tags' data are received in quick succession at their own time

slots. Optimization is used for the time slot allocation so that the requested quality of service is safeguarded while enough space is maintained in the schedule for new tags registration (Zhang 2010).

The signals emitted from *UWB* tags are received by all or some of the sensors in the cell, as shown in Figure 2-17. The signals received by the slave sensors are decoded and the angle of arrival and timing information are sent to the master sensor, using Ethernet or Wi-Fi connection. Trilateration is used to compute the location of tags based on the accumulation of data from all sensors (Zhang 2010; Zhang et al. 2012).

2.4.1 MONITORING AND CONTROL OF EARTHWORK PROJECTS

The monitoring and control of earthwork operations, as a sizable proportion of the large-scale construction projects, have been subject to a comprehensive research.

Peyret et al. (2000) developed two separate systems for the real-time control and monitoring of compactors and pavers. These systems are designed to fulfill the dual objective of: (1) Assisting the operators in performing their tasks through the provision of real-time data; and (2) Helping the site managers in conducting quality control and analysis via collecting the spatio-temporal data of the operations.

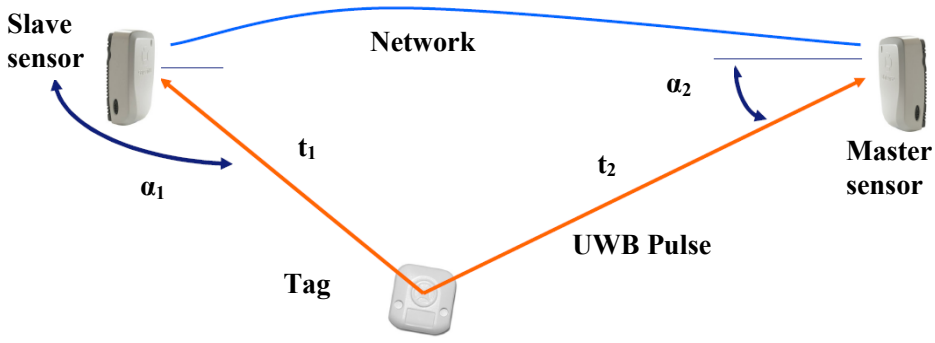


Figure 2-17: Signals Sent to Sensors Are Used to Calculate the 3D Position of a Tag (Ubisense 2015)

Navon (2007) presented several frameworks for the management and control of labor, equipment, materials, and preventive measures against falls from heights. The framework for earthwork equipment control, for instance, suggests that if the equipment's positions are tracked using such technologies as *GPS*, and the schedule, planned productivity and the geometric model of the design are known, it is possible to calculate the duration each piece of equipment spends in different work envelopes that are associated with different activities. These durations can be used to measure the actual productivity of different equipment and further compare them with the planned productivity for analysis and decision making.

A framework was proposed by Alshibani and Moselhi (2007) based on the integration of a Geographical information system (GIS) and *GPS* to collect earthwork data, calculate cycle time and estimate the project time and cost. According to this framework, the *GPS* attached to a truck transfers the data to a GIS-based system for the visualization and analysis. The calculation of cycle time is performed based on the entrance and departure of the truck to/from certain predefined activity-stamped areas, e.g., loading areas. The cost and time forecasting is also performed using the ratio technique and the earned value concept in form of PPI and Cost Performance Index (*CPI*).

UWB was proposed for automated monitoring of construction activities by Teizer et al. (2008). This research provided an overview on the general benefits of the technology as compared to other tracking methods and demonstrated that *UWB* maintains accurate enough measurements, both indoor and outdoor, for applications in the tracking of materials and work task productivity.

Perkinson et al. (2010) proposed a framework to make use of *GPS*-collected data for enhanced management and decision-making purposes. The central objective of their methodology is to

ensure that all the productivity-related data can be efficiently collected in the course of the project development and further used for planning, scheduling and control.

On the application of vision-based technology for automated earthwork monitoring and control, Rezazadeh Azar and McCabe (2011), proposed a framework to use computer-vision techniques, both in form of static image processing and video analysis, for the extraction of data about dump trucks. Such a framework is claimed to serve the real-time estimation of productivity, calculation of service/idle time and detection of accidents. Also, at the detailed level of recognizing actions instead of objects, Golparvar-Fard et al. (2013) proposed an algorithm which is capable of detecting single actions of construction equipment, resulting from articulated motions, based on the video captured form a fixed camera. The success rate for the classification of equipment action was reported to be 86.33% and 98.33% for excavators and trucks, respectively.

Montaser and Moselhi (2012) presented a *RFID* based system for the tracking of earthwork equipment. This framework is based on fixing *RFID* readers at certain locations as the gate system. The data from several readers are later fused to extract the data about the full cycle of a hauling unit in a simple loading-hauling-dumping-returning operation.

The reviewed literature on the automated monitoring of the earthwork operations indicates the significance of automated solutions to reduce cost, increase accuracy and shorten the analysis time. Nevertheless, the computer-vision based techniques are still immature in terms of (1) tracking and capturing the full-cycle of equipment travelling across a wide area, and (2) reliably and economically maintaining the functionality under climatic variations. Other tracking technologies, in the context of their proposed uses, although able to provide full coverage and seamless data fusion between various sources, are limited in the level of detail they can capture and are proposed for equipment that tend to travel over long distances, e.g., trucks. On the other

hand, *LGS* provides accurate set of data that combine a wide array of location and sensory data. These different types of data can be fused and used to capture the location and orientation of equipment in a stable manner and further calculate the cycle time of both relatively fixed equipment, e.g., excavators, and highly mobile units, e.g., trucks.

2.5 SAFETY OF EARTHWORK OPERATIONS

With only less than 5% of the U.S. work force, the construction industry claims around 20% of fatalities and injuries in workplaces (MacCollum 1995). In Britain, in addition to 25000-30000 injured, approximately 1500 people are losing their lives on construction sites in a typical decade (Davies and Tomasin 1996). Figure 2-18 illustrates the fatal injuries on construction sites between 1997 and 2007, happened to different target groups and in total, according to the Health and safety Executive (HSE) statistical data (Howarth and Watson 2009). In Canada, construction industry has accounted for more than 22% of the work-related fatalities between 2011 and 2013, which is more than any other industries (AWCBC 2013).

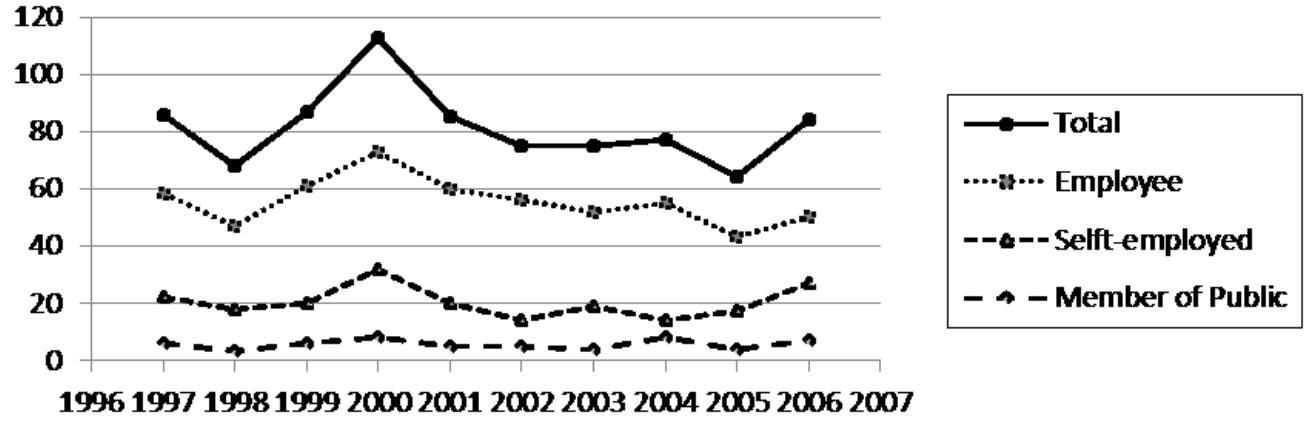


Figure 2-18: Fatal Injuries in Construction Industry in the U.K. (Adapted from Howarth and Watson 2009)

Earthwork equipment accounts for a large share of injuries on the construction site. In the U.K., of the total number of fatalities in the period of 7 years (1996-2003), 14% were identified to have been caused by being struck by a moving vehicle (Howarth and Watson 2009). According to Hinze and Teizer (2011), one-fourth of construction fatalities are due to equipment-related incidents. Equipment-related incidents are usually categorized into struck-by and caught-in/between accidents (Hinze et al. 2005). These two types of accidents differ in that while in the struck-by accidents the impact alone is the cause of injury, in caught in/between cases the injury is caused by crushing between objects (OSHA 2011). In the U.S., 428 equipment-related struck-by and caught-in/between accidents were reported between 1995 and 2008 (Wu et al. 2010). While federal level data is missing for Canada, the statistics from Occupational Health and Safety (OHS) of Alberta suggests that 28% of the reported injuries in construction industry in 2010 are caused by struck-by, struck-against, and caught-in accidents (OHS 2011).

McCann (2006) provided an extensive overview of equipment-related fatalities on excavation sites based on the BLS data. According to this research, in the period between 1992 and 2002, the causes of 30%, 24% and 12% of equipment-related deaths on the excavations sites were identified to be struck by vehicles, struck by objects and caught in-between (e.g., vehicle parts, vehicle loads, or falling vehicles), respectively. Based on the distribution of fatalities amid different types of equipment, trucks, excavators, loaders, and bulldozers collectively account for 79% of the excavation fatalities. Of the different target groups fallen victim to these fatalities, equipment operators and construction labors constitute the dominating part, together adding up to some 63% of fatalities. Different typical accidents on an earthwork operation, risk factors and their root causes have been also investigated by Chi and Caldas (2011), as shown in Table 2-3.

Table 2-3: Accident Risk Factors Identification for Typical Accidents in Earthwork Operations (Adapted from Chi and Caldas 2011)

Accident	Risk Factor	Root Cause
Rolled over Fell over (road) edge Hung up on (road) edge	Operator Error Poor Operating Condition Mechanical/Hydraulic Failure	Limited visibility to objects High operation speed Access to unstable pile, edge or ground Travel through edge or berm Close access to edge or berm Over-loading Inadequate rules or signs Narrow working area or road width Steep grades No berm or poor berm condition Undercut pile or poor ground condition
Collision	Operator Error Poor Operating Condition Mechanical/Hydraulic Failure	Limited visibility to objects High operation speed Access to unstable pile, edge or ground Inadequate rules or signs Narrow working area or road width Steep grades Poor ground condition
Bounced or Jarred	Operator Error Poor Operating Condition Mechanical/Hydraulic Failure	High operation speed Access to unstable pile, edge or ground Over-loading Narrow working area or road width Steep grades
Contacted power line	Operator Error Poor Operating Condition	Limited visibility to objects High operation speed Access to unstable ground Inadequate rules or signs Narrow working area or road width Steep grades Inadequate power line clearance
Pinned between	Operator Error Mechanical/Hydraulic Failure	Carelessness

Safety concerns in excavation are not limited to the collision with human or equipment, and are as well extended to the collision with underground utilities, which can lead to tremendous

financial damages, injuries or even fatalities. Based on the statistics published by Common Ground Alliance (CGA), a total of 232,717 of damages to underground utilities were submitted to the Damage Information Reporting Tool (DIRT) in Canada and U.S. in 2012 (CGA 2012). However, a linear regression model for 11 states, estimated the number of underground excavation damages in 2012 to be approximately 350,000, only in the U.S. Figure 2-19 illustrates the submitted and estimated damages as well as the construction spending ensued in the period between 2004 and 2012 (CGA 2012).

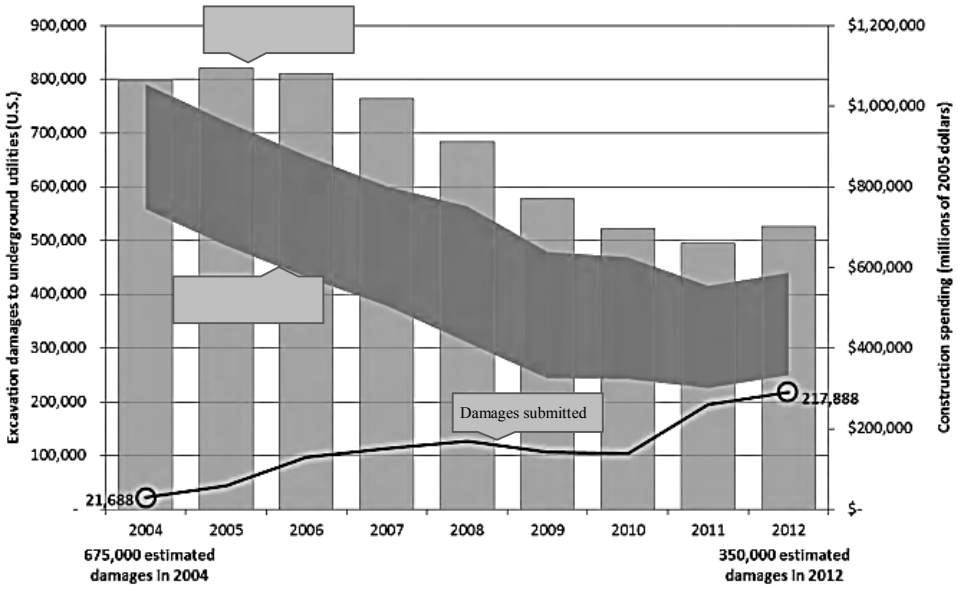


Figure 2-19: Estimated Number of Total Underground Utility Damages Resulting from Excavation in the U.S. (CGA 2012)

2.5.1 ACTIVITY WORKSPACE

The above suggests that earthwork operations are in need of enhanced safety to avoid damages, injuries and fatalities. Therefore, many researchers have investigated the safety of earthwork operations as well as equipment, exploring a wide range of solutions from sensor-based safety systems. With this need in mind, many researchers have explored a wide range of solutions to mitigate the risk of accidents on excavation sites through reducing the possibility of collisions

between equipment through a proper planning method (Chavada et al. 2002, Mallasi and Dawood 2004, Tantisevi and Akinci 2007, Hammad et al. 2007, Moon et al. 2013). These methods identify the spaces required for the safe completion of different activities, i.e., activity workspaces, and try to reduce the overlap between them.

A workspace is an area defined for the purpose of the safety and progress monitoring around locations where several cyclic actions take place. Unlike the strategic plan, the activity workspace does not require the equipment to be at a fixed location at a given time, but alternatively ascertains that the equipment is expected to be working within a certain boundary at a certain time to ensure safety and compliance with the schedule.

In a diagnostic study of the causes of lower productivity, Kaming et al. (1998) identified the work interference, resulting mainly from poor scheduling, as the most influential cause of productivity problem. Mallasi and Dawood (2001) indicated that in excess to 30% of the non-productive time could be attributed to space conflicts resulting from poor time-space planning. On this premise, the concept of workspace, which help demarcate the space that needs to be reserved for a certain work within a period of time, are being recommended for space planning. In addition to productivity related functionalities, workspace analysis can tremendously contribute to the identification and preemption of potentially hazardous circumstances through the concurrent consideration of the temporal and spatial details of various activities.

An automated method for the generation of construction activities' workspaces are proposed by Akinci et al. (2002) based on the application of the qualitative description of the location of an activities and the quantitative description of the size of workspace in an IFC-based 4D production model.

Mallasi and Dawood (2004) proposed a framework for workspace planning using 4D visualization technology. In this framework, work progress rate, the initial schedule, information about the process geometric specifications and the 3D design are integrated into a 4D model, in order to minimize the workspace congestion.

Hammad et al. (2007) proposed a method for the automated generation of workspaces around construction equipment and the analysis of workspace conflicts between various activities based on the project schedule.

Decomposing the workspaces into four main categories, namely main, support, object and safety workspaces, Chavada et al. (2012) have proposed a nD modeling approach that is committed to the identification of spatio-temporal conflicts, analysis of workspace congestion and resolution of the time-space conflicts.

Moon et al. (2013) have developed a schedule optimization framework that uses temporal and spatial information of various activities to identify the schedule with minimized simultaneous interference level of schedule-workspace. The proposed framework further uses 4D modeling for the visualization of the optimum schedule.

While most of the above-mentioned methods are devised to serve at the design stage, where designs are to be made over the constructability of the developed schedule, Setayeshgar et al. (2013) have proposed a framework based on the integration of real-time simulation with the workspace analysis. In this framework, the spatio-temporal analysis of the site is repeatedly performed in view of the site monitoring data and the results of real-time simulation in order to identify the impact of the most recent changes in the site on safety.

Despite the effectiveness of these methods in reducing the possibility of collisions between different teams of equipment at a macro level, they are not fully capable of averting safety risks emanating from human errors and unforeseen circumstances. Additionally, space is a limited resource that many of earthwork projects do not have. These methods are not able to effectively improve the safety in congested sites, given that activity workspaces may overlap in many instances.

2.5.2DYNAMIC EQUIPMENT WORKSPACE (*DEW*)

While there could be various different root causes behind the equipment-related fatalities, as shown in Table 2-3, the majority of equipment-related accidents can be avoided if the dangerous areas around the equipment is monitored in real time and the operators are warned against any intrusions into these areas. As a result, it is of a paramount importance to devise a complementary real-time mechanism to reduce safety risks based on the current pose and state of the equipment. To this end, researchers considered systems that generate warning against dangerous proximities using radar-based proximity sensors (Ruff 2006; Choe et al. 2014), vision-based tracking (Chi and Caldas 2011), and *RTLS*s such as *GPS*, *RFID*, and *UWB* (Burns 2002; Chae and Yoshida 2010; Teizer et al. 2010b; Carbonari et al. 2011; Zhang and Hammad 2012; Wu et al. 2013; Zolynski et al. 2014). These methods are applied at the monitoring phase with the intention to ensure that different pieces of equipment do not collide with one another. Similar to the methods used at the planning phase, these methods consider the space around the equipment that should not be trespassed by other equipment to avoid potential collision in the immediate future. Because these spaces are applied to equipment, as opposed to the activities, and their shapes are dynamically changing based on the current pose of the equipment, they are referred to as Dynamic Equipment Workspaces (*DEW*s) in this research. The correlation between

the two types of workspace is that an activity workspace must be the envelope that contains all the *DEWs* generated by the fleet assigned to that activity over the scheduled period. Although *DEWs* are alternatively termed in the literature as "safety envelopes" (Burns 2002; Zolynski et al. 2014) or "safety zones" (SAFEmine 2014), given the above correlation, it is preferable to use the unified term "workspaces" for both activity and equipment.

Two approaches can be found in the research addressing the generation of *DEW*s. While some researchers use only the equipment geometry and pose for the generation of *DEW*s, others also consider the speed characteristics of the equipment.

2.5.2.1 DEWS BASED ON THE EQUIPMENT GEOMETRY AND POSE

Several methods have been developed to generate *DEW*s based on the application of different types of *RTLS*s. Generally, the methods of generating *DEW*s based on the proximity measurements can be categorized into two groups. Some methods are totally independent of the pose, state and speed data; and therefore they over-conservatively reserve the space within a radius (r) of the equipment (called here cylindrical workspace, Fiure 2-20(a)). For instance, CRC Mining (2015) developed the Shovel Load Assist System that uses the combined data from a laser scanner, GPS and pulse radio to locate the trucks and dozers in the vicinity of the shovel to avoid the potential collision with them. There are many examples of cylindrical workspace in proposed systems for collision avoidance on construction site (Chae 2009; Teizer et al. 2010a; Cheng and Teizer 2013; Marks 2014; Pradhananga 2014; Luo et al. 2014). Other methods detect the shortest distance between the two pieces of equipment and use a minimum acceptable threshold for generating the warnings, which is equivalent to considering only the pose of the equipment and creating a buffer of width (b) around the equipment (called here buffer workspace, , Fiure 2-20(b)).

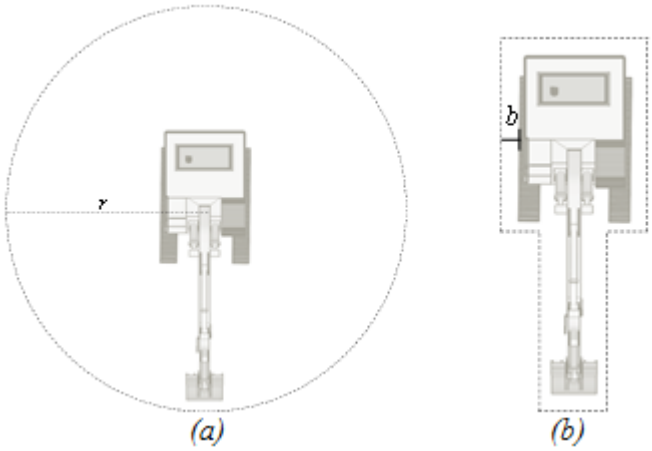


Figure 2-20: (a) Cylindrical Workspace, and (b) Buffer Workspace (the Model of Excavator Is Obtained from Google 3D Warehouse (2015))

For instance, Kim et al. (2006), proposed a method for real-time collision avoidance systems that uses laser range finders to model the obstacles on the site and then calculates the shortest distance of the equipment to various surrounding obstacles. If a threshold distance is violated, the warning is generated.

Teizer et al. (2010b) proposed a pro-active real-time system for automated warning and alerting to the workers and operators using very high frequency active radio frequency technology. In this system, workers and operators are equipped with separate units, called PPU and EPU respectively, that runs on batteries and emit signals for proximity detection. The adjustment for the safety distances are performed based on the equipment-specific blind spot measurement. The system alerts the operators and workers if the safety distance is violated.

A similar RFID-based system was developed by Chae and Yoshida (2010) to automatically estimate the working areas, analyze the collision risk and warn concerned operators and workers. The three main functions supported by the system are: (1) providing early information about the approaching workers to avoid delayed actions, (2) providing information about the workers'

location to avoid operators' misjudgment, and (3) providing information about approaching hazard areas to avoid workers' indiscretion. The *RFID* in the proposed system is used only to identify the presence within a certain radius of the tags/readers and is not applied to measure distances between objects.

Yang et al. (2012) developed an integrative system comprising ZigBee-enabled Wireless Sensor Network (WSN), passive *RFID* and ZigBee *RFID* sensor network structure to record and identify accident precursors. The system tracks and collects data such as staff IDs, photos, access authorities, time of inspection, access control, checking personnel, machine conditions, repair work, etc. The core objective is to provide a good near-miss accident database to the managers so as to facilitate safety decision-making.

A *GPS*-based collision avoidance system was developed by Wu et al. (2013), with the central objective to assist crane operators with handling concrete buckets in a dam construction project. Talmaki and Kamat (2014) proposed the application of hybrid virtuality for the simulation of the actual jobsite and detecting hazardous proximities between various objects using a combination of 3D *CAD* models, terrain models, *GPS* and sensory data, and input from a Geographic Information System (GIS). This method uses proximity measurements as the basis for the collision detection.

Zolynski et al. (2014) developed a two-layer safety mechanism for autonomous excavators that generates a safety buffer around the equipment based on the present pose of the equipment and avoids collisions with the surrounding objects using a laser scanner. Another instance of the methods that use buffer workspace is developed by Guenther and Salow (2012).

However, the cylindrical workspace reserves a large space for the safe performance of the equipment, considerably diminishing its effectiveness for the application in a congested site. On

the other hand, while performing better in terms of economic use of space, the buffer workspace takes more time to detect potential collisions compared to the cylindrical workspace. In both cases, the shape of *DEW*s generated through these methods is determined by the *DOF*s and the geometry of the equipment. Ignoring the movement characteristics of the equipment, i.e., the magnitude and direction of the instantaneous speed, results in reserving a large space around the equipment. However, a portion of this space can be safely used by other equipment if the workspace is defined more efficiently through considering the movement characteristics of the equipment.

2.5.2.2 DEWS BASED ON THE EQUIPMENT GEOMETRY, POSE, AND SPEED CHARACTERISTICS

Other researchers have deployed the information pertinent to the equipment movement characteristics to enhance the proximity measurement with a degree of prediction about the possible states of the equipment in the near future. These methods do not only rely on the proximity between various equipment and objects as the indication of imminent hazards, but also use the movement characteristics of the equipment to foresee if the equipment is likely to engage in potentially risky situations if it follows its current trajectory. For instance, Burns (2002) proposed a method for the generation of workspaces around autonomous equipment based on a set of characteristics such as the current position, trajectory vector, speed, system tolerances, etc.

Oloufa et al. (2003) proposed a *GPS*-based collision avoidance system which is based on the simulation of potential collision through the analysis of equipment's motion vectors. The motion vector is measured by a server computer which wirelessly receives data from the *GPS* receivers mounted on the equipment. The system considers the required braking distance and the speed of the different pieces of equipment to generate warnings.

A probabilistic approach was proposed by Worrall and Nebot (2008) that uses the speed profiles of the equipment to account for the properties of the road network when detecting the impending vehicle intersection.

Having scrutinized the best practices in the mitigation of the risk factors, as discussed earlier, Chi and Caldas (2011) concluded that speed and proximity measurements are the two types of essential data required for the real-time risk assessment. A vision-based system is proposed for the speed and proximity measurements and identification of safety rule violations, e.g., speed limit or dangerous access violations.

Zhang and Hammad (2012) used *UWB* to track the movement of equipment, and further used the location data to create a buffer around the equipment based on its speed, which is in turn used to identify the potential collision between different pieces of equipment.

Cheng (2013) suggested the use of pose and speed data for the generation of the *DEW*. However, the proposed method does not consider the equipment state as a means to economize the use of space around the equipment and does not cover the equipment with rotary movements, e.g., excavators.

In another instance, a *GPS* and Radio Frequency (RF) based collision avoidance system was developed by SAFEmine (2014), where the speed characteristics and the pose of the equipment are used to generate the dynamic workspace of the equipment and generate warnings when workspaces of different pieces of equipment collide with one another.

In a more recent work, Wang and Razavi (2015) suggested considering the breaking distance, the reaction time, the equipment heading, and speed to create a workspace with a lower false alarm rate. Despite making a distinction between moving and static equipment, the method does not

fully exploit the equipment state information. Additionally, the adapted approach in this research treats equipment as a single point and disregards the geometry of the equipment.

In general, the existing methods do not distinguish between different states of the equipment when generating the workspaces. The valuable information about the state of the equipment can help better determine the size of the *DEW* in view of the potential dangers that may emanate not solely from the speed characteristics of the equipment but also from the nature of the equipment's current state. Therefore, it could be argued that if the combination of equipment geometry, current pose, state, and speed characteristics are properly leveraged, it is possible to economize the space usage without sacrificing the effectiveness of the collision detection.

2.5.3 LOOK-AHEAD EQUIPMENT WORKSPACE (*LAEW*)

As mentioned in Sections 2.5.1 and 2.5.2, researchers have addressed the need for the enhanced safety of earthwork equipment in two different streams, i.e., activity workspaces and *DEW*s. The integration of the two approaches can result in the overall mitigation of the equipment-related collision risks through considering the safety both at the planning and monitoring phases. However, there is still a middle level that is left uncovered. This is because, while the activity workspaces can be used to perform the initial path planning of different equipment, such planning tends to lose its efficiency in the face of the multitude of unforeseen circumstances that may occur during a project. On the other hand, the dynamic equipment workspaces are merely designed as "*the last line of defense*" (Zolynski et al. 2014) to warn the operators against imminent collisions and, thus, are not able to provide the information and time window required for the path re-planning of the equipment. Accordingly, there is a need for a middle-level solution at the monitoring phase that is able to reliably predict the operation of the equipment for

a long-enough time window to enable different pieces of equipment adjust their planned paths to avoid collisions in near-real-time.

In one of the efforts to address this need, Hukkeri (2012) proposed a safety mechanism based on the intention mapping technique in which every piece of equipment speculates the potential path of other mobile objects on the site and tries to avoid collisions. However, this method is based on a statistical modeling approach that ignores the underlying logic of a construction operation to predict the future movements of the equipment.

The emerging methods for near-real-time simulation of construction operations mentioned in Section 2.2.3 are providing the adequate inputs for a middle-level solution. Such methods are trying to build on the underlying logic of the operation, which is embedded in a simulation model, and use the data collected from the operation to continuously update its initial simulation model. On this ground, the valuable information about the cyclic pattern of equipment activity, which is an inherent feature of different types of earthwork equipment, and their movement characteristics can be fully leveraged to correlate the shape of the workspaces with the future expected poses and states of the equipment.

Stentz et al. (1999) proposed another middle-level solution based on the prediction obtained from a parametric motion planning technique. Although the presented approach is efficient in finding a collision-free path for a single excavator, it is not able to consider a fleet of equipment and their interactions in determining the potential collisions between different pieces of equipment. Additionally, only predicted collisions are used as the basis for the warning. However, given the uncertainties involved in the predictive models, near-miss instances can present as much risk as collisions. Furthermore, in addition to distance-based risks, the blind spots of the equipment can place the safety of other equipment and crews at risk (Soltani et al. 2002).

It is imperative to develop a method to generate Look-Ahead Equipment Workspaces (*LAEWs*) that consider not only the proximity-based risks but also the visibility conditions of the site vis-à-vis the future states of the equipment. Such a method needs to consider the operation pattern of the equipment and the visibility conditions of the space around the equipment, in addition to the input information used for the generation of *DEW*s, to determine a relatively longer-term spatial risk assessment (e.g., for the next 10 s) of the space surrounding the equipment. The spatial risk analysis leads to the generation of equipment risk maps that represent the risk distribution in the space around the equipment. These risk maps can then be used to generate the *LAEWs* associated with a certain risk level.

2.6 MULTI-AGENT SYSTEM (*MAS*)

The rise of empiricism and logical positivism, and the advent of modern theories in philosophy of mind, cognitive sciences and mathematics have inspired researches to structure human thinking in terms of imitable formalism (Russell and Norvig 2003). Artificial Intelligence (AI) has emerged from such efforts to mimic human ratiocination and to humanize computer systems (Ferber 1999), so much so that the Turing Test (Turing 1950) determines intelligent systems based on the extent to which the computer response to a problem is indistinguishable from that of a human beings. As mentioned in Section 1.2, when a problem grows in the complexity and the extensiveness, the centralized problem solving approach becomes less efficient. According to Ferber (1999), distributed problem solving can be an effective approach for such complex problems. In order to solve problems through distributed intelligence, a system consisting of several agents with local views of the problem are developed. Such systems are referred to as *MAS* wherein a group of autonomous agents interact to collectively solve a problem. The main

ingredients of *MAS* are intelligent agents capable of social interaction, which encapsulate the notions of agency, intelligence and interaction.

2.6.1 INTELLIGENT AGENTS

An agent is defined as an entity situated in an environment with the capability to form a perception of the environment and act upon it, which in the context of AI is materialized through the application of sensors and actuators (Russell and Norvig 2003). Figure 2-21 visualizes the basic architecture of an agent. Inextricably, agency is tightly associated with such notions as situatedness, perception, and autonomy. Agents are situated in an environment which basically represents the problem for which a solution needs to be sought. For agents to create a frame of reference for the environment, they need to perceive. Perception simply refers to the agents' perceptual inputs at any given point in time (Russell and Norvig 2003) and intends to elicit the understanding of the environment in which the agents are situated. Autonomy in agents refers to their ability to maintain partial control over their behavior and act without the intervention of humans or other systems (Weiss 1999). Unlike the conventional reasoning entities in classic AI, agents are capable of reacting to the changes in their environment, i.e., reactiveness, and carrying out actions and changing their environments based on their limited perception of the global environment even without external stimuli, i.e., pro-activeness (Ferber 1999).

According to Russell and Norvig (2003) the typology of agents comprises four main categories, namely: (1) simple reflex agents, (2) model-based reflex agents, (3) goal-based agents, and (4) utility-based agents. Simple reflex agents will react to certain environmental stimulus solely based on the present perception. It is possible for agents to keep track of its perception and maintain an internal state for aspects of the environment that may become un-perceptible later. Such agents tend to possess a form of knowledge that allows them to know how the world works

through a model. Accordingly, this type of agents is known as model-based reflex agents. While the aforementioned two agent types tend to be acting solely based on knowing the state of the world, this may not be sufficient to opt a right course of action when there is a goal to be satisfied. If agents are endowed with the ability to search and plan for an action that contributes to meeting a goal, they are known to be goal-based agents. Finally, although goal-based agents are able to realize a goal, they might choose for a suboptimal course of action. In face of such phenomenon, utility is defined as the extent to which an agent feels happy about a state. A utility-based agent tries not only to attain its defined goals but to realize them with a maximized utility. The first two types are also known as *reactive agents*, being only able to act upon environmental impetus, and the second two types are referred to as *cognitive agents*.

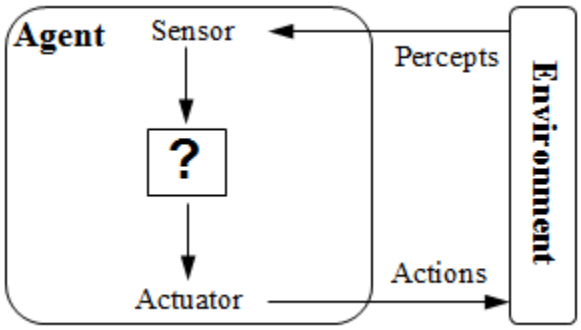


Figure 2-21: Agents Interact with the Environment through Sensors and Actuator (Adapted from Russell and Norvig 2003)

Referring to the last type of agents, *intelligence* is defined as the ability of an agent to pursue its own goals in such fashion to optimize some given performance measures (Weiss 1999), where performance measures are defined as the indication of successful behavior of an agent (Russell and Norvig 2003). As such, an intelligent agent should not be misconstrued as omniscience or omnipotence entity with infallible decisions (Weiss 1999), as it more pertains to the notion of

rationality, which requires an agent to choose a course of action that maximizes the expected, as opposed to the actual, utility (Russell and Norvig 2003).

2.6.2 AGENTS AS SOCIAL ENTITIES

Imposed by the inherent structure of distributed problem solving approach in *MAS*, which disintegrates the problem into segments pertinent to different functional or expertise areas and allows each agent to have only a limited view of the problem, agents in *MAS* are bound to interact in order to solve the problem communally. *Interaction* in agents implies the possibility of agents being affected by humans or other agents while they are in pursuit of their individual goals (Weiss 1999). Various interaction situations can be instigated by different stances of agents against one another in relation to their goals, available/required resources and skills (Ferber 1999). Basically, agents can be indifferent, cooperative or antagonistic towards one another, as shown in Table 2-4. Indifference represents a situation where agents do not need cooperation or competition to satisfy their goals. When agents face a shortage of required resources but have compatible goals their mode of interaction is known to be cooperation. On the contrary, when two agents with incompatible goals compete with one another over the scarcity of resources or insufficient skills, they are said to be in antagonistic towards each other.

With the ultimate objective of the interaction between agents in a *MAS* being the creation of a more coherent system through self-regulated actions and behavior of individual actions, interaction is as well referred to as coordination. The level of coordination is determined by the extent to which extraneous activities can be forestalled by dint of reducing resource contention, avoiding livelock and deadlock, and marinating applicable safety conditions (Huhns and Stephens 1999). With reference to the classification of interaction situations presented in Table 2-4, two types of coordination can be organized between agents based on the compatibility

of goals, namely, cooperation and negotiation. Whereas the earlier is the coordination mode of choice when there exists compatibility between agents' goals, the latter is used among antagonistic agents (Huhns and Stephens 1999).

Table 2-4: Classification of Interaction Situations (Adapted from Ferber 1999)

Goals	Resources	Skills	Types of situations	Category
Compatible	Sufficient	Sufficient	Independence	Indifference
Compatible	Sufficient	Insufficient	Simple collaboration	
Compatible	Insufficient	Sufficient	Obstruction	Cooperation
Compatible	Insufficient	Insufficient	Coordinated collaboration	
Incompatible	Sufficient	Sufficient	Pure individual competition	
Incompatible	Sufficient	Insufficient	Pure collective competition	Antagonism
Incompatible	Insufficient	Sufficient	Individual conflicts over resources	
Incompatible	Insufficient	Insufficient	Collective conflicts over resources	

2.6.2.1 COOPERATION

According to Ferber (1999), when cooperating, agents identify and adopt a common goal and subsequently engage in a common action. Cooperation between agents takes place in six different modes: (1) grouping and multiplication, (2) communication, (3) specialization, (4) collaborating by sharing tasks and resources, (5) coordination of actions, and (6) conflict resolution by arbitration and negotiation.

2.6.2.2 NEGOTIATION

Through negotiation, agents are purported to reach to an agreement based on concessions and compromises so that the collective utility of agents in the conflict is maximized. Huhns and Stephens (1999) introduced two different types of negotiation approaches, namely, environment-centered and agent-centered. The earlier is committed to embed some rules in the environment to

ensure that agents can fairly and productively interact. On the contrary, the agent-centered approach addresses negotiation as a mechanism within agents that determines the best strategy in the encountered conflicts.

2.6.3 ORGANIZATION OF MAS

The organization in *MAS* addresses two different aspects. The first aspect is about the formation of a team of agents in terms of types of different functions in agents, relationship between various agents, team structure, etc. The second aspect of organization revolves around the structure of the agents individually in terms of the process through which a perception results in an action. The criteria that constitute variations in organizations at the team and individual structure level are presented in Ferber (1999).

Different approaches could be adopted in terms of the underlying design mentality that determines a *MAS* structure. Ferber (1999) categorizes these approaches into object-centered, space-centered and functional. In the object-centered approach, agents represent different objects in the real world and the control and analysis are exercised from the point of view of objects. The alternative space-centered approach manages the operation in terms of various space-occupying objects. In other words, in a space-centered design, agents are representing different portions of the environment in which real world objects are engaged. Finally, the functional approach disintegrates the real world system into several functional areas and assigns each function to an agent.

Another taxonomy of the *MAS* organization refers to the centralized versus de-centralized structure. The distinction can be made at two levels, namely at the planning level or at the coordination level. The centralized planning refers to the process where a single agent is delegated as a planner, and usually the same agent ensures that the execution of the plan between

the subordinate agents is synchronized. On the other hand, it is possible to only centralize the coordination of the agents. In this mode, agents develop their own partial plans and send them to the coordinator agent for the synchronization and synergization of the agents plans, through removing the potential conflicts between agents. The de-centralized (distributed) *MAS* structure, on the contrary, requires each agent to develop its own plan and execute it autonomously. In this structure, the conflicts between agents are resolved en route, i.e., while agents are executing their respective tasks, through different modes of negotiation and collaboration explained earlier (Ferber 1999).

2.6.4 MAS AREAS OF APPLICATION

The areas for which *MAS* is used are manifold. However, based on the purpose of the application two main streams can be identified (Rojas and Giachetti 2009). One stream is focused on the enhanced decision making through the application of intelligent agents that can perform tasks and take over part of decision making. Agents, in this context, are used to substitute humans, partially or completely, on tasks that require high computational efforts or complex coordination. Examples of such systems can be found in the work of Tambe (1997), Yen et al. (2006), and Fan and Yen (2007).

Another major line of *MAS* application, as elaborated in Section 2.2.1, is dedicated to the simulation of processes with the objective to make predictions about a system in view of complex organizational, individual and environmental interdependencies among its components. Examples of *MAS*-based simulation were presented in Section 2.2.1.

2.6.5 APPLICATION OF MAS IN CONSTRUCTION

Although, as shown in Section 2.2.1, the *MAS* application for the simulation of construction operations has been investigated, the application of *MAS* as a decision making and planning tool with the capacity to partially or completely substitute human is palpably understudied.

Ren and Anumba (2002) have developed a *MAS* framework with the learning capabilities to facilitate the construction claim negotiations. In this framework, three types of agents are used to represent contractors, engineers and clients, as the main participants of a claim negotiation. A negotiation strategy is proposed that calculates an agent's likelihood of risk acceptability based on the amount of the utility the agent achieves in different scenarios.

Kim and Russel (2003a; 2003b) proposed a framework for an intelligent earthwork system based on the application of a *MAS* architecture. In this framework, three main subsystems, namely task planning, task execution and human control, have been suggested to provide complete coverage for an earthwork project task assignment and execution. The focal points of this research have been threefold: (1) identifying an optimal task sequencing, which uses heuristic rules based on the shortness of distances and low space interference to determine best cut-fill distribution over an equally gridded area; (2) determining the optimal configuration of resources, using simulation, and assigning different pieces of equipment to different tasks based on agent negotiations; and (3) identifying the collision-free path for the equipment. The proposed structure, while providing a good structure for the distribution of responsibilities between a multi-layer agents hierarchy, assumes an un-deviated execution of the schedule and does not cover functionalities to monitor the actual progress of the operations vis-à-vis the generated schedule in order to provide proactive corrective measures for keeping the project on schedule. Additionally, safety issues of construction site are not addressed by the proposed *MAS* structure.

Overall, this research is conducted at a very high level of abstraction, with a limited scope, and refrains from elaborating the communication protocol between agents and the formal description of using agents' knowledge for the analysis.

In another effort, an agent-based system was proposed for the communication in construction project by Lee and Bernold (2008). In this framework, a set of agents are used to collect weather data from internet and on-site instrumentations, forecast hazardous weather conditions and warn the cranes. Also, agent-based communication was used to facilitate site-to-office communication with the intention to reduce, or eliminate, the hindering process of information request. The onsite communication between agents is realized through Wi-Fi network.

A multi-agent approach was developed by Zhang and Hammad (2012) for the real-time collision avoidance and path replanting for cranes. In this research a hybridized planning approach that combines distributed and centralized problem solving method is proposed to represent several agents on a construction site. A three-type agent structure, comprising site state agent, coordinator agent and crane agents, was used to track the movement of cranes on the site, predict possible collisions and generate a new path for the crane with the lower priority to avoid the detected collisions. This research is the starting ground for the present research, in the sense that the present research is aimed to extend the proposed framework to the earthwork operations and to multiple construction equipment.

Overall, although the applications of *MAS* for the conflict resolution at the managerial level and for claim management have been subject of an ongoing investigation, the operational aspect of *MAS* for the autonomous control and monitoring of the construction project is evidently studied only scantly. This research draws on the research of Zhang and Hammad (2012) and Kim and

Russel (2003a; 2003b) to extend the application of *MAS* for the task assignment and monitoring of an entire fleet of earthwork equipment.

2.7 SUMMARY AND CONCLUSIONS

This chapter was dedicated to the review of the literature on the several areas that pertain to the topic of the present research. As shown in Table 2-5, the research gaps were identified based on the extensive review of the literature, and these gaps will be addressed in the following chapters using the proposed framework.

Accordingly, Chapter 3 will present the overall proposed *MAS* framework and together with Chapter 8 they address the limited machine-level scope of the present *LGS*s. Chapter 4 will focus on a method to improve the pose estimation of the equipment based on affordable *RTLS*s. Chapter 5 will address the shortcomings of the existing *NRTS* approaches and presents a novel approach that covers a wider range of tracking technologies and can also account for the environmental impacts on productivity fluctuations. Finally, focusing of the safety management of earthwork sites, Chapters 6 and 7 will present new methods for *DEW*s and *LAEW*s. These two workspaces can conjointly provide a two-layer shield for the equipment that can be used to perform near real-time path re-planning and collision avoidance.

Table 2-5: Summary of Literature Review

Area of Literature Review	Research gaps	Points of focus in this research	Chapter covering the area
Coordination Improvement Using Automation and LGS	• Limited functionalities at the project level	• Developing a mechanism through which the monitoring and guidance of equipment are performed with the consideration of the entire fleet.	Chapter 3 & Chapter 8
	• High cost of retrofitting the equipment for AMC/G and the insufficient accuracy of affordable RTLS	• Developing an optimization-based method to improve the accuracy of low-cost RTLS	Chapter 4
Real-time Simulation	• Technology-dependent frameworks • Limited equipment coverage • Equipment-restricted view • Indifference to environmental factor	• Developing an environment-aware framework that accommodates a range of technologies for a fleet of equipment with the consideration of the entire fleet for state-identification	Chapter 5
Project Monitoring and Control	• Limited coverage of equipment (mainly trucks) • Low level of captured detail about the equipment	• Fusing a wide array of location and sensory data that can be used to capture the location and orientation of equipment and further calculate the cycle time of both relatively fixed equipment, e.g., excavators, and highly mobile units, e.g., trucks.	Chapter 4 and 5
Earthwork Safety	• Dependence on the proximity measurement using different technologies • Lack of predictive safety evaluation based on the results of NRTS	• Considering a wide range of sensory data to develop effective DEW. • Building on the results of NRTSs to evaluate and identify safety risk in a predictive manner in a wider time frame than possible with existing proximity-based methods.	Chapters 6 and 7
MAS	• Limited coverage for the monitoring and control of earthwork equipment.	• Extending the application of MAS for the task assignment and monitoring of the entire earthwork equipment fleet.	Chapter 3 & Chapter 8

CHAPTER 3: OVERVIEW OF THE PROPOSED FRAMEWORK

3.1 INTRODUCTION

As mentioned in Section 2.6, although the applications of *MAS*s for the conflict resolution at the managerial level and for claim management have been subject of an ongoing investigation, the operational aspect of *MAS*s for the autonomous control and monitoring of construction projects has been studied only scantly. This research draws on the research of Zhang and Hammad (2012) and Kim and Russel (2003a; 2003b) to extend the application of *MAS*s for the task assignment and monitoring of an entire fleet of earthwork equipment.

3.2 OVERVIEW OF THE PROPOSED FRAMEWORK

Figure 3-1 shows an overview of the scope for the proposed *MAS* framework. The main assumptions are that every piece of equipment on the construction site has a sufficient number of *RTLS DC*s attached at specific locations to track its movement, and that every equipment operator is supported by an agent that can communicate with other agents in a *MAS* framework. The proposed *MAS* supports the project at three different levels, namely (1) Planning, (2) execution and monitoring, and (3) re-planning. At the planning level, the *MAS* is able to streamline the operation and task assignments to different equipment as well as to support equipment path planning (Zhang and Hammad 2012), which is operationalized in terms of strategic and tactical planning as explained in Chapter 2.

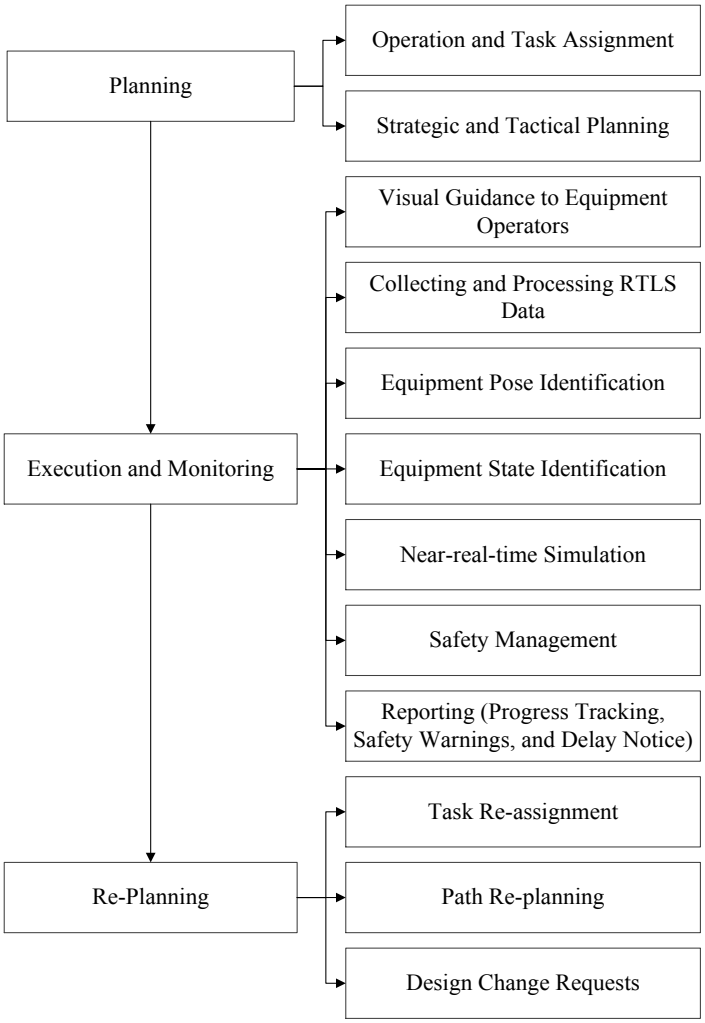


Figure 3-1: Overview of the Scope for the Proposed MAS Framework

At the execution and monitoring level, the *MAS* is committed to: (i) providing visual guidance to equipment operators, (ii) collecting and processing *RTLS* data, (iii) applying appropriate error correction techniques to identify the pose of the equipment, (iv) identifying the state of the equipment, (v) applying the *NRTS*, (vi) performing safety management, i.e., generating *DEW*s and *LAEW*s, and (vii) reporting the necessary information to pertinent agents. The aforementioned two types of workspaces differ in that while *DEW*s are generated based on the equipment pose, state, geometry, and speed in real-time to form a safety buffer around the equipment that can help prevent collisions, *LAEW*s are built based on the predicted future motion

of equipment and operator visibility in near real-time to help find a collision-free path for equipment. More details about the two workspaces can be found in Chapters 6 and 7. Finally, at the re-planning level, the proposed *MAS* framework addresses the need for task-reassignment, path re-planning, and design change requests, which may become necessary in view of the potential unforeseen safety risks identified at the monitoring level. As can be seen in Figure 3-1, while the proposed *MAS* framework offers advantages at the operational and managerial levels, only the operational aspect of the framework is addressed in this research.

3.3 MAS STRUCTURE

A multi-layer agent architecture is proposed in which agents supporting the operators of equipment constitute the lowermost layer of the agent hierarchy. These agents process and manage the huge amount of sensory data, provided by *RTLS*, into useful information that can be used in decision-making at different operational and managerial levels.

Figure 3-2 shows the proposed *MAS* architecture where several teams working in the proximity of each other are supported by different types of agents with different tasks and project views.

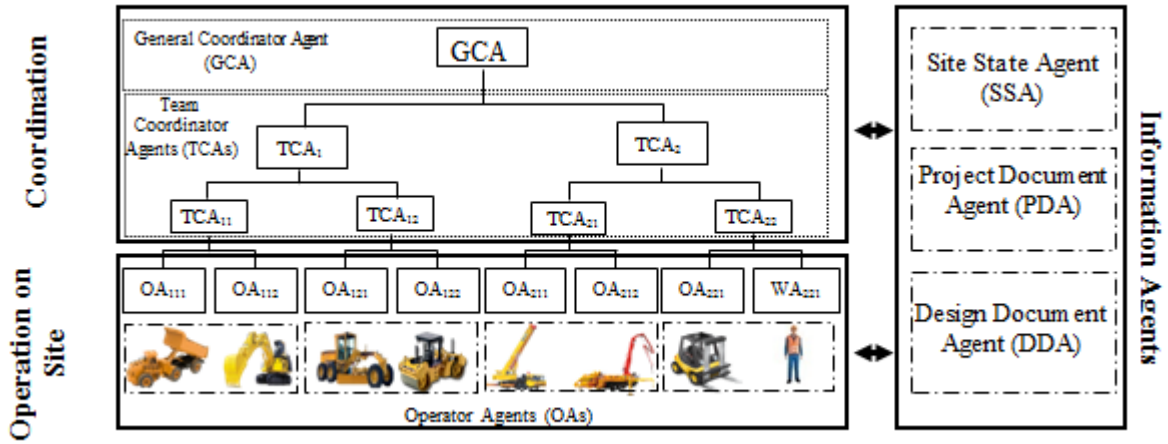


Figure 3-2:Multi-Agent System Architecture

Three functional agent types can be distinguished according to the distribution of the responsibilities, namely, operator, coordinator and information agents. In a nutshell, Operator Agents *(*OAs) support the equipment operators and have the essential knowledge about their current task, state and pose. In a construction site, often a group of equipment is teamed up to serve one particular operation, for instance several trucks and an excavator work together to move the earth. The team coordinators are supported by Team Coordinator Agents (TCAs), whose main objective is to track the progress of operations based on the data gathered from their subordinate *OA*s and to ensure safe and smooth delivery of the operations' objectives. Depending on the level of coordination each *TCA* offers, several layers of *TCA*s and a General Coordinator Agent (GCA) can be defined. Furthermore, these different types of agents are fed by Information Agents who provide the required site, project and design-related data to the site agents, and frequently get updated based on the changes happening in the site as the construction progresses. These encompass the Site State Agent (SSA), Project Document Agent, *(*PDA), and Design Document Agent (DDA).

Although it is outside the scope of the present research, in an extended architecture, workers, surveyors, designers and project managers can also be represented by unique agents, to further enhance the smooth human-human and human-equipment interaction, negotiation, conflict resolution and document management. Furthermore, the consideration of workers-on-foot is of a great importance for effective safety management of earthwork sites. All the workers working on the site must be supported by dedicated agents who can track workers' movements and provide necessary information (primarily safety warnings) using personal digital assistants or smart phones. Therefore, this research assumes that all the workers are supported by simple Worker Agents (WAs) that can send and receive such information.

It is worth mentioning that in addition to the technical aspects of the proposed framework, some administrative issues need to be considered for the future implementation of the proposed *MAS* in earthwork sites. The details of these administrative aspects are presented in Appendix A.

3.4 RESEARCH COMPONENTS AND THESIS STRUCTURE

In line with the research objectives presented in Section 1.3, of the various scopes covered by the proposed *MAS* and shown in Figure 3-1, this research focuses on the execution and monitoring phase of earthwork operations. Therefore, the subsequent chapters focus on the major functions of *MAS* in the planning and execution phases. Figure 3-3 shows the functions that will be discussed in this research. Accordingly, Chapter 4 addresses the second objective of this research (i.e., to provide a method for improving the performance of pose estimation based on low-cost *RTLS*s) and describes a novel pose estimation method that can be utilized to improve the quality of low-cost *RTLS*. In this optimization-based method, a set of equipment geometric and operational constraints are used to apply required corrections to *RTLS DC*s. The proposed method will be tested using a case study and will be compared with a method previously proposed by Zhang et al. (2012).

Focusing on the third objective of this research (i.e., to devise a generic approach for *NRTS* based on data from *LGSs*), Chapter 5 elaborates the methods for state-identification and *NRTS* approach. In this approach, the input from the pose estimation method is used to identify the equipment states (e.g., loading, swinging, and dumping) and to detect the potential discrepancy between the expected and actual productivity of the operation. If the operation is found to be progressing slower or faster than expected, the simulation model is updated and a new schedule is generated. The proposed approach will be validated using an implementation and two case studies.

Chapters 6 and 7 aim to pursue the fourth objective of this research (i.e., to develop a mechanism for improving the safety of earthwork operations using the capabilities of the *LGSs*, *NRTS* and *MASs*) and introduces two types of workspaces, namely *DEW* and *LAEW* respectively. Using the outputs of pose estimation and state identification methods, these two types of workspaces, conjunctively enable the equipment to perform real-time collision avoidance and near real-time path re-planning, as explained in Section 3.2. These two methods will be verified through the developed implementations and case studies.

Chapter 8 provides an in-depth discussion of the *MAS* structure and aims to address the first objective of this research (i.e., to enable the project-level coordination, monitoring and control through the integration of a *MAS* architecture and a *LGS* to help better resolve operational and managerial conflicts in earthwork projects) by demonstrating how the smooth integration of functions presented in Chapters 4 to 7 can be realized in the proposed *MAS*. The high-level definitions of the structure of each type of agents, their functionalities, inputs, and outputs are presented. Also, the communication scheme between different agent types for certain scenarios are introduced and tested. Several simulated case studies will be used to demonstrate the feasibility of the proposed framework. Finally, Chapter 9 presents the contributions, conclusions, limitations, and future work of this research.

3.5 SUMMARY

This chapter presented the overview of the proposed *MAS* framework, where every operator and worker on-foot on the construction site is represented by a dedicated software agent. The overall scope and functionalities of the proposed framework were discussed. As mentioned in Section 3.4, the following chapters present the main functions of the proposed framework that have been

addressed in this research. Also, a more detailed account of this framework is presented in Chapter 8.

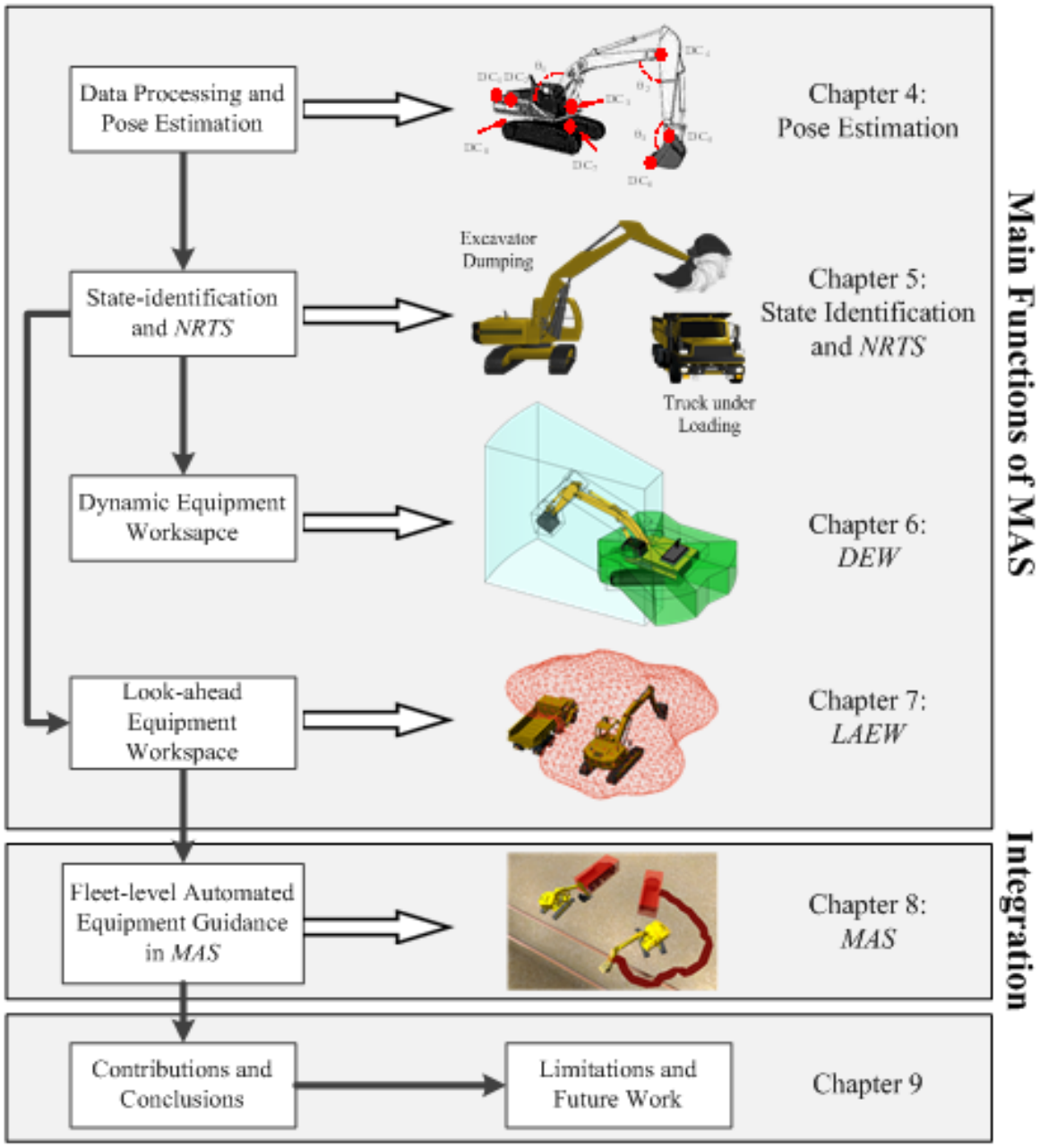


Figure 3-3: Research Methodology

CHAPTER 4: OPTIMIZATION-BASED EQUIPMENT POSE ESTIMATION

4.1 INTRODUCTION

The pose of a piece of equipment (i.e., its location and orientation) is the most basic and fundamental information required for the monitoring and analysis of the earthwork operations. When the pose of a piece of equipment is known, it can be effectively used to detect hazardous conditions and identify the state of the equipment, which in turn can be used to calculate the productivity-related indices. Advanced technologies embedded in *AMC/G* are capable of generating the tracking information with a high level of accuracy. However, as explained in Chapter 1, the high cost of procuring this technology for every piece of equipment limits the availability of *AMC/G* for small and medium size contractors. On the other hand, the lower accuracy of more affordable *RTLS*s, renders them impractical for the use in *LGS*. As a result, a method needs to be devised to process and refine the captured *RTLS* data, as explained in Section 2.3.1.1.

The present chapter aims to propose a novel and robust optimization-based method that uses the geometric and operational characteristics of a piece of equipment to improve the quality of the data captured by low-cost *RTLS*s so that the pose of the equipment can be estimated with a good accuracy. The method is purported to ensure that while the pose compliance with the machine-imposed constraints is maximized, it is estimated with a minimum amount of required corrections.

The proposed method is explained and framed using excavators as the main focal point. The excavators are chosen due to their level of complexity and the fact that they are able to concurrently perform two types of movement, namely rotational and traversal. Also, of various types of equipment engaged in earthwork projects, excavators account for the largest proportion of fatalities on site, with 36 reported incidents in the U.S. only in 2012 (BLS 2012). Nevertheless, the proposed method is generic in nature and can be applied to all types of earthwork equipment.

It is noteworthy that the pose estimation method, as will be explained in Section 8.2.1, is performed by the *OA* of each piece of equipment. The output of this method is then used by *OA*s to generate the *DEW*s, as will be discussed in Chapter 6. Similarly, *TCAs* use the pose data to enable the execution of *NRTS*, as will be explained in Chapter 5, and to generate *LAEW*s, as will be discussed in Chapter 7.

4.2 PROPOSED EXCAVATOR POSE ESTIMATION METHOD

The pose of a piece of equipment is a spatial feature that encompasses its location and orientation at any given point in time. The location can be measured in various formats, e.g., the triplet of longitude, latitude, elevation or the Cartesian coordinates in relation to a local reference frame. The orientation can be manifested in terms of the roll, pitch and yaw between the frame representing a rigid body and the reference frame (Schaub and Junkins 2003). The orientation reference frame can be established using either three *RTLS DC*s installed on the upper structure of the excavator or alternatively one *RTLS DC* and a digital gyroscope.

In this research, it is assumed that an excavator is equipped with a set of *RTLS DC*s, which continuously provide the location data of the objects to which they are attached with a certain

update rate, and the pose is estimated during certain intervals (ΔT), e.g., every 5 s. Note that the value of ΔT depends on the maximum *dt* that can be offered by the used *RTLS* under the site condition. To decrease the estimation interval, a high rate of *RTLS* update is required. Also, in order to improve the accuracy of the location data, it is assumed that every rigid part of the equipment is represented by at least two *DC*s. This redundancy helps increase the data accuracy and the visibility, i.e., the chance of missing data is reduced. In order to compensate for the missing or erroneous data, the raw data gathered from the *RTLS DC*s require a multi-step processing before they can be used for the pose analysis. Figure 4-1 shows the high-level flowchart of the proposed method, which consists of several steps to increase the accuracy of the pose estimation. The process begins with the averaging of data over a period of time and applying interpolation for filling the missing data. Next, the optimization-based correction is applied and the pose is calculated.

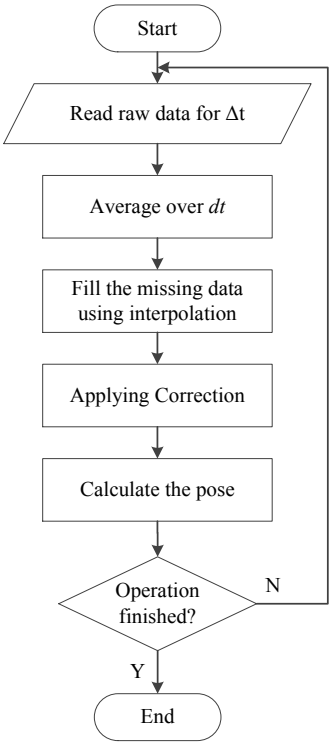


Figure 4-1: Flowchart of the Proposed Pose-estimation Method

4.2.1 AVERAGING OVER A PERIOD OF TIME AND FILLING THE MISSING DATA

The first step in the method is averaging over a period of time (*dt*), which is less than the analysis interval (Δt). For instance, Δt can be 5 s and *dt* can be 1 s. It should be noted that the length of Δt determines the extent to which the processing is real-time. The impact of averaging over a period of time on improved location data is elaborately discussed by Zhang et al. (2012). However, in a nutshell, knowing that the *RTLS* records the location data with a certain frequency, which depends on the specifications and configurations of the used system, it is possible to increase the accuracy of data by averaging several consecutive data over *dt*. The rationale behind this step is that random errors that typically exist in the captured *RTLS* data commonly follow a normal distribution, which enables the averaging to improve the accuracy of the *RTLS* data by allowing the negative and positive errors to level each other. The degree of impact of the averaging depends on the number of consecutive location data averaged. In determining the required averaging period (*dt*), it should be noted that while a long period debases the smoothness of the data by introducing jumps between two successive averaged data, a too short period may lead to insignificant improvements. The optimal averaging period should be devised according to the required smoothness and the update rate of the *RTLS*. Figure 4-2 shows an example of the *RTLS* data capturing pattern and the rows and columns in this figure represent the indices *j* and k of DC_j and dt_k, respectively. In this pattern, every cell contains the averaged location data over *dt*, i.e., x_{jk}, y_{jk}, z_{jk}. On the other hand, Δt represents a longer period of time over which a near-real time analysis is performed. Every Δt is represented by the index q and contains several *dt* periods, which are represented by the index *k*.

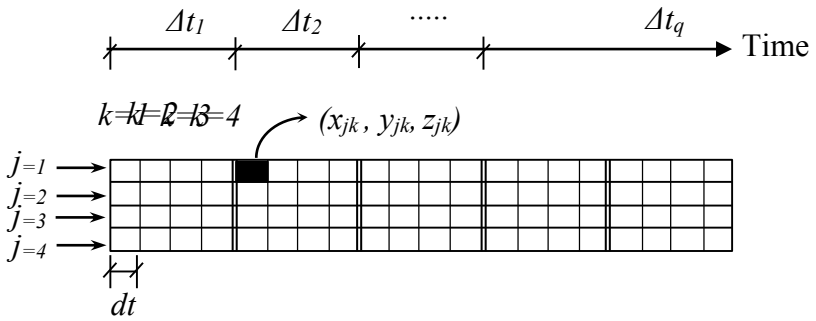


Figure 4-2: The data capturing pattern for RTLS

Once the data is averaged over *dt*, there is a chance that there are some missing data due to *RTLS* failure in registering data. In order to provide a smooth stream of data for further analysis, the missing data can be filled out using linear interpolation between preceding and succeeding data points.

4.2.2 APPLYING CORRECTION USING SIMPLIFIED CORRECTION METHOD

The *DC*s attached to an excavator are governed by a set of Geometric Constraints (*GC*s) and Operational Constraints (*OC*s). A GC requires the *DC*s data to respect a specific geometric relationship that exists between the parts of the excavator to which the *DC*s are attached, e.g., fixed distance or fixed orientation. The number and type of *GC*s depend on the *DOF*s of the excavator and the number of the *DC*s. An *OC*, on the other hand, pertains to the type of restrictions put on the excavator based on a particular condition of the terrain, e.g., rolling resistance and slope, or the operational specifications of the equipment. While this research presents a comprehensive optimization-based method for correction, the simplified method previously presented by Zhang et al. (2012) is explained in this section for completeness and in order to evaluate the proposed optimization-based method against the simplified method.

The *OC*s are applied through ensuring that the difference between two consecutive data entries pertaining to a *DC* does not violate the maximum operational speed limit of the equipment or the

part they represent. Figure 4-3 can be used as an example to clarify this method. In this figure, the subscripts of letters, e.g. 1 and 2 in A_1 and A_2, represent the sequence in time, and the superscripts ' and ", as in A' and A'', show already corrected points based on the *OC*s and already corrected points based on the *OC*s and *GC*s, respectively. For example, as shown in Figure 4-3(a), in the context of location data, if the distance between two consecutive readings, e.g. B_1'' and B_2, of a *DC* attached to a piece of equipment is *L*, while based on the maximum operational speed of the equipment this distance cannot be more than L_{Max}, a correction is applied to the latter point and the new location is calculated (B_2').

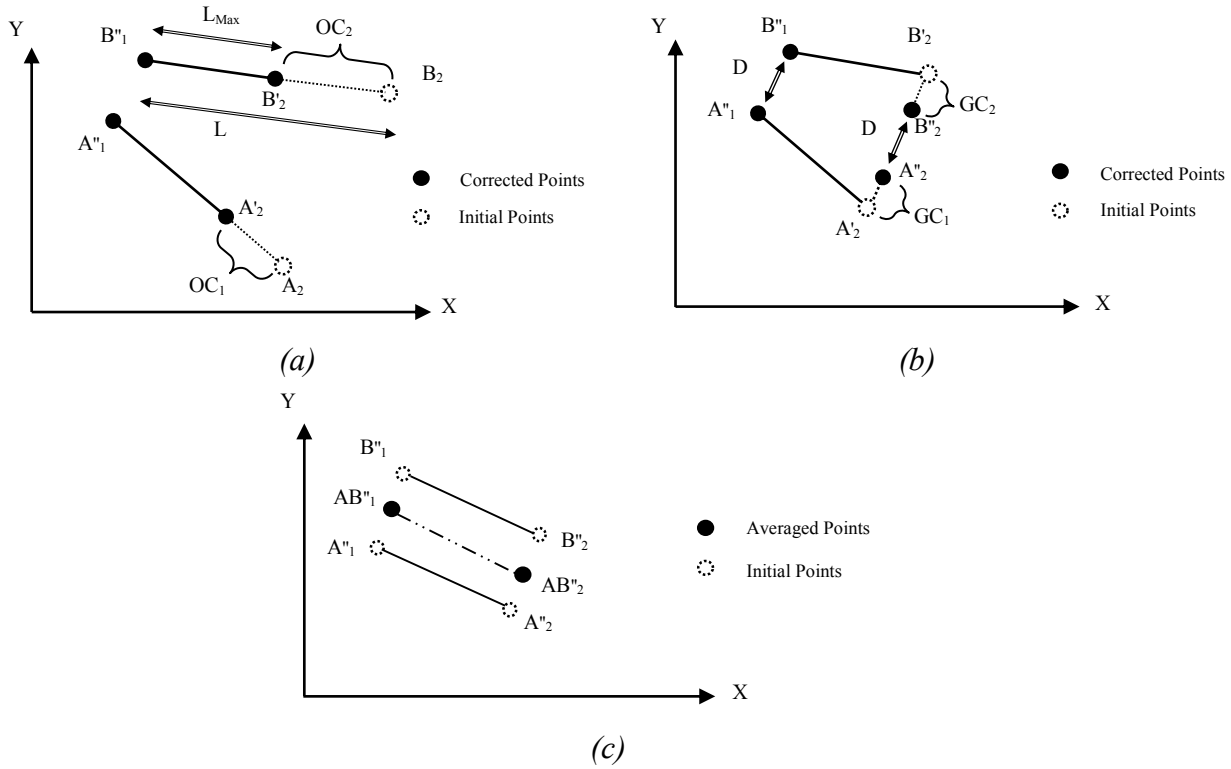


Figure 4-3: (a) Corrections Based on the Operational Constraints, (b) Corrections Based on the Operational Constraints, and (c) Averaging of Several Data Collectors that Are Attached to the Same Part

It is noteworthy that the maximum operational speed is the speed typical of various types of earthmoving equipment in a given project type and terrain conditions. The *GC*, on the other

hand, is applied through adjusting the data based on a fixed geometric relation between any given two *DC*s attached to a rigid body at any point in time. For example, as shown in Figure 4-3(c), if the calculated distance based on the readings of A_2' and B_2' is different from what it is measured to be (*D*), the amount of error is calculated, and the correction is equally distributed between the two parts, resulting in points A_2'' and B_2''.

It should be noted that the data correction is an iterative process of operational and geometric corrections. The flowchart of the iterative correction process is given in Figure 4-4. In this figure, *i* represents the *ID* of the *DC*s, *t* represents the sequence of data entries, *j* denotes the number of GCs between a certain pair of *DC*s and $\mathrm{DC}_1^{\mathrm{j}}$and $\mathrm{DC}_2^{\mathrm{j}}$ implies the *DC*s involved in the GC_{j}. Δt, on the other hand, represents the interval between two successive entries. The process begins with checking the *OC*s of every single *DC*s. For this purpose, the speed of every adjacent readings of a *DC* is analyzed and if a violation of the *OC* of that *DC* is identified, the error is corrected as shown in Figure 4-3(a). Once all the *OC*s are controlled, the data will be analyzed for the compliance with the *GC*s. To materialize this, first the number of existing *GC*s and their values are determined based on the user input. Then, if the standard deviations of all *GC*s are acceptably small, based on the user requirements, the process will end; else the correction will be applied to the *GC*s that have high standard deviation as shown in Figure 4-3(b). If any *GC* correction is applied, then the data needs to be checked for the possible violation of *OC*s that may have been caused by the *GC* corrections. The process will continue to the point that the standard deviations of all *GC*s are acceptably small. The reason for the possibility of convergence is that the operational correction is an intra-*DC* process while the geometric correction is an inter-*DC* process; and thus, corrections are made along different lines.

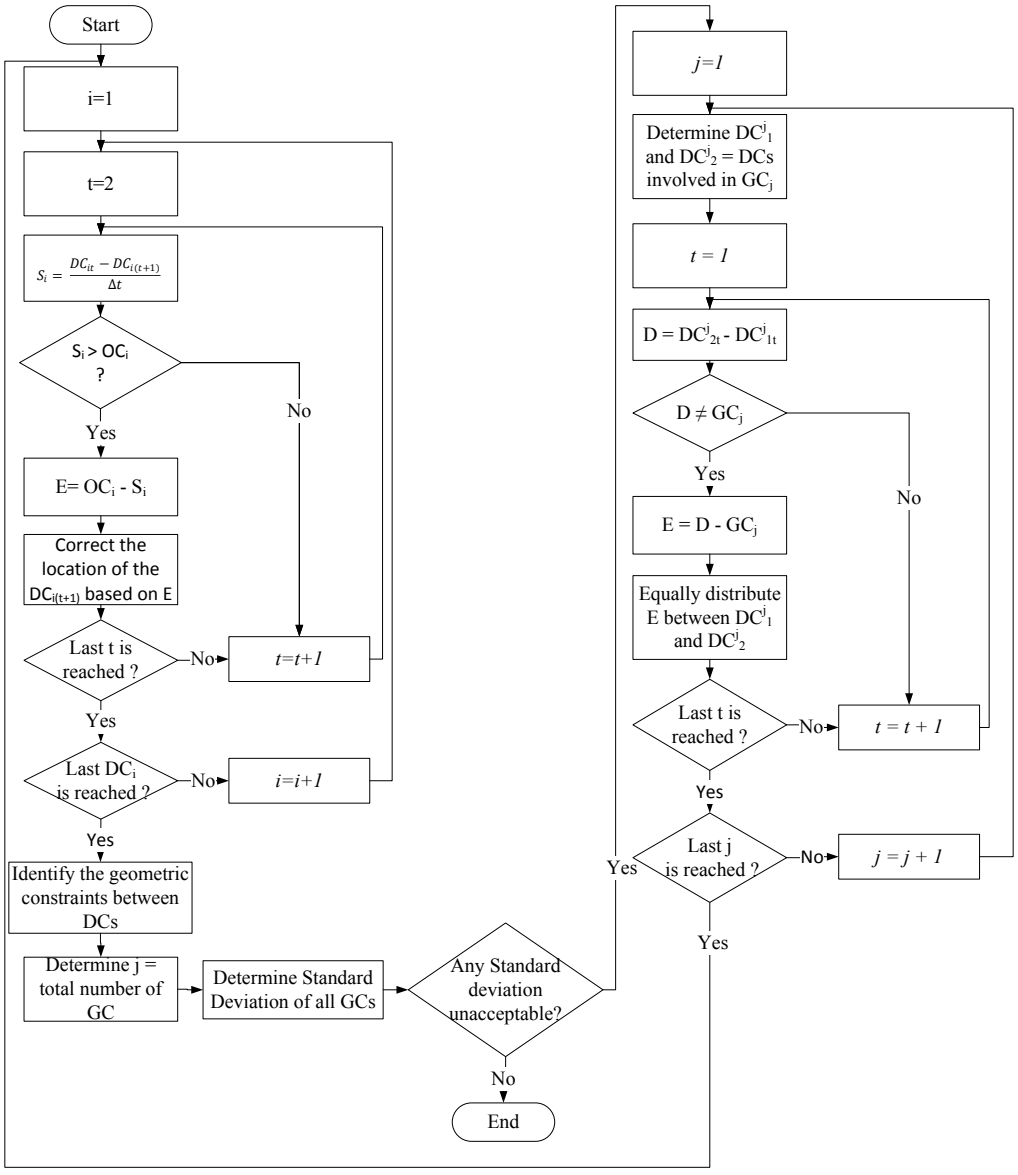


Figure 4-4: Flowchart of the Simplified Correction Method

The various user inputs required for the data-processing is given to the system only once at the inception of the project, based on the consideration of the project-specific characteristics, *RTLS* features and used equipment. The process flows automatically thereafter, for as long as the project is running and the input values are valid. These inputs, e.g. distance between tags, the update rate, number of tags, etc., are seldom subject to variations in the course of a single project.

Although the earlier two steps can be sufficient to obtain accurate results, under certain circumstance and if several *DC*s are not required for further analysis, the data can be further enhanced through averaging the data of several *DC*s attached to the same solid object. For instance, if the blade of a loader is represented by two *UWB* tags on the right and left sides, and if the angle of the blade is not required for further analysis, the data of the two tags can be averaged and represented by the intermediate point AB_1''. Figure 4-3(c) illustrates the processes where corrected readings of *DC*s ($A''$and B'') are averaged and represented by AB''.

4.2.3 APPLYING CORRECTION OPTIMIZATION-BASED CORRECTION METHOD

The simplified correction method has several shortcomings: (1) the increased number of required iterations when the number of constraints increases, (2) the inaccurate assumption that the error is equally distributed between two *RTLS DC*s along the axis formed between them, which may lead to a failed convergence when the number of constraints increases, and (3) the inefficiency in capturing non-distance geometric constrains, e.g., angular constrains. On this premise, it can be argued that the above-mentioned iterative correction method is not enough to maintain the full consistency of the data required for accurate pose estimation, especially when the number of *DC*s attached to one piece of equipment increases.

The optimization-based correction for the excavator location data includes two phases, each of which has an optimization algorithm. A Standard Genetic Algorithm (GA) (Goldberg and Holland 1988) is proposed to solve the optimization problems. GA is a robust optimization algorithm that is proven to be capable of searching the complex solution spaces with the reduced likelihood of restriction to a local optimum (Li and Love 1997). GA has been widely used for solving complex optimization problems in a wide array of engineering contexts (Feng et al. 1997; El-Rayes and Khalafallah 2005; De Giovanni and Pezzella 2010; Peralta et al. 2014).

As shown in Figure 4-5, the first phase of the correction is dedicated to finding the Center of Rotation (*CR*) of the excavator. It should be highlighted that two forms of CRs can be distinguished: (1) *CR* as a mechanical element of the excavator that represents the swinging hinge of the excavator underneath the upper structure, which can be calculated based on the geometric characteristics of the excavator and is referred to as CR_{cal} in this research; and (2) *CR* as a fixed point on the site with respect to which swinging takes place, which can be measured using the statistical method introduced in Section 4.2.3.1, and is referred to as *CR** in this research. The first phase of the proposed optimization-based method is trying to locate *CR**. It should be added that while CR_{cal} always exists as an indivisible part of the excavator's geometry, *CR** is only available when the excavator is not relocating. Therefore, when the excavator is swinging during a non-relocation period, the two *CR*s have to overlap as will be explained in Section 4.2.3.1. This should not be perceived as a limitation on the possible swinging action during the relocation, since CR_{cal} is always calculable based on the excavator's geometry. Instead, it uses extra information about the existence of *CR** during the non-relocation period to further improve the data.

The second phase of the correction aims to minimize the *DC*s errors in such a manner that a number of geometric and operational constraints of the excavator, including the rotation around *CR**, are respected.

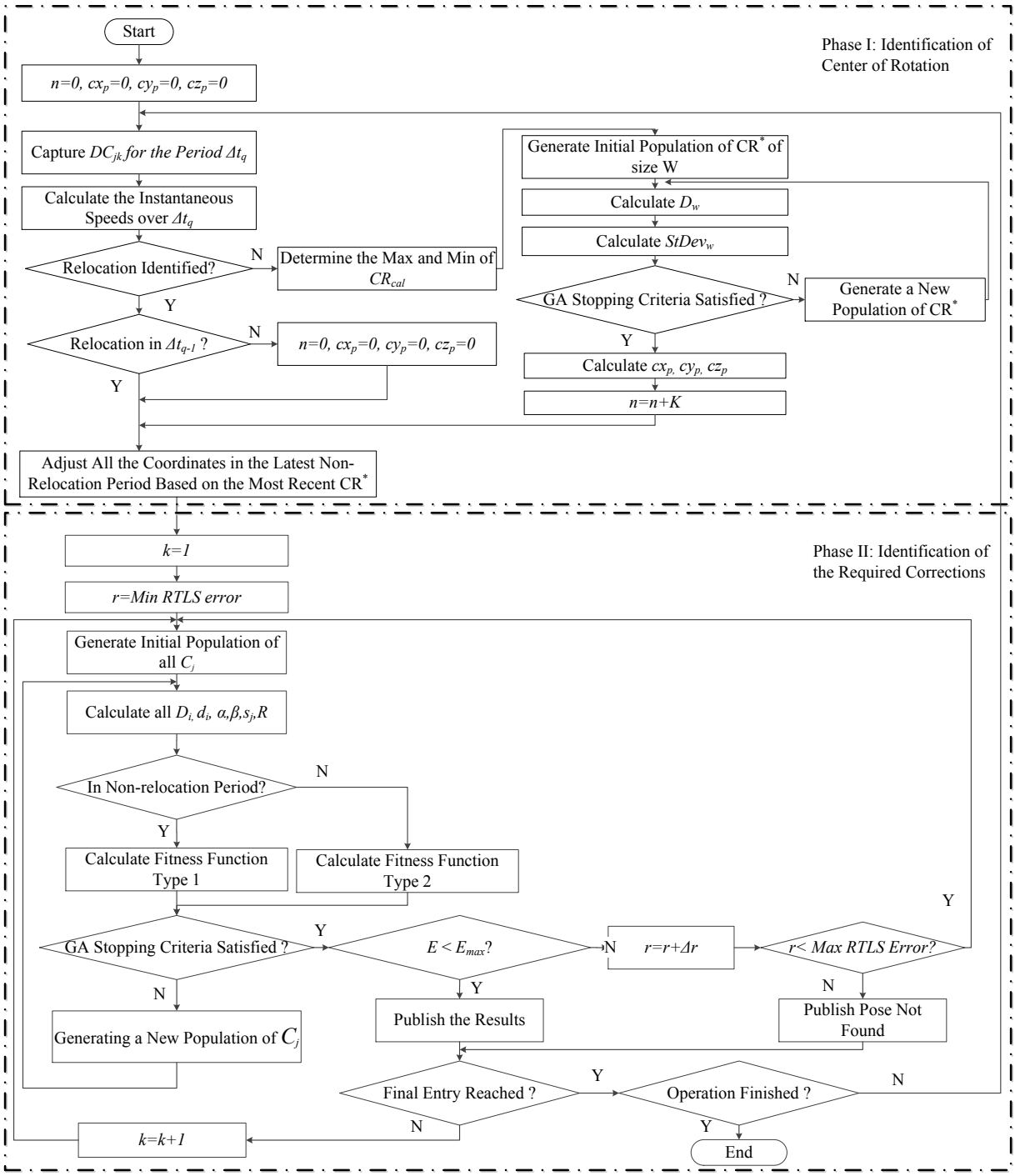


Figure 4-5: Flowchart of Optimization-based Correction

4.2.3.1 PHASE I: IDENTIFICATION OF CENTER OF ROTATION

Given that the *CR** imposes a time-dependent constraint that needs to be respected based on the status of equipment over time, i.e., relocation or non-relocation, and knowing that there is no pre-determined location for the *CR** during each period when the excavator is not relocating, its approximate location needs to be determined using statistical analysis.

As shown in Figure 4-5, the determination of *CR** begins with capturing the *DC* data over a period of Δt. Next, it is required to isolate the data pertaining to the period when the excavator is not relocating. This can be done using a simple calculation of the excavator speed at every point in time. The relocation is identified when the value of the measured speed is greater than a threshold for all *DC*s over a long-enough period of time. The rule for the identification of relocation state of an excavator and the definition of long-enough period of time for the state-identification of equipment is presented in Chapter 5.

If a period of Δt is identified as non-relocation period, the algorithm uses the characteristics of *CR* to determine *CR**. *CR* is characterized as a fixed point to which every *DC* on the upper structure of excavator maintains a unique but fixed distance throughout the operation during every non-relocation period. The upper structure is defined as the rigid part of the excavator's super structure that includes the operator cabin and engine, excluding the boom/stick/bucket system. For this purpose, first the range of possible CR_{cal} is determined over the period of Δt. The minimum and maximum possible CR_{cal} are measured by the two extreme values of CR_{cal} obtained from the raw data of tags attached near the actual CR during a non-relocating period, and are constant for all the solutions in every Δt. The CR_{cal} of the equipment can be calculated considering the relationship between the location of the rotation engine on the equipment and the positions of the *DC*s on the upper structure. For instance, in the excavator shown in

Figure 4-7(a), the hinge is positioned at the intersection between (1) the line crossing DC_3 and perpendicular to the boom axis and (2) the boom axis itself. It should be highlighted that although in majority of the models CR_{cal} is located on (or very near to) the axes of boom/stick/bucket, the shown relationship is a representation of the CR_{cal} in the model of the excavator used as an example in this chapter.

Next, the GA optimization is used to find the coordinates of CR^* (cx, cy, cz) within the range of values of CR_{cal} that results in the smallest total standard deviation ($StDev_w$) of the distances between CR^* and all the DCs attached to the upper structure of the excavator over Δt. The GA begins with generating a population of CR_w^* of size W. D_{jkw} represents the distance between each DC_j and every CR_w^* in the population during dt_k within Δt, where CR_w^* is member w of the population of CR^* in the current generation as shown in Equation 4-1.

$$D_{jkw} = \sqrt{(x_{jk} - cx_w)^2 + (y_{jk} - cy_w)^2 + (z_{jk} - cz_w)^2} \qquad \textbf{Equation 4-1}$$

Where:

j: index representing the ID of DC_j, $j \in [1, m]$;

k: index representing an averaging period dt_k within Δt, $k \in [1, K]$;

w: index representing a solution in the CR* population of size W generated by the GA, $w \in [1, W]$;

x_{jk}, y_{jk}, and z_{jk}: the coordinates of DC_j during dt_k; and

cx_w, cy_w, and cz_w: the coordinates of a CR_w^* during Δt.

D_w is the matrix that includes all the D_{jkw} for each CR^* in the population as shown in Equation 4-2.

$$D_w = \begin{vmatrix} D_{11w} & \dots & D_{1kw} & \dots & D_{1Kw} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D_{j1w} & \dots & D_{jkw} & \dots & D_{jKw} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D_{m1w} & \dots & D_{mkw} & \dots & D_{mKw} \end{vmatrix}$$ **Equation 4-2**

$StDev_w$ is the summation of the standard deviations of rows in the matrix D_w. $StDev_w$ represents the fitness function of the GA optimization that should be minimized in Phase I of the proposed optimization-based method, and is calculated as shown in Equation 4-3 and Equation 4-4.

$$\mathrm{StDev_w} = \sum_{\mathrm{j=1}}^{\mathrm{m}} \sigma[\mathrm{D_{jkw}}] = \sigma[\mathrm{D_{11w}}, \mathrm{D_{12w}}, \dots, \mathrm{D_{1Kw}}] + \sigma[\mathrm{D_{21w}}, \mathrm{D_{22w}}, \dots, \mathrm{D_{2Kw}}] + \cdots$$ **Equation 4-3**

$$Fitness\ Function = Min(StDev_w)$$ **Equation 4-4**

Subject to $CR_{cal\ min} \leq CR^* \leq CR_{cal\ max}$

Where:

m: the number of *DC*s.

The merit of Equation 4-3 and Equation 4-4 over other methods that require the actual measurement of the distances of *DC*s to the hinge is that it does not require these values. Given the difficult accessibility to the excavator's hinge, the calculation of the actual distance to the hinge can be a cumbersome process.

It should be noted that the assumption at this phase is that the captured locations of tags on the upper structure of the excavator, which as explained before are the result of averaging several readings over *dt*, are error-free. However, the locations have residual errors that will be dealt with in Phase II of the method. Figure 4-6 shows an example of the optimization-based method for finding the *CR** in a 2D space over *Δt*, where the rows and columns represent different members of a populations of CR_w* and different averaged locations, respectively. The minimum

and maximum possible CR_{cal} (x_{min}, x_{max}, y_{min}, y_{max}) are represented by the dotted rectangles in Figure 4-6. For instance, in Figure 4-6, $StDev_w$ for the first and second members of a population equals σ[D_{311}, D_{321}, D_{331}]+σ[D_{411}, D_{421}, $D4_{31}$] and σ[D_{312}, D_{322}, D_{332}]+σ[D_{412}, D_{422}, D_{432}], respectively. It is worth noting that the coordinates of CR_w* in each solution in the population, i.e., in each row in Figure 4-6, are fixed.

Once the stopping criteria of GA are satisfied, the *CR** is identified. However, given that in every Δt only a certain number of *RTLS* data (K) is recorded, which are typically much less than the total duration of a non-relocation motion, the *CR** needs to be corrected incrementally. Therefore, *CR** is calculated as a weighted average of the *CR** as measured to the point of the latest Δt and the *CR** measured during the latest Δt, provided no relocation has happened between the intervals. For instance, assuming that (1) the excavator is in non-relocation period, (2) K averaged location data of each tag are obtained over each Δt, and (3) so far n number of data is registered from the beginning of this non-relocation period up to the beginning of this Δt, the coordinates of the progressive *CR* (cx_p, cy_p, cz_p) are defined as shown in Equation 4-5 to Equation 4-7.

$$\mathrm{cx_p} = \frac{\mathrm{n} \times \mathrm{cx_{p-1}} + \mathrm{k} \times \mathrm{cx}}{\mathrm{n} + \mathrm{k}} \qquad \textbf{Equation 4-5}$$

$$\mathrm{cy_p} = \frac{\mathrm{n} \times \mathrm{cy_{p-1}} + \mathrm{k} \times \mathrm{cy}}{\mathrm{n} + \mathrm{k}} \qquad \textbf{Equation 4-6}$$

$$\mathrm{cz_p} = \frac{\mathrm{n} \times \mathrm{cz_{p-1}} + \mathrm{k} \times \mathrm{cz}}{\mathrm{n} + \mathrm{k}} \qquad \textbf{Equation 4-7}$$

Where:

cx_p, cy_p *and* cz_p: The coordinates of *CR** in the non-relocation period from the beginning of the current non-relocation period up to p^{th} interval;

cx, $cy\ and\ cz$: The coordinates of CR^* found by the GA for the data captured over the last interval p; and

n: The number of data registered in the q^{th} non-relocation periods up to the p^{th} interval.

On the other hand, if a period of Δt is identified as relocation, it is determined whether this is the first instance of relocation or a succession of a previously identified relocation. If the beginning of a relocation period is identified, the averaging of CR^* will reset. Lastly, all the coordinates of previous intervals will be adjusted so that they respect the updated CR^* constraint. Although this post-processing step is not effective for the pose estimation for safety-related decisions, it can help improve the quality of data to be used for the identification of state of the equipment and calculating the cycle-time of the excavator operation, as will be explained in Chapter 5.

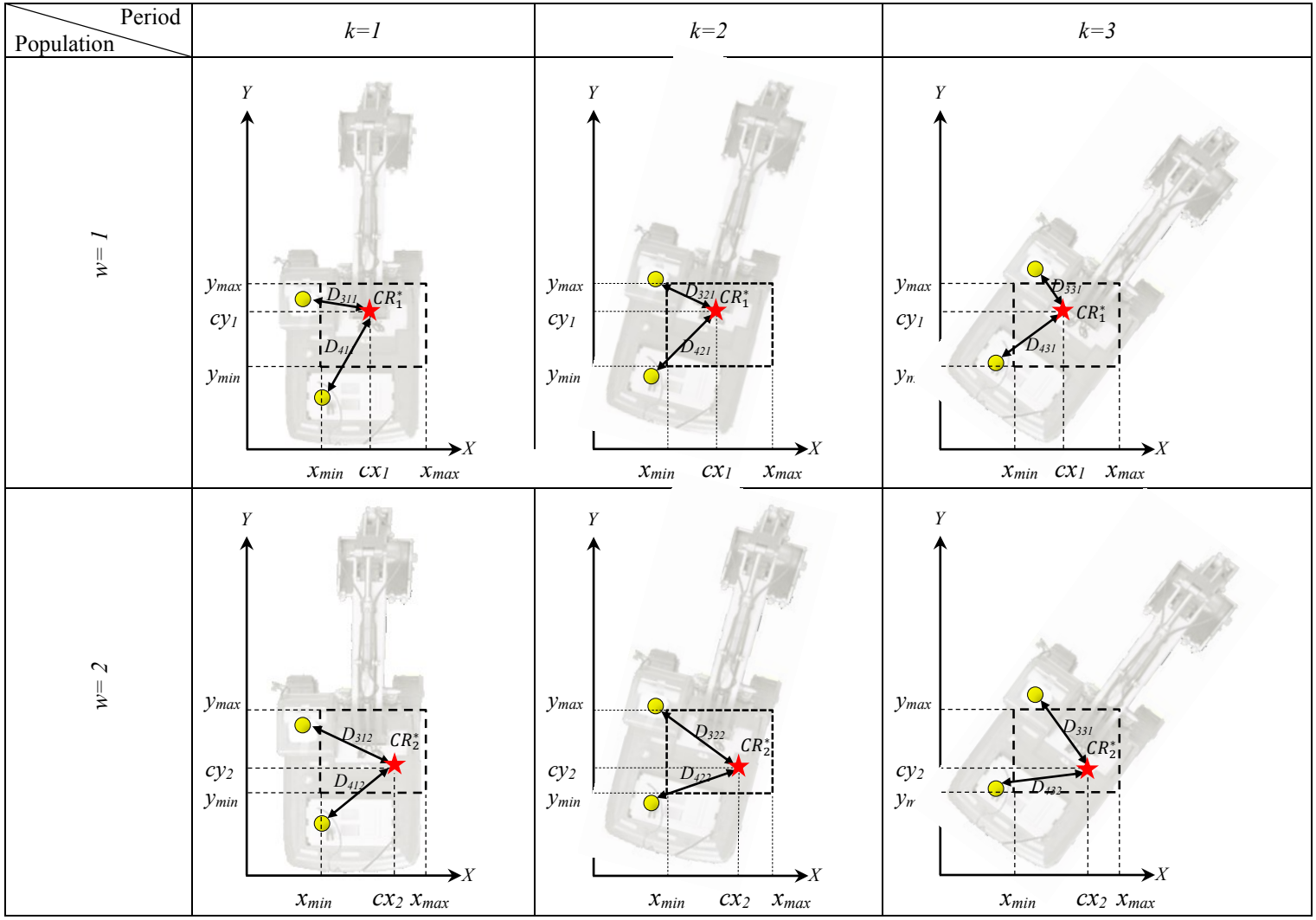


Figure 4-6: Example of Phase I of the Proposed Optimization-based Method

4.2.3.2 PHASE II: IDENTIFICATION OF THE REQUIRED CORRECTIONS

In the face of errors associated with the *DC* data, the proposed correction method, shown in Phase II of Figure 4-5, is committed to determining the minimum amount of correction applicable to each *DC* that will result in a pose with minimum amount of violation of all *GC*s and *OC*s. This converts the problem from a simple error distribution problem, as in the case of simplified correction method, to an optimization problem.

Figure 4-7(a) shows examples of *GC*s that exist between several *DC*s attached to the excavator and Figure 4-7(b) shows the elements used in the calculation of the total violations of constraints (*E*) for this excavator.

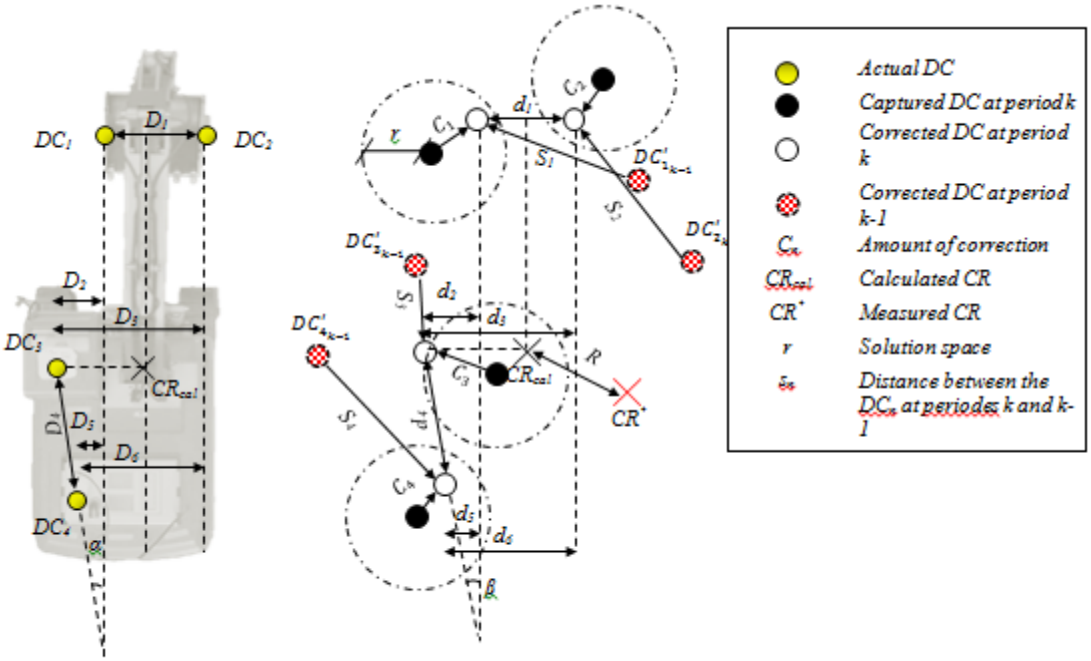


Figure 4-7: Schematic representation of GCs and OCs

Phase II begins with moving back to the first instance of dt_k in the period of Δt. As opposed to Phase I of the method, where the period of Δt was analyzed by considering the entire dataset in one optimization algorithm, Phase II of the method performs a separate optimization for every

dt_k to apply the required corrections to every tag in the period of Δt. As stated earlier, this means that the method is near real-time with the update rate of Δt. Next, the radius of possible corrections in the solution space (r) is set to a threshold defined as the minimum nominal *RTLS* error (e.g., 15 cm). The method proceeds with the generation of an initial population for the amounts of correction (C_j) to be applied to each DC_j. Subsequently, the values of all *GC*s (d_i) and OCs (s_j) are calculated and the overall fitness function is evaluated. These values are used to calculate the summation of violations (E) for this excavator, where E is defined as the summation of the normalized violations from the *GC*s and *OC*s, as shown in Equation 4-8.

$$E = \sum_{i=1}^{u} \frac{|D_i - d_i|}{D_i} + \frac{|\alpha - \beta|}{\alpha} + \sum_{j=1}^{m} \left[\frac{s_j}{s_{max}}\right] + \frac{R}{r} \qquad \textbf{Equation 4-8}$$

where:

E: summation of violations from *OC*s and *GC*s;

i: index representing the number of distance-related *GC*s, $i \in [1, u]$;

D_i: value of a distance-related *GC* as measured;

d_i: value of a distance-related *GC* as calculated (after correction);

α: value of the indicated angle as measured;

β: value of the indicated angle as calculated (after correction);

j: the index representing the ID of DCj, $j \in [1, m]$;

s_j: distance between two consecutive locations of DC_j within dt;

s_{max}: maximum distance between two consecutive locations of DC within dt based on the speed OC;

R: distance between CR_{cal} after correction and CR*; and

r: the radius of possible corrections in the solution space.

The violation from *GC*s is represented by the summation of (1) violations from the distance-related *GC*s, i.e., $\sum_{i=1}^{u} \frac{|D_i - d_i|}{D_i}$, and (2) violations from the angle-related *GC*s, i.e., $\frac{|\alpha - \beta|}{\alpha}$. Similarly, the violation from *OC*s is defined in terms of the summation of (1) the normalized relative distance ($\frac{R}{r}$) between CR_{cal} and *CR**, and (2) the integer value of the ratio between the distance (s_j) and the maximum distance that is allowed by the *OC*s (s_{max}). The reason for the integer value is that if the distance is less than s_{max} it is acceptable because it represents a speed less than the maximum possible speed.

Depending on whether or not the excavator is in relocation, two different types of the fitness function can be defined, where the first type considers the term ($\frac{R}{r}$) in Equation 4-8 and the second type disregards it.

Next, the penalty function *C* is calculated, where *C* is a weight *A* multiplied by the total amount of corrections applied to the *DC*s normalized by r. As shown in Equation 4-9, *C* helps identify the optimum solution with the minimum amount of imposed correction to each *DC*. Equation 4-9 shows the calculation of C.

$$C = A \times \sum_{j=1}^{m} \frac{C_j}{r} \qquad \textbf{Equation 4-9}$$

where:

A: the weight of the penalty function;

C_j: the amount of correction applied to DC_j.

The weight of the penalty function (*A*) is a ratio that defines the relative value of the corrections ($\sum_{j=1}^{m} \frac{C_j}{r}$) versus the summation of violations from OCs and *GC*s (*E*). The greater the value of *A*,

the less inclined is the objective function to reduce the value E at the cost of increased corrections. While a high value of A results in poor performance in terms of ensuring the compliance with *OC*s and *DC*s, a low value may result in high error rate in the detection of the orientation of the equipment.

The fitness function of the optimization in Phase II of the method is defined as the summation of E and C as shown in Equation 4-10. It should be clarified that although the problem can be formulated as a multi-objective optimization problem, where the minimization of errors E and the amount of corrections C are sought, in order to enable the automation of the process, it is favorable to formulate it as a single objective problem that results in a unique pose.

$$Fitness\ Function\ = Min\ (E + C) \qquad \textbf{Equation 4-10}$$

Subject to $C_j \leq r$

If the stopping criteria of the optimization are not met, the optimization proceeds with the GA population re-production routine, i.e., elite selection, crossover and mutation. The optimization continues until the stopping criteria are met. Subsequently, if E of the found solution is more than an acceptable threshold, the optimization will be repeated with an increased radius of the possible corrections in the solution space (r). The solution space is incrementally enlarged, with a user-defined rate (Δr), until it is equal to the maximum nominal *RTLS* error (e.g., 30 cm). If E is greater than a maximum acceptable threshold (E_{max}), the pose for that data point is considered to be undefined. Upon the completion of the correction of one set of data in the time series, the method repeats the optimization until all the data entries are corrected over Δt and then returns to the first phase of the method for a new set of data over the next Δt until the operation is over.

4.2.4 CALCULATING THE POSE

As stated earlier, the pose is defined as the combination of location and orientation of an excavator. Once the errors are corrected, the sensory data can be used to identify the pose of the equipment. The method for the pose calculation depends on the *DOF*s and the configuration of *DC*s on the equipment. Nevertheless, regardless of these conditions, the main ingredient of the pose calculation is the corrected *DC* data of all the rigid elements of the excavator.

It is worth mentioning that, the excavator kinematic chain, which represents the types of joints between different rigid bodies in an excavator, can be used to guide the calculation of the pose of the excavator. A selected set of location and orientation parameters may suffice for this purpose. For instance, if an excavator is equipped with a set of *RTLS DC*s as shown in Figure 4-8, the pose could be represented in terms of the location of a single point on the central axis of the excavator, e.g., DC_1, the orientation of the frame representing the upper structure of the excavator (α, β, Υ), the angle between the upper structure and the base of the excavator (θ_0), and the angles between the several rigid parts of the boom, the stick and the bucket (i.e., $\theta_1 \sim \theta_3$). Figure 4-8(c) shows the angles $\theta_1 \sim \theta_3$ when projected onto a plane parallel to the central axis of the upper structure of the excavator.

The reason for the installation of three *DC*s on the upper structure of excavator, as shown in Figure 4-8(b), is that the determination of the representing frame of the excavator upper structure requires the location of at least three points that are arranged in a non-linear pattern. Therefore, DC_2 is placed so as to form a triangle with DC_1 and DC_3. The remaining *DC*s are placed on the central axis of the excavator's boom. More *DC*s can be added to increase the visibility and reduce the chance of missing data. As a result, it is preferable to use tags attached to the upper structure of the excavator (DC_1, DC_2, and DC_3) as the representing frame rather than those

attached to the base (DC_7, DC_8). The angle between the upper structure and the base (θ_0) can be easily measured using $\overrightarrow{r_{7,8}}$ and $\overrightarrow{r_{1,3}}$, as shown in Figure 4-8(b).

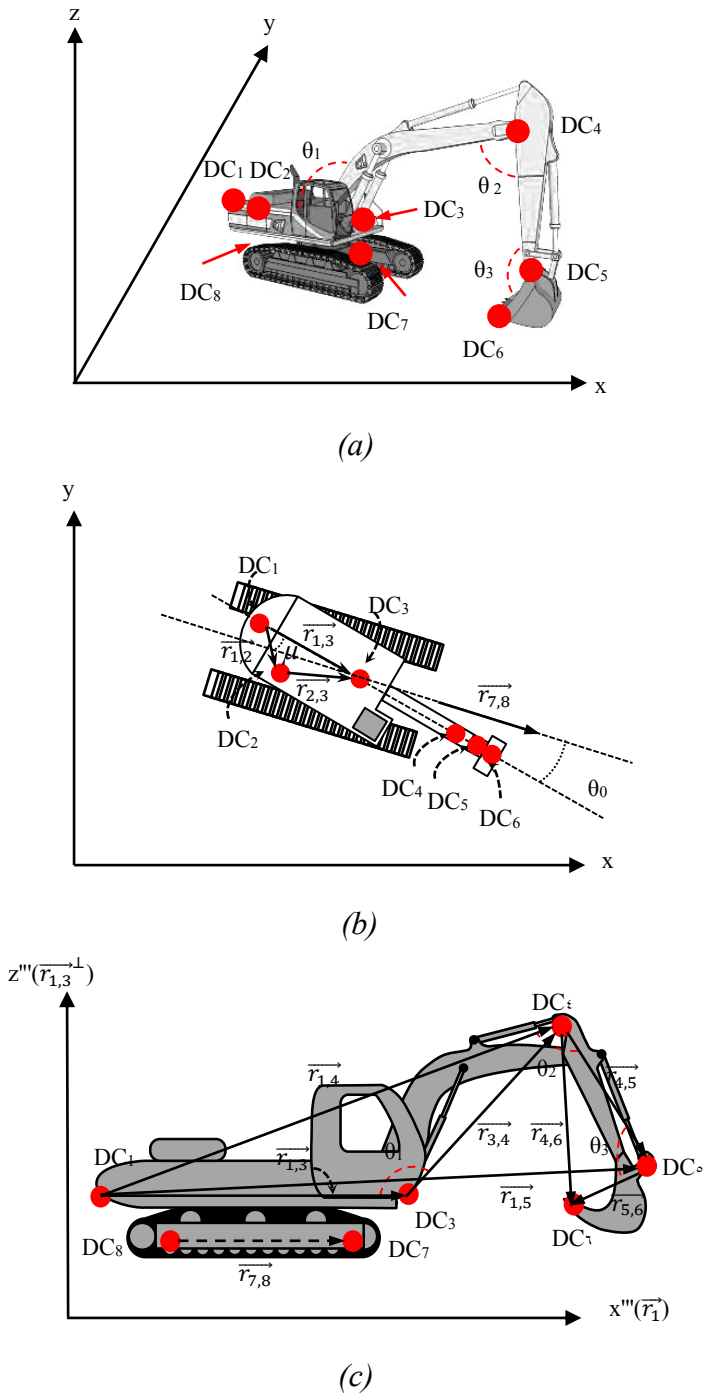


Figure 4-8: Example of Pose Identification for an Excavator Shown in (a) 3D View, (b) Top View, and (c) Parallel to the Central Axis View (The Excavator in Figure (a) is Taken from Trimble 3D Warehouse (2014))

With regard to the orientation of the excavator upper structure, the three angles (α, β, Υ) are calculated through Equation 4-11 to Equation 4-13. Of these angles, α is the angle between the projection of the vector $\overrightarrow{r_{1,3}}$ onto the *x*-*y* plane and *x* axis. β is the angle between $\overrightarrow{r_{1,3}}$ and its projection onto the x-y plane. Υ represents the angle between axis y', i.e., the vector on the horizontal plane of the excavator and perpendicular to $\overrightarrow{r_{1,3}}$, and the perpendicular vector from DC_2 to $\overrightarrow{r_{1,3}}$, as shown in Figure 4-9. In order to calculate Υ, it is required to calculate μ, which is the angle between $\overrightarrow{r_{1,3}}$ and $\overrightarrow{r_{1,2}}$. Figure 4-9 shows the underlying geometric relationships between the vectors $\overrightarrow{r_{1,3}}$, $\overrightarrow{r_{1,2}}$ and $\overrightarrow{r_{2,3}}$, the excavator frame, and the reference frame.

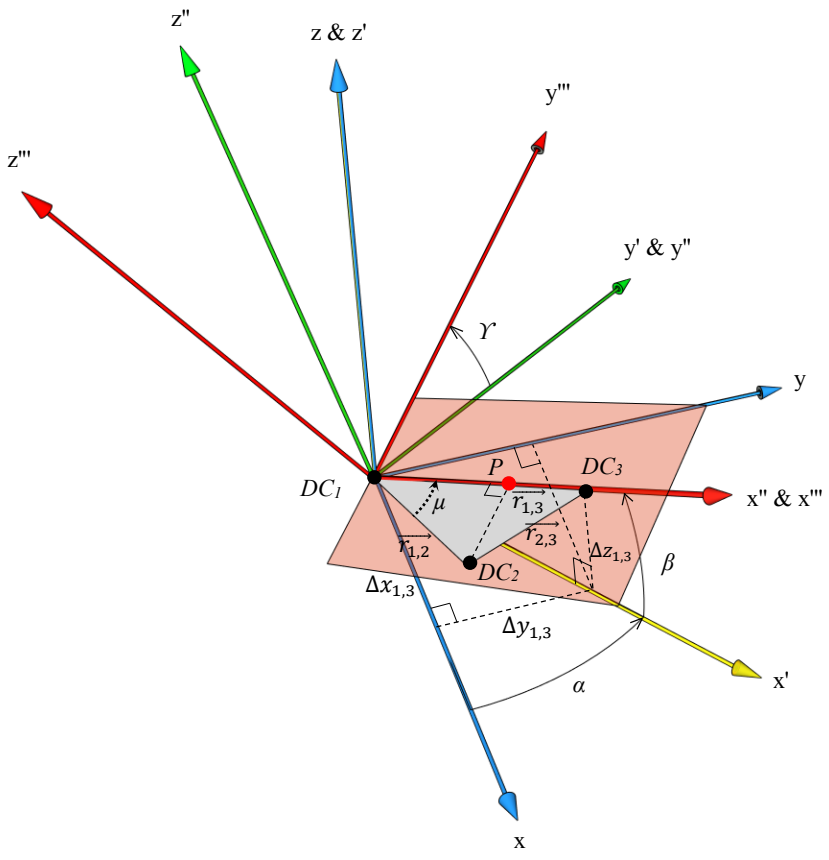


Figure 4-9: Orientation of the Excavator Body Using Three Vectors

$$\alpha = \cos^{-1}\left(\frac{\Delta x_{1,3}}{\sqrt{(\Delta x_{1,3})^2 + (\Delta y_{1,3})^2}}\right) \qquad \textbf{Equation 4-11}$$

$$\beta = \sin^{-1}\left(\frac{\Delta z_{1,3}}{\left\|\overrightarrow{r_{1,3}}\right\|}\right)$$ **Equation 4-12**

$$\gamma = \sin^{-1}\left[\frac{\Delta z_{1,2} \times \cos\beta - (\Delta x_{1,2} \times \cos\alpha + \Delta y_{1,2} \times \sin\alpha) \times \sin\beta}{\left\|\overrightarrow{r_{1,2}}\right\| \sin\mu}\right]$$ **Equation 4-13**

The angles θ_0~θ_3 can be easily determined. For instance, the calculation of θ_1 in Figure 4-8 is presented in Equation 4-14.

$$\theta_1 = \cos^{-1}\left(\frac{\left\|\overrightarrow{r_{1,3}}\right\|^2 + \left\|\overrightarrow{r_{3,4}}\right\|^2 - \left\|\overrightarrow{r_{1,4}}\right\|^2}{2 \times \left\|\overrightarrow{r_{1,3}}\right\| \times \left\|\overrightarrow{r_{3,4}}\right\|}\right)$$ **Equation 4-14**

4.3 IMPLEMENTATION

As a proof of concept, a module is developed using Visual Basic for Applications (*VBA*) in Microsoft Excel, as shown in Figure 4-10. This module is capable of connecting to two GA-based optimization models developed in Matlab (Mathworks 2015), and collecting the output value of *CR* and the corrected location data for an excavator. While the Standard GA used in Phase I is structured to find the optimum solution based on a population size of 100 generations, the number of generations in the standard *GA* used in Phase II is set to 400. In order to reduce the chance of the local optima phenomenon, the Gaussian mutation with the scale factor of 7 is used. Additionally, the hybrid function *fminsearch* is applied to the final results from the *GA* to increase the optimization accuracy (Mathworks 2015). As a stopping criterion, the tolerance of e^{-20} is specified.

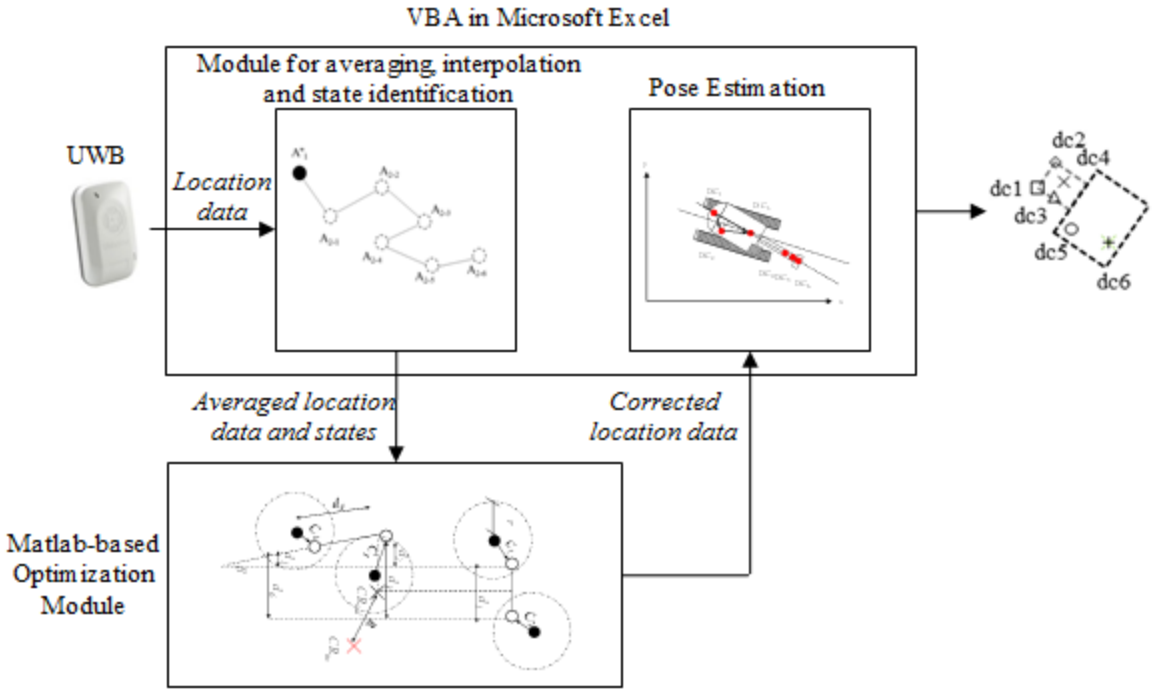


Figure 4-10: Architecture of the Implementation for Pose-estimation

The input data, presented in the form of time-stamped Cartesian coordinates of the *DC*s, are fed to the module and the averaging over a period of time and interpolation are performed. Upon the completion of this step, the *VBA* module invokes Matlab and runs the required optimization. This module determines the status of the equipment, and if the non-relocation period is identified, the associated optimization model for *CR* is invoked.

Finally, it should be elaborated that the pose estimation of every period of *dt*, took approximately 5.5 s and 3.5 s on a personal computer with Intel Core i7-2600 CPU 3.40 GHz for the optimization based on data from 6 and 4 tags attached to the scale model excavator, respectively.

4.4 CASE STUDIES

Two case studies were conducted to demonstrate the feasibility of the proposed optimization-based method. Figure 4-11 illustrates the equipment used in the case studies and the location of tags. All pieces of equipment are radio-controlled (*RC*) using remote controls with different

buttons and joysticks that allow the movement of one *DOF* at a time. As for the *DOF*s of the equipment, the truck is empowered by two motors which propel the body (drive forward/backward, turn right/left) and the bed of the truck (up/down). The excavator, on the other hand, is able to move the body (drive forward/backward, turn right/left), the boom (swing and move up/down) and the stick (move in/out) using its five motors (Setayeshgar et al. 2013). The reason for selecting this environment is that it can be fully controlled and can be repeated as many times as necessary. The crane was kept stationary as will be explained in the second case study.

The case studies were developed based on the 2D location data because the utilized *UWB* has a lower accuracy in the *z* axis than in the x and y axis, which when combined with the small scale of the excavator used in the case studies, renders the z values unreliable.

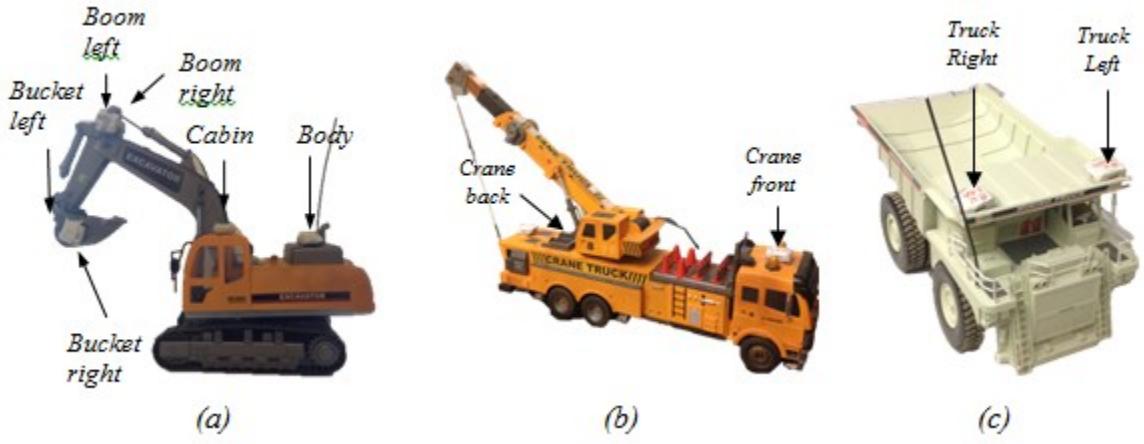


Figure 4-11: Equipment Used in the Case Studies and Location of the Attached Tags

Furthermore, two other outdoor case studies were conducted on an actual construction sites for the purpose of roller and excavator pose estimation. The description of these case studies and the analysis of the results are presented in Appendix B. These field case studies have been done in a collaborative project with Siddiqui (2014).

4.4.1 FIRST CASE STUDY

In the first case study, a test was designed to validate and verify the proposed optimization-based method. The *RC* excavator is employed to simulate the swinging task of the equipment. As shown in Figure 4-11(a), six tags were strategically attached to different components of the excavator for the collection of location data. Two tags were attached to every part of the excavator for which location data is required in order to provide redundancy. However, as explained in Section 4.2.4 and Figure 4-8, three *DC*s are required to be attached to the upper structure in order to obtain the pose of the excavator in 3D. The other pieces of equipment shown in Figure 8 are used in the second case study.

For the tracking purpose, Ubisense *UWB* technology, with a nominal accuracy of 15 cm in ideal conditions, was utilized (Ubisense 2015). In order to evaluate the proposed optimization-based method, it is required to calculate the accuracy of the *UWB* system under the case study environment. For this purpose, a static test was performed in which a single tag was placed at the location of the excavator's CR and its actual location was compared to the averaged location data from the *UWB* system. In this case study, as shown in Figure 4-12, the excavator was working near the edge of the room. Based on this test, the accuracies corresponding to 90% and 95% confidence levels are 45 cm and 48 cm, respectively.

A Sony Network Rapid Dome camera with the update rate of 30 frames per second was used to record the video of the test to provide a ground truth for comparison. In order to enable an accurate validation of the pose estimation method, the setting of the pan and tilt angles of the camera were adjusted so as to achieve a view which has two characteristics: (1) the view plane is parallel to the room floor, and (2) the edges of the view plane are parallel to the edges of the room, as shown in Figure 4-12.

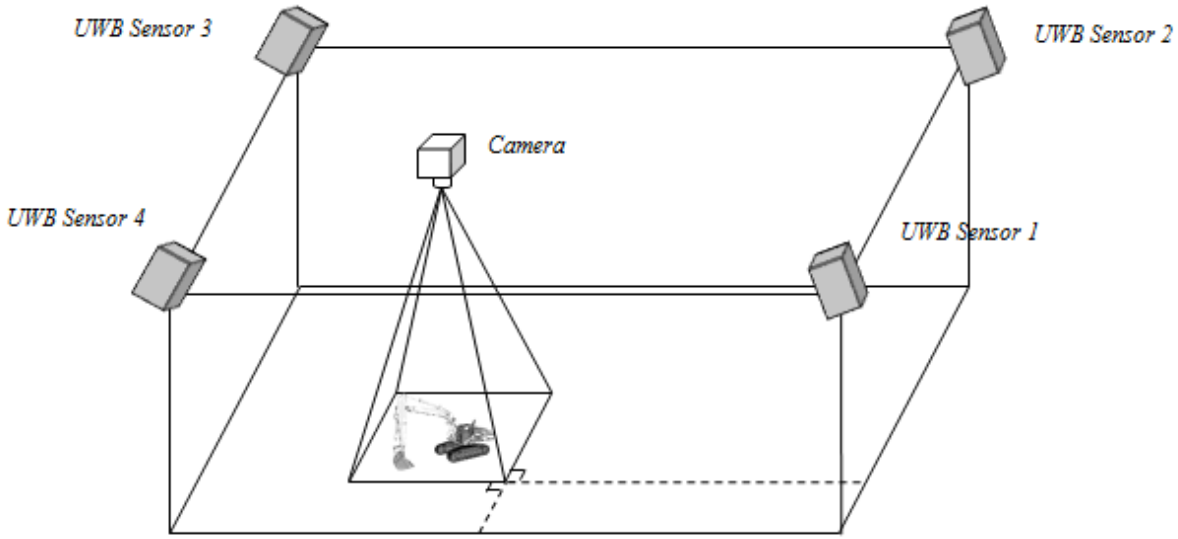
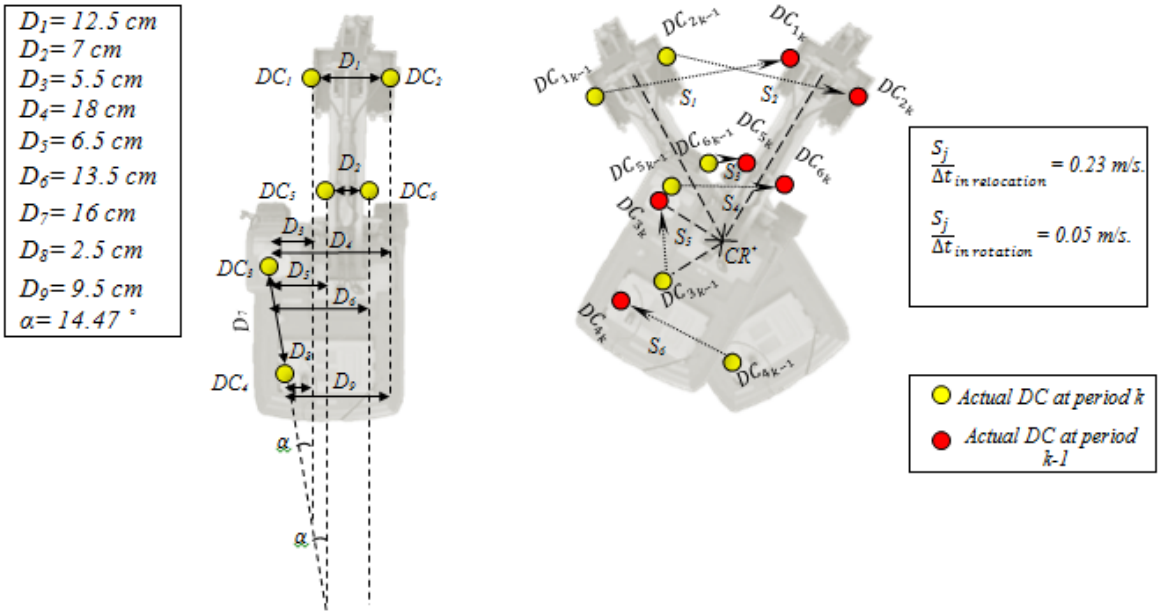


Figure 4-12: Adjustment of Camera View

The excavator was controlled to perform two swings, one counter-clockwise and one clockwise, over a period of 91 s. The update rate for the *UWB* was set to 17 readings per second. The experiment corroborated this update rate by showing the average rate of 16.04 reading per second.

In the implementation of the proposed method, *dt* was set to 500 *ms*, the minimum and maximum *RTLS* errors were set to 20 *cm* and 50 *cm*, respectively, and the threshold for the error after correction E_{max} was set to 5. Also, the value of the penalty weight *A* in Equation 4-9 was set to 0.1. These values were determined based on several rounds of trial and error for identifying the best combination of parameters. The best combination of the parameters is defined based on the lowest error in the estimation of the location of *CR* and the orientation of the equipment. Given that the purpose of this case study was to validate the accuracy of the pose estimation method, Δt was set as the entire length of the test, i.e., 91 *sec.* The *GC*s and *OC*s applied to this case study are shown in Figure 4-13.



(a) Geometric Constraints (b) Operational Constraints

Figure 4-13: Constraints Used in the First Case Study

As no relocation has taken place during this test and Δt was defined as the entire length of the test, *CR* was found using the algorithm discussed in Section 4.2.3.1 over all the captured data. The proposed optimization-based correction method, discussed in Section 4.2.3.2, was applied to all data entries after averaging over *dt* at three different levels: (1) considering only *GC*s, (2) considering *GC*s and OCs, and (3) considering *GC*s, OCs, and *CR*. Of the 182 *UWB* averaged location data, corresponding to *dt*= 500ms, in two cases the algorithm was not able to find a pose. The orientation of the equipment was measured in 2D, meaning that of the angles shown in Figure 4-9 only α was calculated.

The location error was calculated based on the comparison of the known location of *CR* and the location of *CR* as calculated before applying any correction, applying simplified correction method, and after the three levels of optimization-based correction mentioned above. Figure 4-14

shows the cumulative distribution of location errors for the five scenarios and the scenario with static tag. As expected, the location error of *CR* before correction is approximately equal to the location errors calculated through placing a static tag at the location of *CR*, as explained above. Contrary to the expectations, the simplified method of correction had adverse impact on the location accuracy. This can be attributed to the fact that the angular *GCs* were not captured by this method. However, the application of each type of correction resulted in a degree of improvement in the accuracy of the location data. For example, when all constraints are applied, the error corresponding to 90% and 95% confidence levels are equal to 16.8 *cm* and 18.2 *cm*, respectively. This shows more than 60% improvement in the accuracy of location estimation when compared to the results of the data with before correction. The sharp slope of the error distribution of data after applying the correction with all constraints is due to the fact that the proposed optimization-based method is adjusting the locations of *CR* based on the location found in the first phase, explained in Section 4.2.3.1, and this had 16.4 *cm* errors. Therefore, all the *CR*s are centered very tightly around the found CR location.

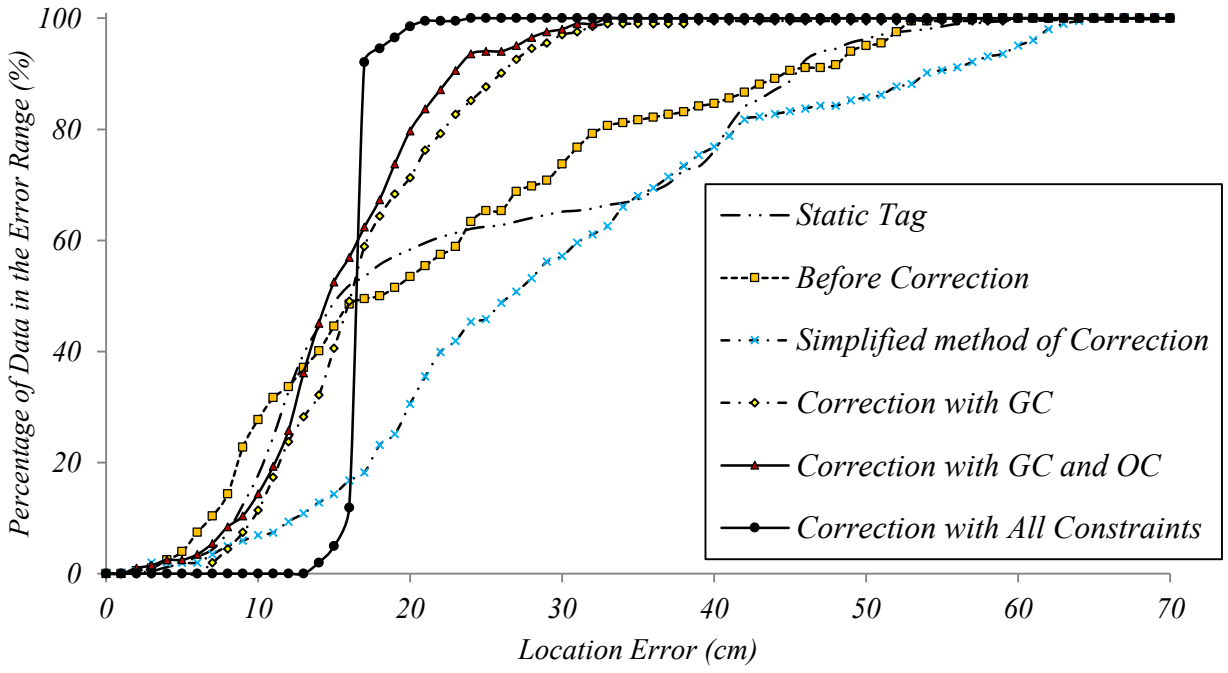


Figure 4-14: Cumulative Distribution of Location Error for Different Levels of Correction

For the validation of the estimated orientation, the captured video of the test was used. For this purpose, the video is decomposed into frames, and one frame per second was chosen for measuring α. Each frame was imported into AutoCAD Civil 3D (AutoDesk 2014) in order to measure α. Since the video is time-stamped, each frame and the corresponding value of α were used for the comparison with the orientations from the corrected *UWB* data.

Figure 4-15 shows three frames obtained from the test with their corresponding corrected *UWB* data, orientations and calculated errors for the case where all constraints are considered. A statistical analysis is performed in order to determine the accuracy of the estimated orientation at the simplified method of correction and, the three levels of optimization-based correction. For this purpose, the orientations (before and after correction) from a sample of 91 *UWB* data points, i.e., 50% of the available dataset were plotted, as shown in Figure 4-16. It is worth mentioning that the swinging speed of the scale equipment is low compared with this speed in the case of actual excavators. With the actual equipment, if the high update rate of *RTLS* is maintained, high swinging speed of the equipment is not expected to affect the performance of the proposed method.

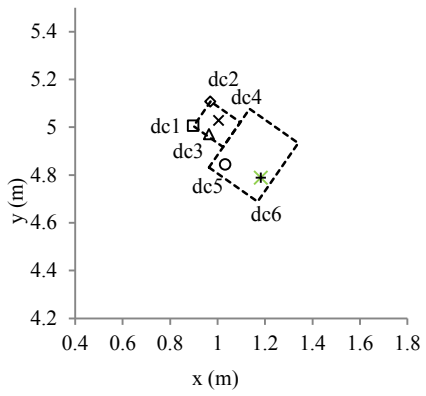
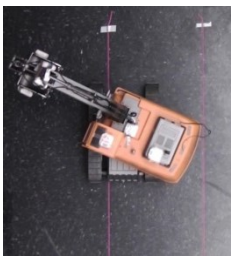
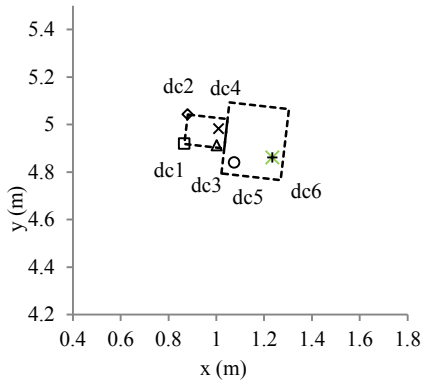
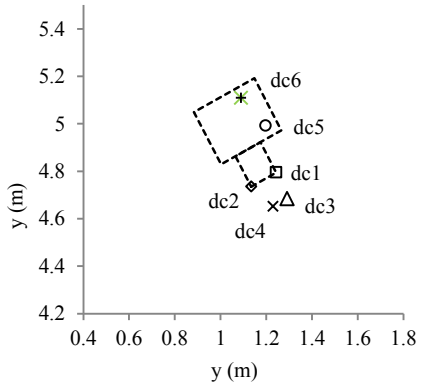
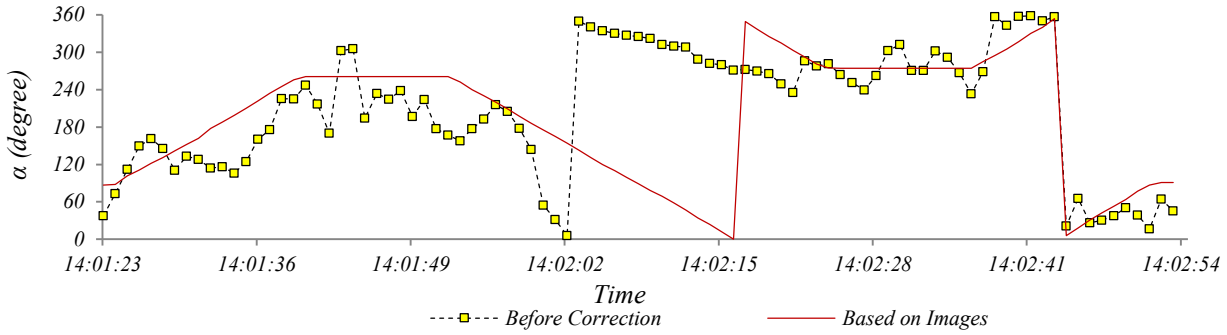
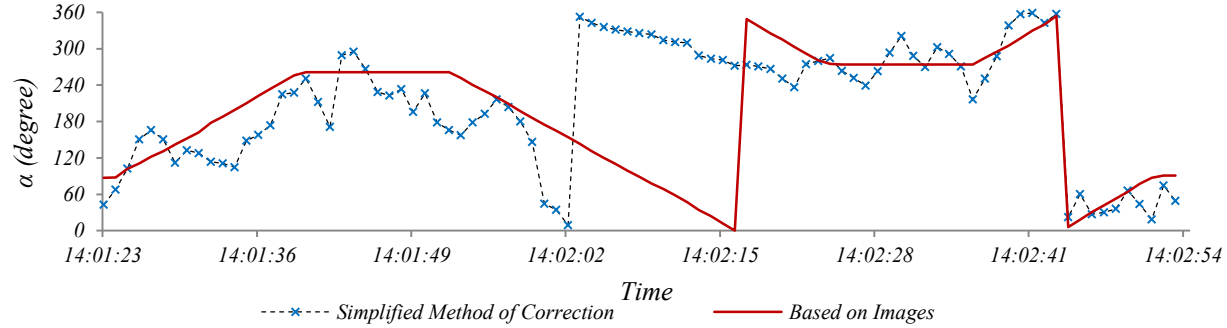
Time	Plot of Corrected UWB Data	Corrected Orientation	Corresponding Frame from Video	Orientation from Images	Error
14:01:30		153.42˚		152˚	1.42˚
14:02:00		178.39˚		175˚	3.39˚
14:02:30		287.1˚		274˚	13.1˚

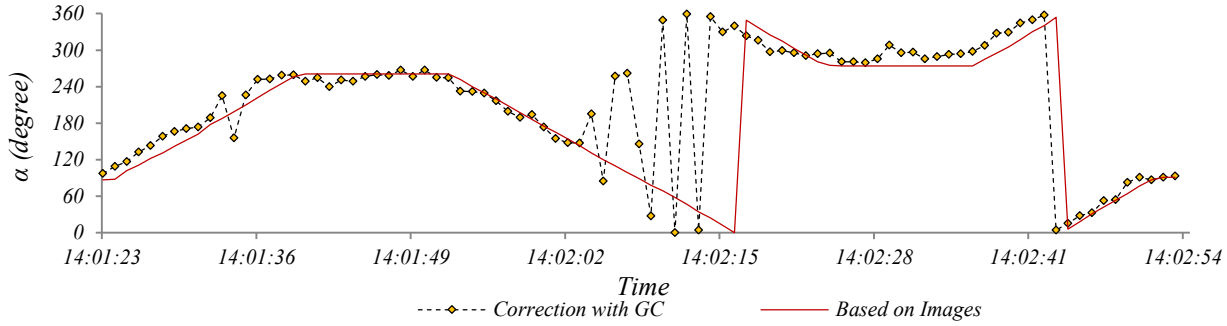
Figure 4-15: Comparison of the Result of Pose Estimation



(a) Comparison of Before Correction and Data from Images

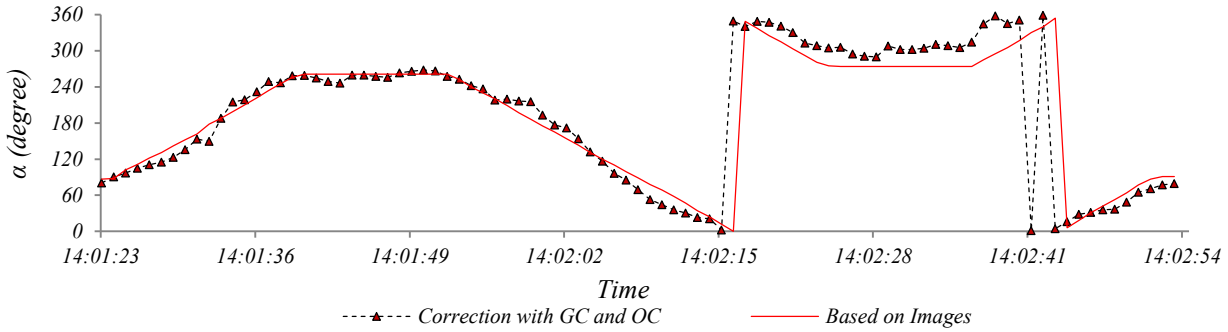


(b) Comparison of simplified Correction and Data from Images

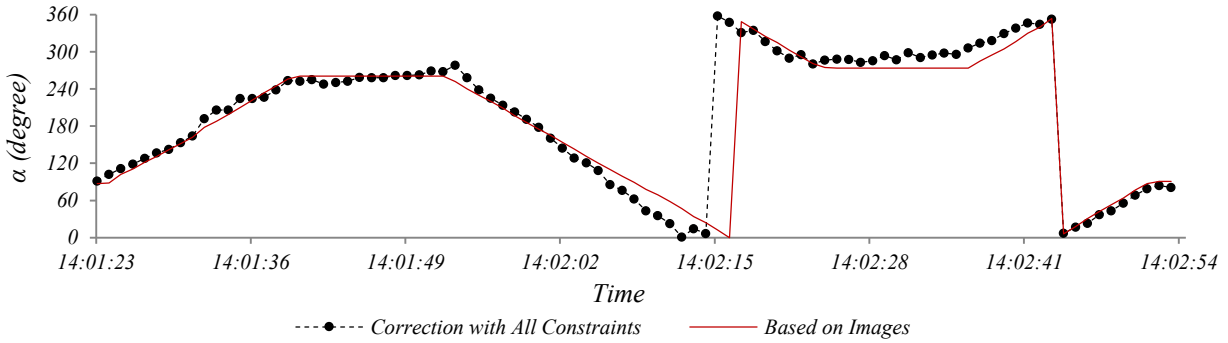


(c) Comparison of Optimization-based Correction with GC and Data from Images

Figure 4-16: Comparison of the Orientation Based on the UWB Data Before and After Correction with the Orientation Based on Images



(d) Comparison of Optimization-based Correction with GC and OC and Data from Images

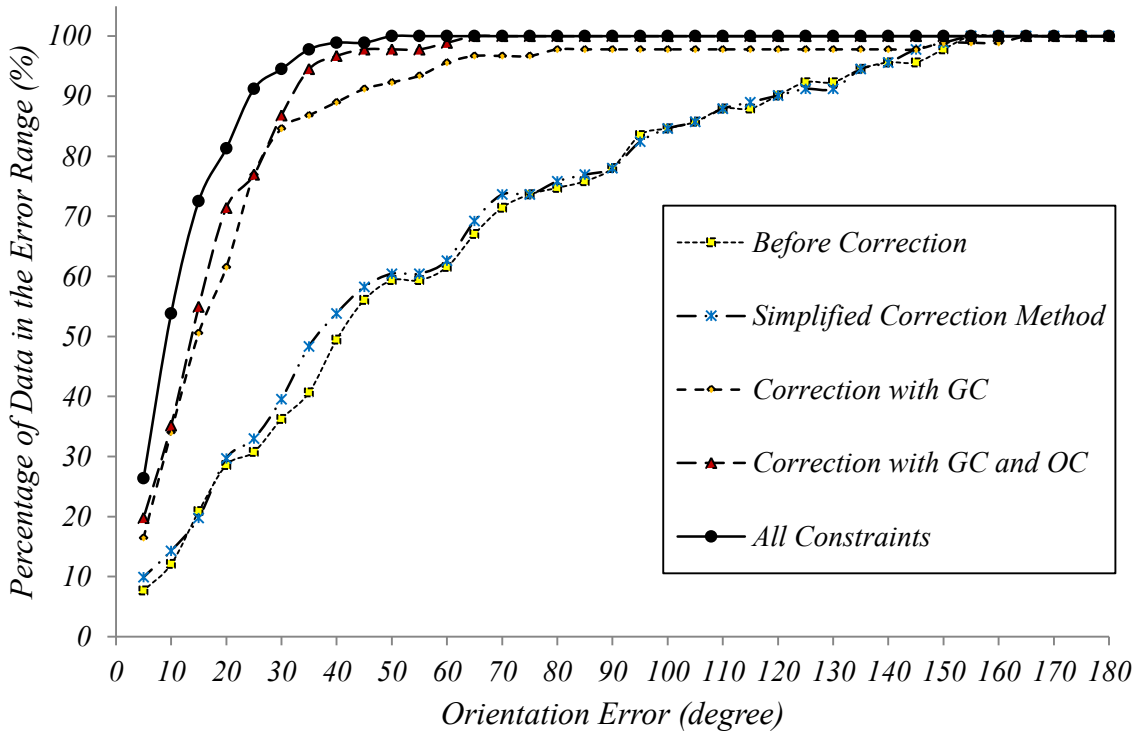


(e) Comparison of Optimization-based Correction with all Constraints and Data from Images

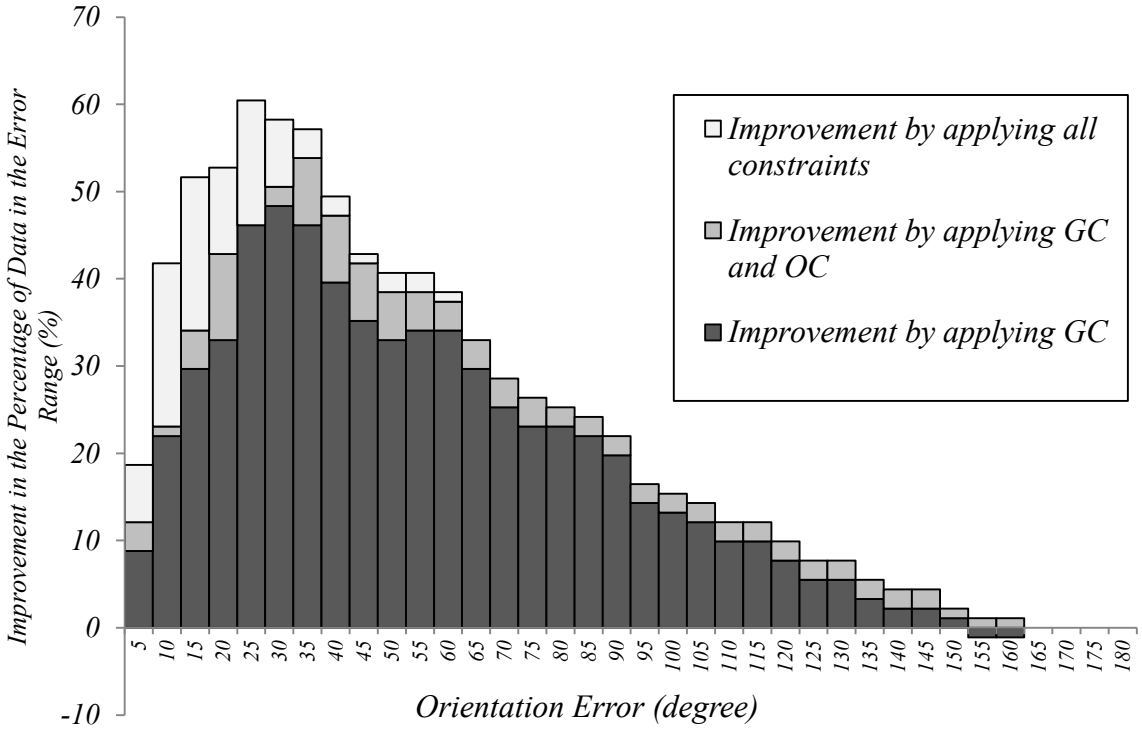
Figure 4-16: Comparison of the Orientation Based on the UWB Data Before and After Correction with the Orientation Based on Images (Cont.)

Figure 4-16(a) compares the orientation of the *UWB* data before the correction with the ground truth obtained from the analysis of images. Similarly, Figure 4-16(b) compares the orientation of the *UWB* after the simplified method of correction with the ground truth. Figure 4-16(c), (d), and (e) illustrate the same comparison for the cases where only *GC*s, the combination of *GC*s and *OC*s, and all the constraints are considered in the correction method, respectively. Based on the comparison of the five figures, it can be observed that while the simplified correction method did not improve the orientation accuracy, the addition of each type of constraint (i.e., *GC*s, *OC*s and *CR*) to the optimization-based method has contributed to reducing the orientation error to a certain degree. As can be seen in Figure 4-16(e), the proposed method of optimization-based correction improved the accuracy of estimation considerably. Note that in Figure 4-16 what seems to be a great error around angles 0˚ and 360˚ is in fact a small fluctuation around angle 0˚.

Figure 4-17(a) shows the cumulative distribution of errors in five different scenarios, namely (1) before correction, (2) after the simplified correction, (3) after optimization-based correction considering only *GC*s, (4) after optimization-based correction considering *GC*s and OCs, and (5) after optimization-based correction considering all constraints. The *X* axis in Figure 4-17 represents the absolute value of the boundaries of an error range, i.e., δ for the range of error between [$-\delta, \delta$] and the *Y* axis represents the percentage of data within that range.



(a) Cumulative Distribution of Error for Different Levels of Correction



(b) Improvement After Each the Addition of Each Level of Optimization-based Correction

Figure 4-17: Distribution of Errors in Orientation Estimation

According to Figure 4-17(a), only 20.9% of the estimated orientations are within an accuracy of ±15˚ before the correction. However, the percentage of data within the same accuracy increased to 72.5% when all constrains were applied using optimization-based correction. Similarly, the errors associated with 90% and 95% confidence levels improved from ±119.75˚ and ±137.25˚, before correction, to ±24.39˚ and ±30.75˚, after optimization-based correction, respectively. This corresponds to more than 77% improvement in the estimation of orientation. Nevertheless, the simplified correction method only marginally improved the accuracy of orientation estimation. It can also be discerned from Figure 4-17(a) that each type of constraints in optimization-based method improved the data to a certain degree.

To better put the contribution of each type of constraints in perspective, the percentage of data added to each error range after the introduction of each type of constraint to the optimization-based correction is plotted in Figure 4-17(b). For example, after applying the correction considering only *GC*s, 29.7% of the data improved their accuracy to ±15˚. After adding the OCs to the correction method, 4.4% was added to the data within this accuracy. Finally, by adding the CR to the correction method, there is another 17.6% increase in the percentage of data with the accuracy of ±15˚.

Another interesting observation is that the average amount of violations from distant-related *GC*s and angle-related *GC*s before correction are 21 cm and 25.23˚, respectively. However, after optimization-based method, these average violations decreased to 0.09 cm and 0.09˚, which represents a high degree of compatibility with the geometry of the equipment. As for the *OC*s, the raw *UWB* data indicated 163 instances of violations of the *OC*s. This number has dropped to 8 instances after the optimization-based correction with all the constraints.

It should be noted that the orientation accuracy is relative to the scale of the equipment, and given that a small piece of equipment is used in this case study, it is expected that in a setting with an actual equipment the accuracy would increase considerably.

4.4.2 SECOND CASE STUDY

In the second case study, the data from another lab test was used to further study the proposed method in the context of an operation. This case study pursues two objectives: (1) Show that the geometric integrity of the excavator is preserved through the proposed optimization-based method even when the excavator is relocating; and (2) Demonstrate that the smooth rotational movement of the excavator during non-relocation movement is captured, even after the relocation of the excavator.

This test was carried out in the laboratory environment, where the three *RC* machines shown in Figure 4-11 were utilized to simulate a simple loading-hauling-dumping operation. The three pieces of equipment are equipped with *UWB* tags. Careful scrutiny was carried out as to identify how many tags to use and where to attach the tags to get the best results. Every part of the equipment for which location data is required is represented by two tags, in order to reduce the errors. A total of eight tags, four for the excavator, two for the truck and two for the crane, were used. Except the elimination of tags on the boom, i.e., DC_5 and DC_6, the location of tags and the *GC*s and *OC*s used in the second case study are identical to those of the first case study shown in Figure 4-11 and Figure 4-13. The elimination of the two tags from the boom was necessary in order to keep the update rate of *UWB* as close as possible to the first case study. The truck is not considered in the scope of this chapter because it was only equipped with two tags, which have only one *GC*. In this case, the simplified correction method is sufficient for the pose estimation.

The case study covered four full cycles of loading and hauling, which consisted of loading, hauling, dumping and returning operation for the truck and the loading, swinging, dumping, swinging back and relocation operations for the excavator. The crane is used only to show that this system can still work in a congested site where several teams are working on different operations. Zones within which each of the afore-mentioned operation takes place were marked and their corresponding coordinates were measured. These zones later establish the base for the state identification of the equipment, as will be shown in Chapter 5.

Figure 4-18 shows this setting of the case study which contains four main parts, namely, the excavation area, loading area, hauling area and the dumping area. The length and width of the simulated site are 7 m × 3.5 m, respectively. The truck is loaded in the loading area by the excavator; it moves to the dumping area and dumps its load. On the other hand, the excavator obtains a load from the borrow pit, swings to the loading area, dumps the material into the truck and swings back. The excavation area is divided into two work spaces. After two cycles of the truck, the excavator relocates to a new work space within the same excavation zone, which is marked by *WS2* in Figure 4-18(a), and continues to load the truck for two more cycles. With the intention to create a congested site, a crane was placed near the excavator without actively engaging in the simulated operation. Also, upon the completion of the fourth loading cycle, the excavator was intentionally steered to collide with the crane, to evaluate the effectiveness of the proposed *DEW* in preempting the potential collision, as will be explained in Chapter 6. The test included four cycles of loading-and-dumping, which took 4 minutes and 45 seconds. The update rate was set to 9.6 readings per second. However, after the analysis of the test results, the actual average update rate was found to be about 8 readings per second. The proposed correction

method was applied to the excavator with the same setting as in the first case study except *dt* which was set to 1 s and Δt was set to 5 s.

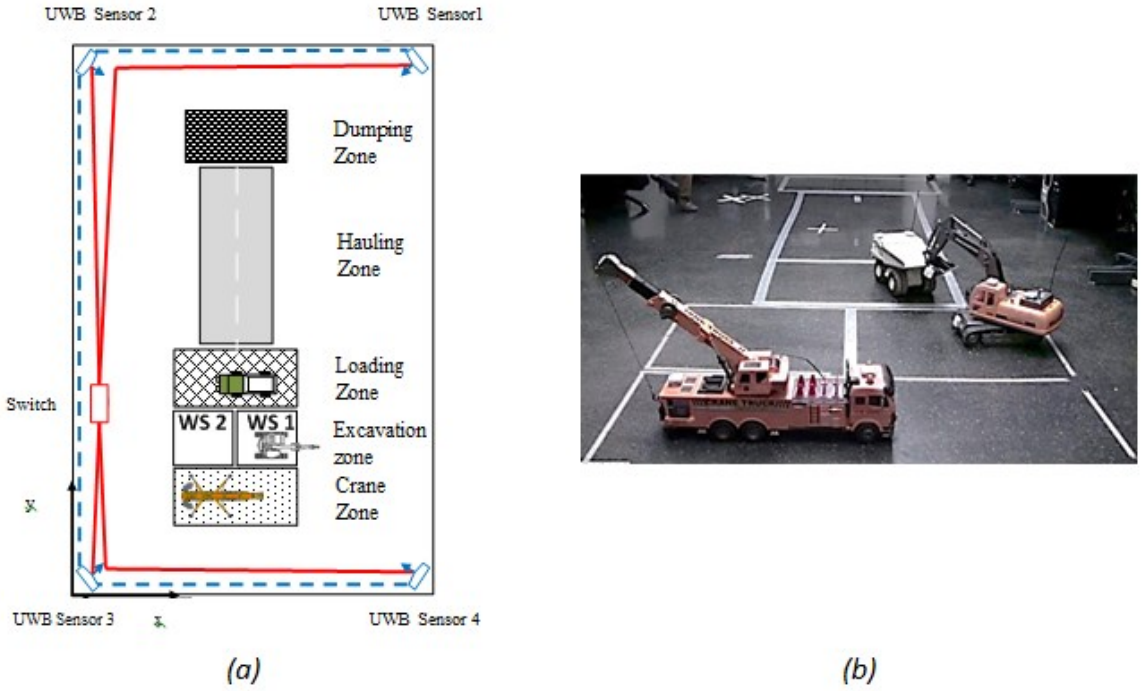


Figure 4-18: (a) Schematic Layout and (b) Snapshot of the Site Layout in the Second Case Study

The truck is not considered in the scope of this chapter because it was only equipped with two tags, which have only one *GC*. In this case, a simple averaging method is sufficient for the pose estimation. The results of the correction for each *UWB* tag attached to the excavator, i.e., DC_1 to DC_4, over the entire period of the case study are presented in Figure 4-19.

As can be seen in Figure 4-19(c) and (d), the proposed optimization-based method managed to preserve the rotational motions with minimal amount of correction applied to each tag. Unlike Figure 4-19(c) and (d), the smooth curve did not appear in Figure 4-19(a) and (b). This is because DC_1 and DC_2 can move towards or away from the base, whereas DC_3 and DC_4 are attached to the upper structure of the excavator that can only rotate around *CR*. Presenting the

results of several instances of the extended correction method, Figure 4-20 depicts the effectiveness of the method in preserving the geometry of the equipment throughout the case study. The analysis of the location estimation was not possible in the second case study because the excavator was moving, which makes the measurement of the ground truth difficult.

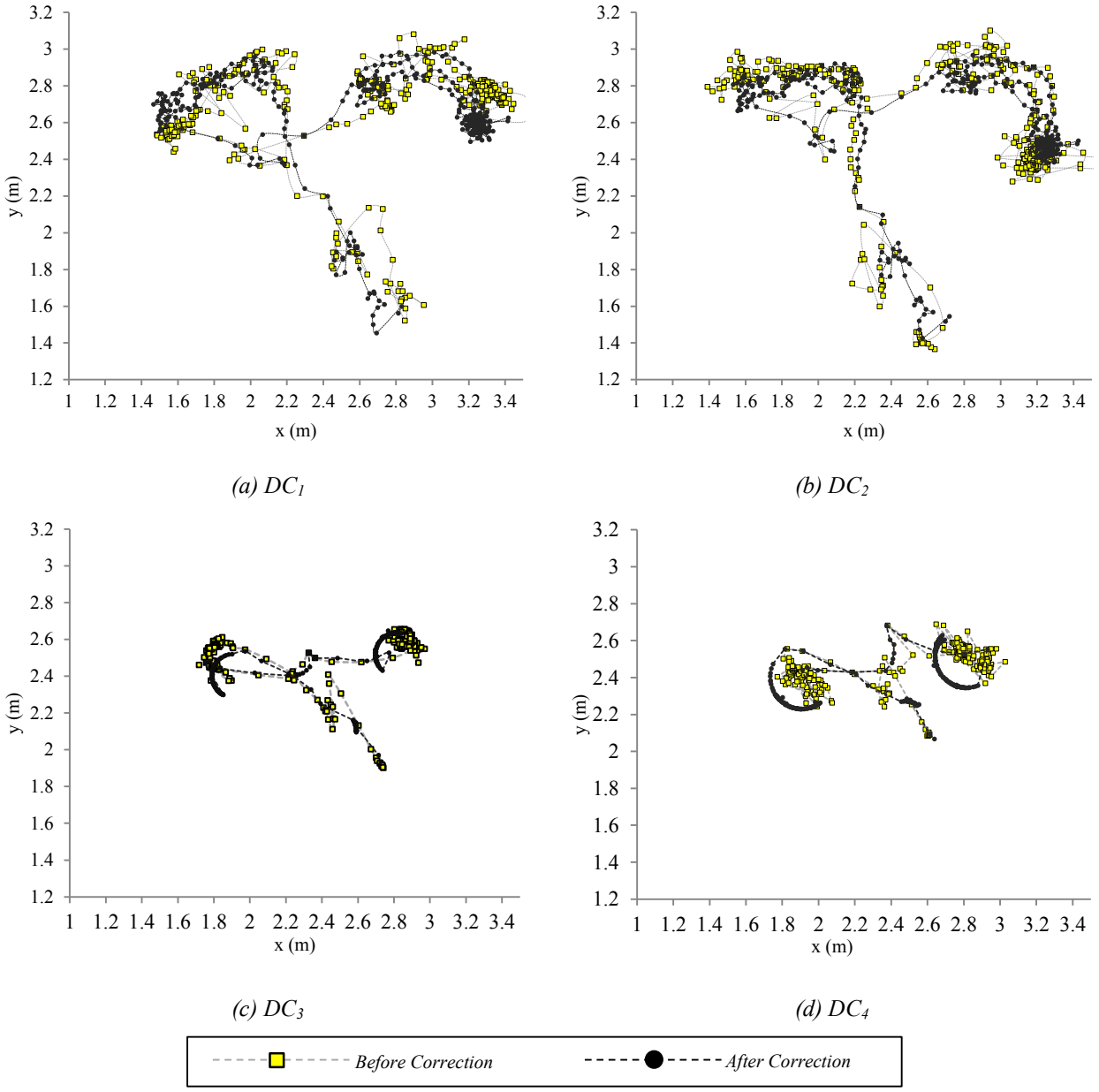


Figure 4-19: Effect of the Correction Method on Different Tags over the Entire Period of Case Study

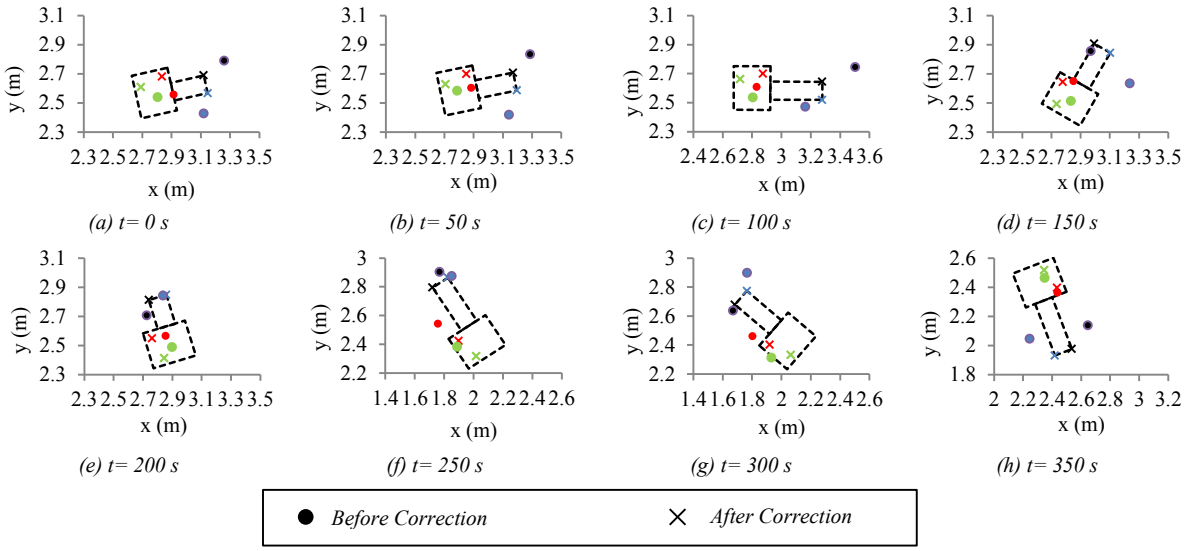


Figure 4-20: Results of the Correction Method

In order to demonstrate the improvement caused by the proposed optimization-based method in the orientation estimation, the orientation data before correction and after correction were compared to the results extracted from the video of the test, as shown in Figure 4-21.

Given that in this case study the excavator was not stationary and the size of the test bed was considerably bigger than the first case study, the adjustment of the camera location for the full top view of the operation, as shown in Figure 4-12, was infeasible. However, Matlab's spatial transformation from control point pairs (Mathworks 2015) was used to apply the required transformation on the images obtained from the video of the operation so that the orientation of the excavator can be accurately extracted from the images. It is clearly demonstrated in Figure 4-21 that the proposed optimization-based method improved the orientation estimation. To better put the level of improvement in perspective, the cumulative distribution of orientation errors before and after simplified and optimization-based corrections is presented in Figure 4-22. This Figure suggests that while only 34.9% of data were within the accuracy of ±15˚ before correction, the percentage improved to 62.7% after applying optimization-based correction. As

for the simplified method of correction, the percentage of data Similar analysis to case study one suggests that errors associated with 90% and 95% confidence levels improved from ±43.93˚ and ±54.16˚, before correction, to ±35.87˚ and ±41.32˚, respectively. This is equivalent to more than 18% improvement in the estimation of the orientation.

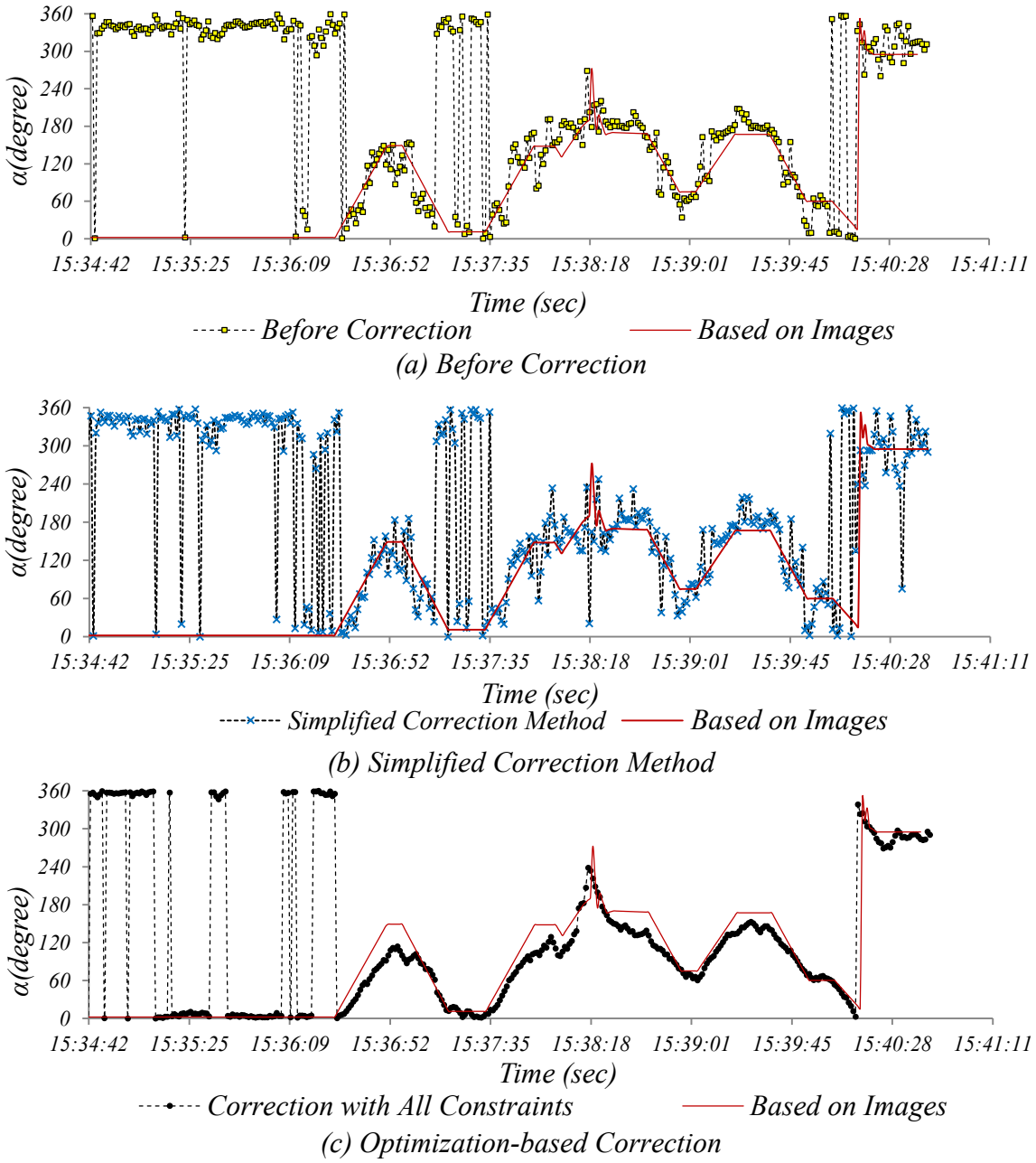


Figure 4-21: Comparison of the Results of Orientation Estimation

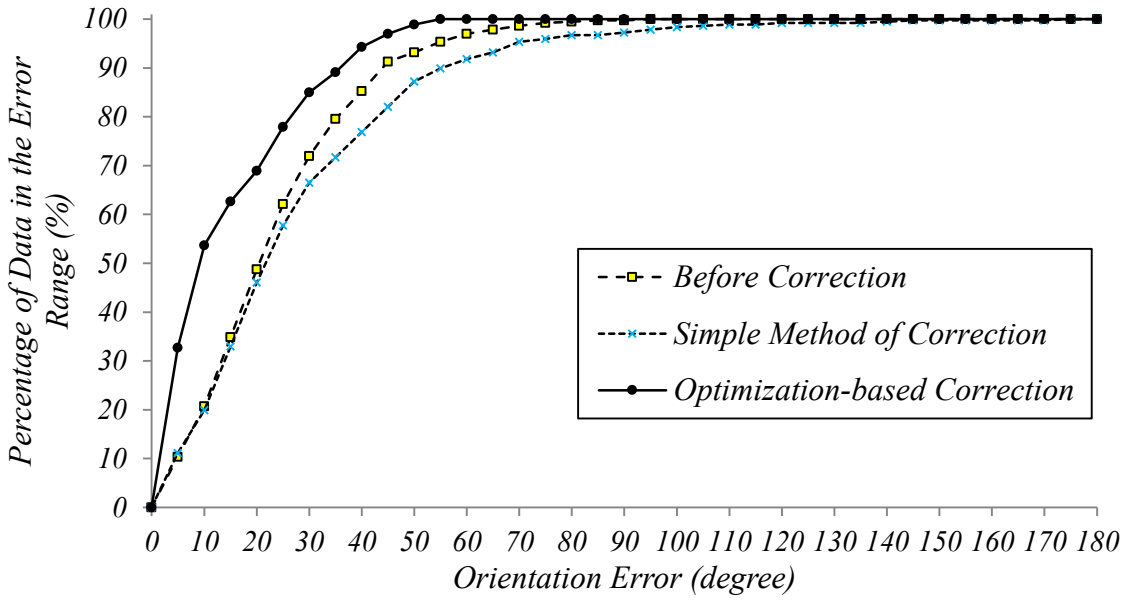


Figure 4-22: Cumulative Distribution of Orientation Error Before and After Correction

The improved quality of the pose estimation combined with a safety buffer around the equipment can help identify safety hazards. For instance, near the end of the test, the excavator was deliberately moved near the stationary crane to simulate a potential collision. As shown in Figure 4-23, a warning can be triggered when the DEWs of the two pieces of equipment collided. The descriptions of *DEW* generation can be found in Chapter 6.

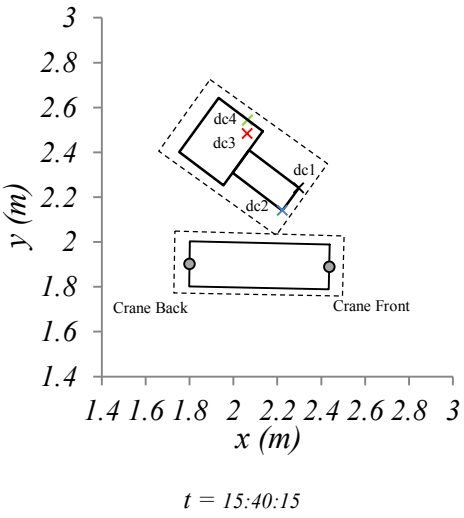


Figure 4-23: Warning Triggered by the Identification of Potential Collision

4.5 DISCUSSION

Given the difference in the level of orientation improvement offered by the proposed optimization-based method in the two case studies, two observations can be made. First, the accuracy of the orientation estimation before correction is significantly better in the second case study, where the average error was 23.40˚, as compared to the first case study, where the average error was 53.07˚. This can be attributed to the fact that in the second case study, as shown in Figure 4-18(a), the excavator was working near the center of the room, which gives a better visibility, while in the first case study the excavator was working near the edge of the room, as shown in Figure 4-11. To test this hypothesis, another static test was conducted by placing a tag at nearly same location as the excavator in the second case study and its results were compared to the results of the static test presented in the first case study, Section 4.4.1. It is observed that the location errors associated with 90% and 95% confidence level in the static test of the second case study decreased to 27 cm and 46 cm, respectively, compared to 45 cm and 48 cm calculated in the static test of the first case study. Similar observations are reported in the literature (Maalek and Sadeghpour 2013). Second, although the average error after the optimization-based correction for the first and second case studies are very close, i.e., 12.02˚ and 14.28˚ respectively, the proposed optimization-based method performed better in the first case study. The slight reduction in the performance of proposed optimization-based method can be attributed to the effect of the excavator's relocation and the reduction of the number of tags on the excavator. Based on these two observations, it can be concluded that the change in the performance of the *UWB RTLS* based on the location of the tags is the main cause of the difference in the level of improvement offered by the proposed optimization-based method.

It is worth noting that data redundancy is the underlying principle for the proposed optimization-based method, meaning that for the method to be applicable it is required that every rigid body part of the equipment is represented by one or more *DC*. This redundancy enables the identification of the *GC*s that are represented in terms of fixed distances on the rigid body and angular dependencies with regards to the relative orientation of various parts of a piece of equipment, as explained in the case studies. Therefore, as stated in Section 4.1, this method is particularly applicable to the less expensive *RTLS*s, e.g., *UWB*. Additionally, the proposed optimization-based method is not restricted to location data, and can as well work with other types of sensory data (e.g., digital gyroscope, Inertial Measurement Units (IMUs), sensors for measuring the level of extension of the hydraulic cylinders, etc.) provided the redundancy requirement is met.

Also, as mentioned in Section 4.1 it should be noted that although in this chapter the pose estimation method was presented in the context of an excavator, it is easily extensible to other types of equipment. Depending on the type of equipment and the number of *RTLS* tags used for data collection, specific *GC*s and OCs can be identified for that type. However, the first phase of the proposed optimization-based method, discussed in Section 4.2.3.1, is exclusive to the equipment with swinging movements, and if the method is to be applied to equipment without this type of movements (e.g. rollers), this phase is not necessary.

4.6SUMMARY AND CONCLUSIONS

Although advanced *AMC/G*-enabled equipment are capable of providing highly accurate pose, the high cost of the technology renders a large number of contractors unable to have access to accurate information about the pose of the equipment. This chapter presented a novel approach to improve the quality of data captured by less expensive *RTLS*s so that the pose of the

equipment can be estimated with an acceptable level of accuracy. A two-phase optimization-based method was proposed that uses a set of geometric and operational features of the equipment and the information about the position of *DC*s attached to different parts of the equipment to correct the location data captured by the *RTLS*. The method identifies the minimum amounts of corrections that need to be applied to the data captured by various *DC*s attached to an excavator to achieve a pose with minimum violations of the *GC*s and *OC*s. The first phase of the proposed optimization-based method applies a statistical approach to identify the *CR* of an excavator when it is not relocating. Two case studies were conducted to validate and verify the method and to demonstrate its applicability. In the light of the results of this research, it can be concluded that (1) the proposed optimization-based method improved the location accuracy for more than 60%; (2) the proposed optimization-based method improved the accuracy of orientation estimation by 77%, in the first case study, and 18% in the second case study; and (3) the estimated pose can be used to identify the potential safety hazards and also to determine different states of the excavator, which can be later used to calculate the cycle time as will be explained in Chapters 5 and 6.

Finally, although the proposed optimization-based method is very effective in improving the accuracy of the pose estimation, the computationally intensive GA detracts from the real-time-ness of the data. Given that the pose of the equipment is an integral part of the safety analysis of earthwork site, it is required to study other optimization algorithms that are capable of solving the presented objective functions faster.

CHAPTER 5: STATE-IDENTIFICATION AND NEAR REAL-TIME SIMULATION

5.1 INTRODUCTION

Although the planning and optimization of small-size and some medium-size projects can rely on the professional intuition and experiences of project managers, large scale earthmoving projects require a very meticulous and delicate planning, if productivity and efficiency are not to suffer. Therefore, in order to compare different scenarios for a given operation, and thus optimize the operation planning and resource allocation, it is indispensable to develop a digital simulation model of operations, in terms of time-sequenced activities and the flow of resources.

However, as mentioned in Section 2.2.3, in the traditional modeling approach, the modeler intuition and the historic data are the main input for the development of the simulation model. This approach is unable to capture the project-specific characteristics that tend to introduce uncertainties to the initially developed model parameters. Therefore, there is a need to develop a method that can leverage the accurate pose data, which were discussed in Chapter 4, to identify different state of equipment and then use the state information to continuously update the simulation model.

This chapter presents a comprehensive approach for *NRTS* using the pose information as the main input. The structure of the chapter is as follows: first the state-identification method is introduced in Section 5.2 and then the approach for *NRTS* is presented in Section 5.3. Finally, the proposed method is verified using two case studies in Section 5.5.

5.2 STATE-IDENTIFICATION

The data obtained from the pose estimation is not usable for state-identification unless put under another round of processing. The *DC*s provide information about the location or sensory data of particular objects and the time in which the readings were made. However, for these readings to be transformed into meaningful information usable for the simulation purpose, the data need to be converted to modeling parameters, e.g. durations and speeds. To materialize this, it is required to develop a knowledgebase which encapsulates expert rules that help interpret and transform the data. For instance, from an array of *GPS* readings coming from a receiver attached to a truck, we need to identify when the truck starts traveling to the dumping location or how long it has been in a queue before it is serviced by an excavator.

As shown in Figure 5-1, the proposed state identification method uses a rule-based system that applies a set of expert rules to the pose data to identify the states of different equipment and measure the cycle times.

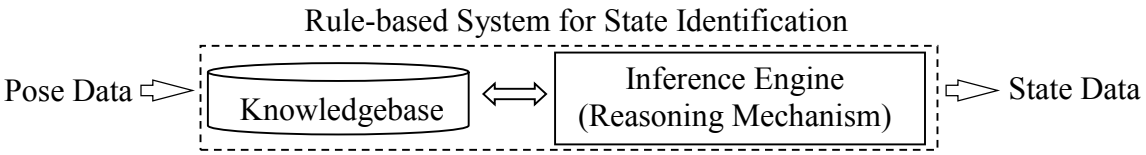


Figure 5-1: Structure of State Identification Method

The rule-based system transforms location/sensory/time data to states and simulation parameters and comprises a knowledgebase and a reasoning mechanism. The knowledgebase contains all the rules and heuristics which determine the states of machines and their current phase in the operation. These rules are experience-driven and case-dependent. On the other hand, using the knowledgebase, the inference engine helps determine the modeling parameters that are of interest for the simulation.

Expert rules need to be developed based on the exclusive characteristics of the project, the types of machines in the operation, the condition of the site, and the types of available data and the employed tracking technology(s). Of these parameters, the last two play an integral role in the formation of the expert rules. Different types of data that may be used for the expert rules include location data, sensory data, e.g. accelerometer and Inertial Measurement Unit (*IMU*), video data and any combination of these types. The accuracy of the expert rules varies according to the type and number of available data types and the rate of expected error associated with each type. In case *AMC/G* is used, a variety of sensory data is available along with the location data coming from *GPS*. These sensory data include measurements indicating the level of extension of the hydraulic cylinders and the trigger signs for different control units available on the cabin (Halder and Vitale 2010). It should be noted that with the increased number of available data types, although the accuracy increases, so does the complexity of the rule-based systems. However, the state-of-the-practice suggests that possible scenarios for the combination of data types, in order of the lowest to the highest accuracy, include: (1) only sensory data, (2) 2D location data, (3) 3D location data, and (4) Location data (2D or 3D) augmented with the sensory data (*AMC/G*). Consequently, it is of crucial importance to develop the rule-based system in view of the available data types. Akhavian and Behzadan (2012) showed that a minimal state identification can be performed using only sensory data. Moreover, it is possible to structure the rule-based solely based on the 2D location data, even in the absence of elevation data. This possibility offers the opportunity to take advantage of the proposed approach even when the sophisticated *AMC/G*-enabled fleet is not available, using alternative low-cost *UWB* technology or *GPS* receivers.

To provide a tangible example, a simple hauling-dumping project is used. In this example, a team of several trucks, an excavator and a conveyor belt, as the indicator of the dumping point,

are assigned to a hauling task. Trucks move to the loading area, get loaded, haul the load to the conveyer belt, dump and return. The operation for the excavator, on the other hand is to swing to the load, obtain a load, swing back to the truck and dump.

Depending on the type of tracking technologies used, many different rules can be applied. For instance, the state of dumping can be identified by the rise of the trucks' bed if 3D orientation tracker is used (Akhavian and Behzadan 2012). Alternatively, in the presence of the location data and with the assumption that the locations where different activities take place are fixed and unchanging, geo-fences can be used for the state-identification. Geo-fences are virtual contours drawn around an area that help detect if a unit has entered a known area (Reclus and Drouard 2009). In such cases, the dumping state of a truck, for instance, can be detected when the truck is in the dumping area, it is the closest unit to the conveyer belt and it is not moving. If the elevation data is present, the dumping is identified when the truck is in the dumping area and the elevation of the bed is increasing. Finally, if sensory data are also available, the dumping is identified when the truck is in the dumping area and the bed rise control is activated. However, the assumption of the fixed locations for activities can often be violated in practice. In such scenarios, the activities are taking place over dynamic zones (Pradhananga and Teizer 2013). One approach for the detection of dynamic zones is to apply pattern recognition methods, e.g. k-means clustering (Hartigan and Wong 1979). However this approach is more suitable for post processing, where a large pool of data is available and also a priori knowledge of the number of times the locations have been changed exists. Ostensibly, in near real-time applications, the rule-based approach is more practical. For instance, the dumping state can be identified when the truck is further than a threshold from the excavator and has a low speed. Table 5-1 summarizes

the states of a truck in the hauling operations and how they can be identified for the fixed and dynamic zones in four different scenarios where different types of data are available.

As for the excavators, given their higher *DOF*s and finer motions that segregate different states, the state-identification rules are more sophisticated. The states can be identified without zone detection on the account that the information regarding the velocity of the bucket and the direction of its move suffice for the state identification of the excavator. Like trucks, several states can be identified based on different types of data. For instance, a swing to the load is identifiable by sensory data through the detection of the rotational motion of the boom of the excavator. In case 2D location data are available, this state can be identified when the bucket is moving toward the truck with the swinging speed. If elevation data is available, the additional rule of predominant horizontal motion can help more accurately identify this state. Finally, in the presence of sensory data, the swing to the load state is identified when the excavator is in the swinging mode and it is approaching the truck. Six different states can be identified for the excavator in this example, as shown in Table 5-2.

Table 5-1: Example of Rule-Based System for Truck in a Hauling-Dumping Operation

Unit	State	Identification Rule			
		Sensory data	**2D Location Data**	**3D Location Data**	**Location (2D or 3D) and Sensory Data**
Truck (Fixed Zones)	In the Dumping Queue	Unidentifiable	The truck is in the dumping zone, it is not the closest unit to the conveyor belt and its velocity is zero	The truck is in the dumping zone, its velocity is zero and the elevation of the bed is not increasing	The truck is in the dumping zone, its velocity is zero and the bed is not raised
	Maneuvering for Dumping	Unidentifiable	The truck is in the dumping zone and it is moving		
	Dumping	The bed of the truck is raised	The truck is in the dumping zone., it is the closest to the conveyor belt, and its velocity is zero	The truck is in the dumping zone, and the elevation of the bed is increasing	The truck is in the dumping zone and the command for raising the bed is triggered
	Returning	Truck is moving	The truck is in return zone and its velocity is not zero		
	Idle	Unidentifiable	The truck is in the return zone and its velocity is zero		
	Maneuvering for Loading	Unidentifiable	The truck is in the loading zone and it is moving		
	In Loading Queue	Unidentifiable	The truck is in the loading zone but it is not the closest truck to the excavator and its velocity is zero		
	Waiting for Loading	Unidentifiable	The truck is in the loading zone, it is the closest to the excavator and its velocity is zero but the excavator is not in dumping		
	Under Loading	The excavator's bucket is opened	The truck is in the loading zone, it is the closest to the excavator, its velocity is zero and the excavator is dumping		
	Hauling	Truck is moving	The truck is in the hauling zone and its velocity is not zero		
Truck (Dynamic Zones)	In the Dumping Queue	Unidentifiable	The truck is in the vicinity of the conveyor belt, its velocity is zero but it is not the closets to the conveyor belt	The truck's velocity is zero, it is in the vicinity of the conveyor belt, and the elevation of the bed is not increasing	The truck's velocity is zero, it is in the vicinity of the conveyor belt and its bed is not raised
	Maneuvering for Dumping	Unidentifiable	The truck is in the vicinity of the conveyor belt and its velocity is not zero		
	Dumping	The bed of the truck is raised	The truck's velocity is zero, it is in the vicinity of the conveyor belt and it is the closest to the conveyor belt	The truck's velocity is zero, it is in the vicinity of the conveyor belt and the elevation of the bed is increasing	The truck's velocity is zero, it is in the vicinity of the conveyor belt and the command for raising the bed is triggered
	Returning	Truck is moving	The truck is moving towards the excavator		
	Idle	Unidentifiable	The truck is not in the vicinity of the excavator or conveyor belt and its velocity is zero		
	Maneuvering for Loading	Unidentifiable	The truck is in the vicinity of the excavator and its velocity is not zero		
	In Loading Queue	Unidentifiable	The truck is in the vicinity of the excavator, its velocity is zero and it is not the closest truck to the excavator		
	Waiting for Loading	Unidentifiable	The truck is in the vicinity of the excavator, its velocity is zero, it is the closest to the excavator but the excavator is not dumping		
	Under Loading	The excavator's bucket is opened	The truck is in the vicinity of the excavator, its velocity is zero, it is the closest to the excavator and the excavator is dumping		
	Hauling	Truck is moving	The truck is moving away from the excavator		

Table 5-2: Example of Rule-Based System for Excavator in a Hauling-Dumping Operation

Unit	State	Identification Rule			
		Sensory data	**2D Location Data**	**3D Location Data**	**Location (2D or 3D) and Sensory Data**
Excavator	Idle	Any state that does not meet the conditions of the other states			
	Relocation	Excavator is moving	The excavator track is moving (or alternatively there is a break in the swing pattern and a shift in the excavator's location is discerned)		The excavator is in the tram mode
	Swing to the Load	The body of the excavator is rotating	The bucket is moving toward the truck with the swing speed	The bucket is predominantly moving in a horizontal plane and it is moving toward the truck	The swinging is triggered and the bucket is approaching the truck
	Loading	The bucket is lowered or raised	The bucket's position is almost stationary and it is far from the truck	The bucket is predominantly moving in a vertical plane with a low velocity	The excavator is in bucket tilt mode and the bucket is being lowered to the ground
	Swing to the Truck	The body of the excavator is rotating	The bucket is moving away from the truck with the swing speed	The bucket is predominantly moving in a horizontal plane and the bucket is moving away from the truck	The swinging is triggered and the bucket is moving away from the truck
	Dumping	The bucket is lowered or raised	The bucket is almost stationary and its location intersects with the truck's bed	The bucket is relatively stationary vertically and horizontally and its location intersects with the truck's bed	Bucket open is triggered and the bucket intersect with the truck's bed

As can be seen in Table 5-1 and Table 5-2, the rules about the state of equipment can be reformulated in view of the types of available data. For instance, the rules are different when 3D location data are available as compared to when 2D location data are used. Also, the presence of additional sensory data helps enhance the accuracy of the rule-based system significantly. It is noteworthy that, as shown in the research of Akhavian and Behzadan (2012) and Table 5-1, although sensory data alone can be used to identify different states, in case of the truck, the level of detail that can be obtained is very low. Figure 5-2 and Figure 5-3 represents the state-identification flowcharts for trucks and excavators developed for 2D location data, respectively.

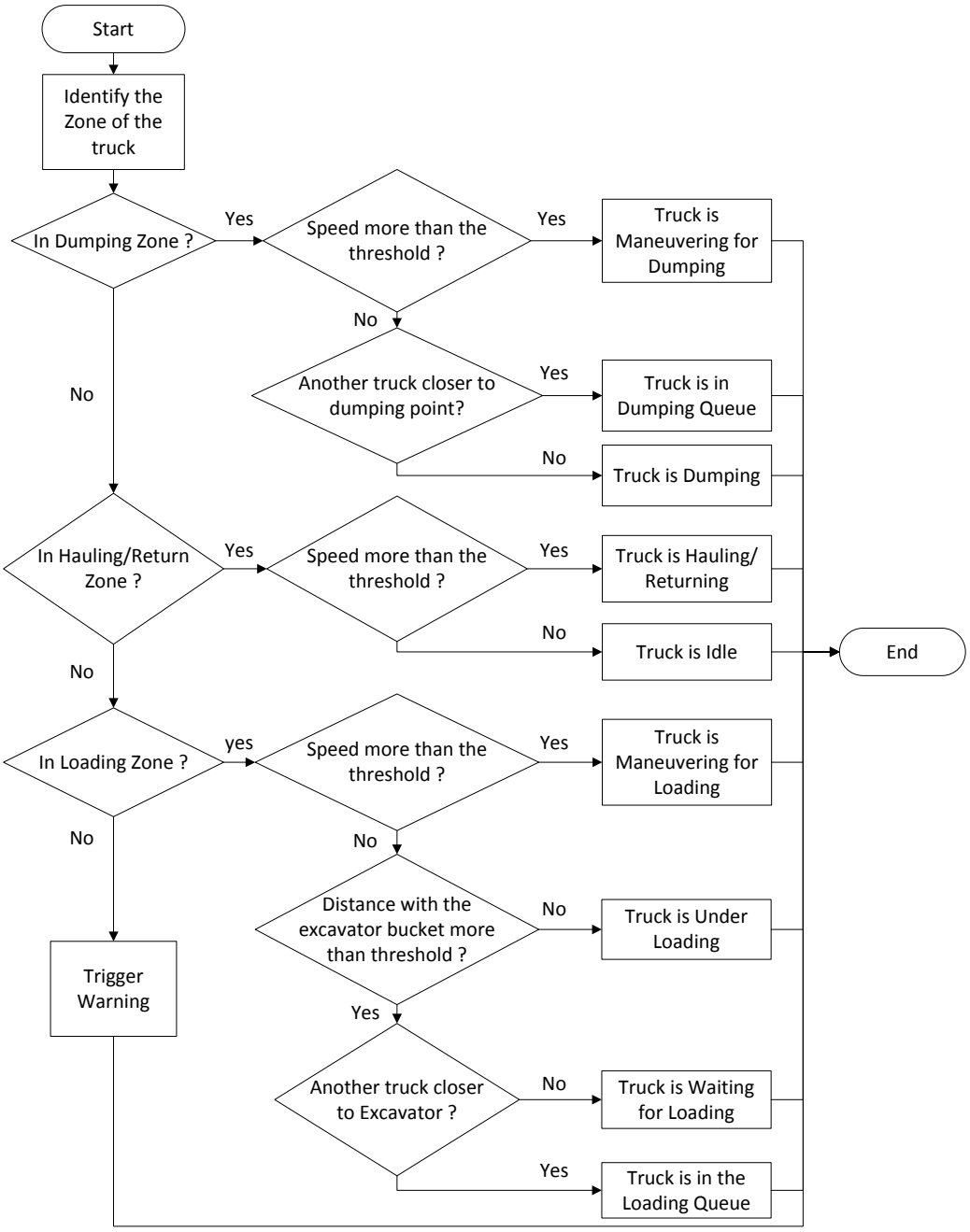


Figure 5-2: Flowchart of the State-Identification Based on Location Data for Trucks in the Rule-Based System

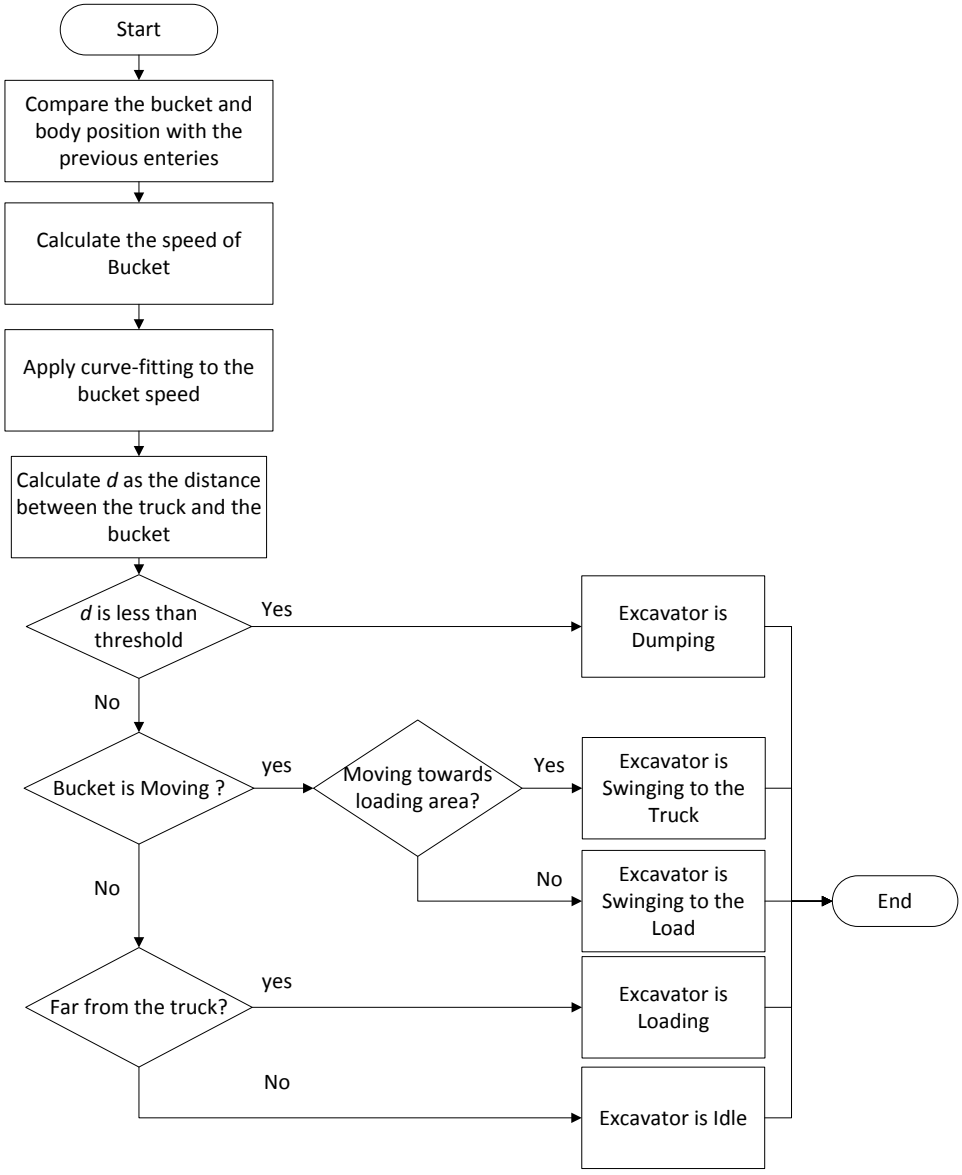


Figure 5-3: Flowchart of the State-Identification Based on Location Data for Excavators in the Rule-Based System

5.3 NEAR REAL-TIME SIMULATION

As stated earlier, *NRTS* is a concept based on the continuous update of a model initially developed for a particular process. The initial model is inherently established on a large amount of speculations and postulations, extracted mainly from similar previous projects. This model requires constant adjustment and tuning to remain accurate in the face of many anomalies and

discrepancies that are rampant in the course of the work execution. The model refinement can be performed using near real-time data from the *OA*s. With reference to the shortcomings of the previous *NRTS* methods reviewed in Section 2.2.3, a generic *NRTS* approach needs to: (1) accommodate different types of tracking technologies and function with different levels and types of information; (2) adopt a holistic view of the operation for the identification of the states of different pieces of equipment involved in an operation; (3) provide a detailed account of the equipment's states that can, in turn, furnish different levels of information to managers for a variety of purposes such as productivity measurement and safety improvements; (4) account for the environmental factors that determine the persistency of the captured discrepancies.

In the proposed *MAS* architecture, *OA*s are capable of performing detailed state-identification based on a variety of the sensory data at their disposal and with the consideration of the location and position of other *OA*s in the team. These capabilities already fulfill the first three of the above-mentioned requirements for a generic *NRTS* approach. The forth requirement will be discussed in this section.

In the proposed MAS, *NRTS* is performed by the *TCA*s based on the state-identification data provided to them by their subordinate *OA*s. Data about the current as-built status of the operation and site conditions are provided via the *DDA* and *SSA*, respectively. Integrating this information, the *NRTS* can more realistically perform operation forecasts at desired intervals or at critical times. It should be emphasized that the notion of real-time should not be construed as an uninterrupted chain of simulations as the operations proceed, but instead it should be understood as periodically performed simulation in intervals that are determined by the criticality of the operations or when a simulation is requested by the *GCA* in response to an urgent unforeseen occurrence. These intervals are referred to as "updating interval" hereafter.

In this sense three roles for *NRTS* are defined in the proposed MAS: (1) To periodically report to the *GCA*; (2) To provide updated forecast about the operations triggered in response to the design change update; (3) To provide predictions about the future states of other pieces of equipment to every *OA* for the purpose of *LAEW* generations, as will be explained in Chapter 7; and (4) To simulate the various scenarios requested by *GCA* when an unforeseen circumstance befalls. The *GCA* will generate a set of possible scenarios through varying the time and resource values associated with the elements of an operation and request from a *TCA* the feedback on the performance, e.g., productivity and cost, of different scenarios. Subsequently, the *TCA* will present the results to the *GCA* for selecting the optimum scenario, based on *NRTS*.

The structure of the proposed *NRTS* approach is presented in Figure 5-4. The proposed approach encompasses the following steps: (1) Once the data transfer from all the subordinate *OA*s is completed, it is required to direct the state data through another round of post-processing to ensure that individually captured *OA* states can be integrated to form meaningful and coherent simulation-usable operation cycle-time data; (2) Upon the completion of the data post-processing, if a mismatch is detected between the estimates, coming from a module called *Model Analyzer*, and the *OA*s' data, the sates are sent to another module called *Information Filter*, which will collect and analyze the external and environmental factors, from *SSA*, to determine whether or not the mismatch emanates from a temporary cause; (3) Depending on the decision made by the *Information Filter*, either or both of the following is performed: (*i*) the simulation model is updated by the *Model Refiner* based on the captured site data, and/or (*ii*) only the schedule is updated based on the amount of the measured delay; (4) The updated simulation model is then used by the Simulation Engine to update the schedule; (5) If the *GCA* decides that

an action is required, different scenarios can be consulted through the sensitivity analysis of the simulation model by a module called Scenario Analyzer.

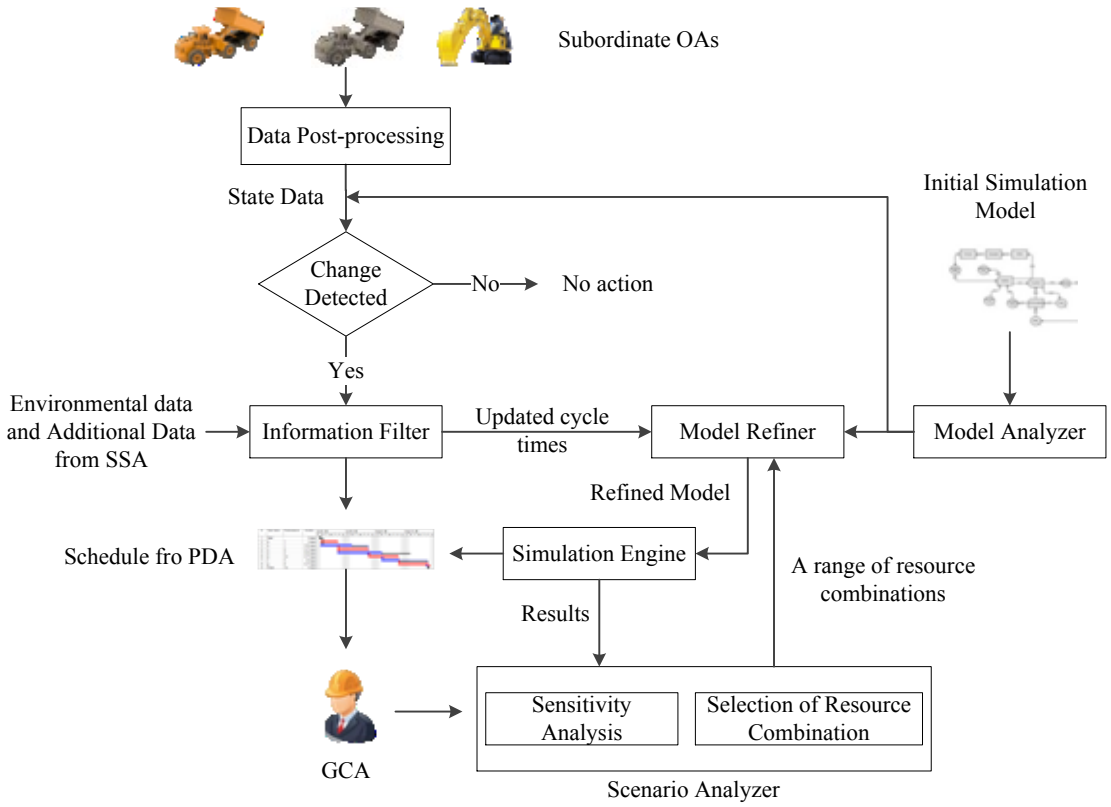


Figure 5-4: Architecture of the Proposed NRTS

The proposed refinement is based on the assumption that in every round of simulation update, if the simulation parameters are found to have been altered, the previous assumptions are invalidated and a new simulation model needs to be generated and executed. The backbone of this assumption is to notify the project manager about the expected finishing time of the project, should the operation follow the same regime as the last update interval. The merit of this assumption over simply accruing the sample data over time is that it represents the most recent working pattern based on the most recent changes in the site, with the condition that no major one-time or short-time occurrences have contributed to the mismatch between existing model

and the captured data. On the other hand, this assumption requires capturing enough number of cycle times before evaluating the need for simulation update.

5.3.1 DATA POST-PROCESSING

Data post-processing comprises two layers of analysis, namely, *Pattern Analysis* and *Cycle Logic Check*.

5.3.1.1 PATTERN ANALYSIS

State-identification based on the strict rules explained in Section 5.2 is susceptible to errors, even after the error correction of the raw data. The cause of these errors is the fact that the motion of equipment in practice is not uninterrupted and has a degree of intermittence. As a result, the analysis of equipment state at one instance of time without the consideration of the preceding and succeeding states is prone to inaccuracy. In order to offset these errors, a pattern analysis needs to be performed, as shown in Figure 5-5. The pattern analysis runs through the results of state-identification and pinpoints states that do not constitute a pattern.

This threshold can be defined in view of the averaging timespan used in the correction phase (dt), the typical duration of the shortest state in operation (t_s), and the logical shortest duration for an activity to be legitimized as a valid activity and not a fluctuation (t_{min}). Based on these values, for every entry in the data, n entries on each side of the current entry, i.e. preceding and succeeding, need to be included in the pattern analysis. Nevertheless, at least m entries should have the same state for the current entry to be recognized as a part of a pattern. In other words, the threshold states that p percent of the total range chosen for the analysis must have an identical state. The n, m and p can be obtained using the Equations (1~3).

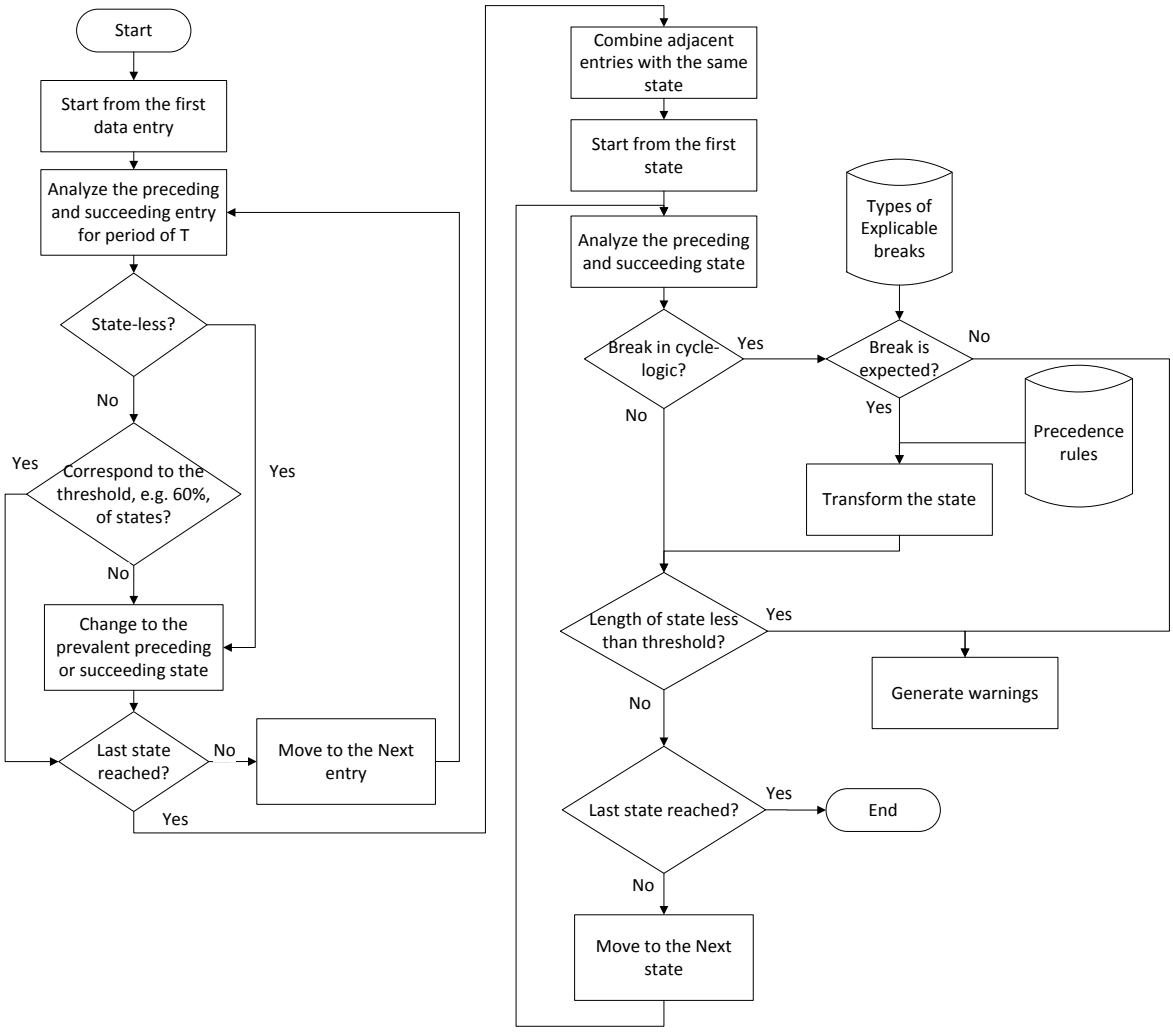


Figure 5-5: Flowchart of the Pattern Analysis and Cycle Logic Check in the Rule-Based System

$$n = \frac{t_s - dt}{2 \times dt} \qquad \textbf{Equation 5-1}$$

$$m = \frac{t_{min}}{dt} - 1 \qquad \textbf{Equation 5-2}$$

$$p = \frac{m+1}{2 \times n + 1} \times 100 = \frac{t_{min}}{t_s} \times 100 \qquad \textbf{Equation 5-3}$$

For example, if the duration of the typical shortest state in the operation (t_s) is 10 sec., the logical shortest duration of that state (t_{min}) is 6 sec., and the data is averaged over 2 sec. (dt), then the analysis should be performed over a range of 2 entries (n) before and after every entry, altogether encompassing 5 entries. Of these range, at least 2 entries (m) must have the same state as the

current entry, corresponding to a threshold (p) of 60%. Nevertheless, very high p will result in the elimination of a large quantity of the data and setting it too low will cause no change in the state-identification. The pattern analysis eliminates the states that do not satisfy the threshold. Then, it assigns a state to the state-less entries based on the prevalent state preceding or succeeding the state-less entry, using the same threshold. The impact of pattern analysis on the smooth data capturing correlates directly with the timespan used in the pose estimation. The greater the timespan, i.e. the more data are bundled together, the less the need for the pattern analysis.

5.3.1.2 CYCLE LOGIC CHECK

Although the pattern analysis suffices to capture a good account of the equipment states, an additional cycle logic check is required. Given that the result of state-identification needs to be used in the simulation, and that the simulation model has a predefined logical sequence of states, the captured data need to be mapped to this logical sequence, as shown in Figure 5-5. The need for this operation stems from the fact that in practice there can be minor breaks in the logical sequence. For instance, take a truck that enters the dumping state, maneuvers for dumping and stops for dumping but, due to some reasons, readjusts its position before the actual dumping. In this case, the rules given in Table 5-1 identify the following states: (1) maneuvering for dumping, (2) dumping, (3) maneuvering for dumping, and (4) dumping. Thus, the main functionality of cycle logic check is to reinterpret the states in the context of the logical sequence of states. The incongruity with the simulation logical sequence is induced by two possible types of breaks. First, the captured sequence of states represents an expected break, i.e. irregularity caused by the repetition of the same states in a close proximity or interchanged neighboring states, e.g. the example given above where an extra maneuvering for dumping is identified

between two dumping states. Second, the captured sequence of states contains an inexplicable break, caused by missing states or non-neighboring states, e.g. the state of loading is not identified for the truck and the truck is identified to proceed to the hauling after maneuvering for the load. Although cycle logic checks can identify both types of breaks, only the first type can be subject to modifications. In the case of inexplicable breaks, the cycle logic can only generate a warning to inform the modeler or manager of the anomaly. For the modification of the first type of breaks, a predefined knowledge of the ranges of possible breaks needs to be used to transform the captured states. The decision about what state should be transformed to what other state depends on the hierarchical precedence rules, and can be determined by the simulation modeler. For instance, in the above example, two options exist. Either the middle maneuvering state is transformed to dumping state or the middle dumping state is transformed to the maneuvering state. The selection is made based on the preference of the manager or simulation modeler to truncate maneuvering time towards dumping time or vice versa. Also, cycle logic can help identify the states that have been unreasonably long, using user-defined thresholds. It is worth mentioning that unlike pattern analysis, which is performed for every entry, cycle logic check combines all the adjacent entries with the same state and operates at the state level, i.e. it only compares the states.

Upon the completion of the cycle logic check, the rule-based system calculates the durations of every identified state and transfers the data to the *Information filter*.

5.3.2 INFORMATION FILTER

The *Information Filter* monitors various sources, and should it detect discrepancies that demand a change of the model, it will send a request to the *Model Refiner* for further actions. The flowchart of the *Information Filter* is shown in Figure 5-6.

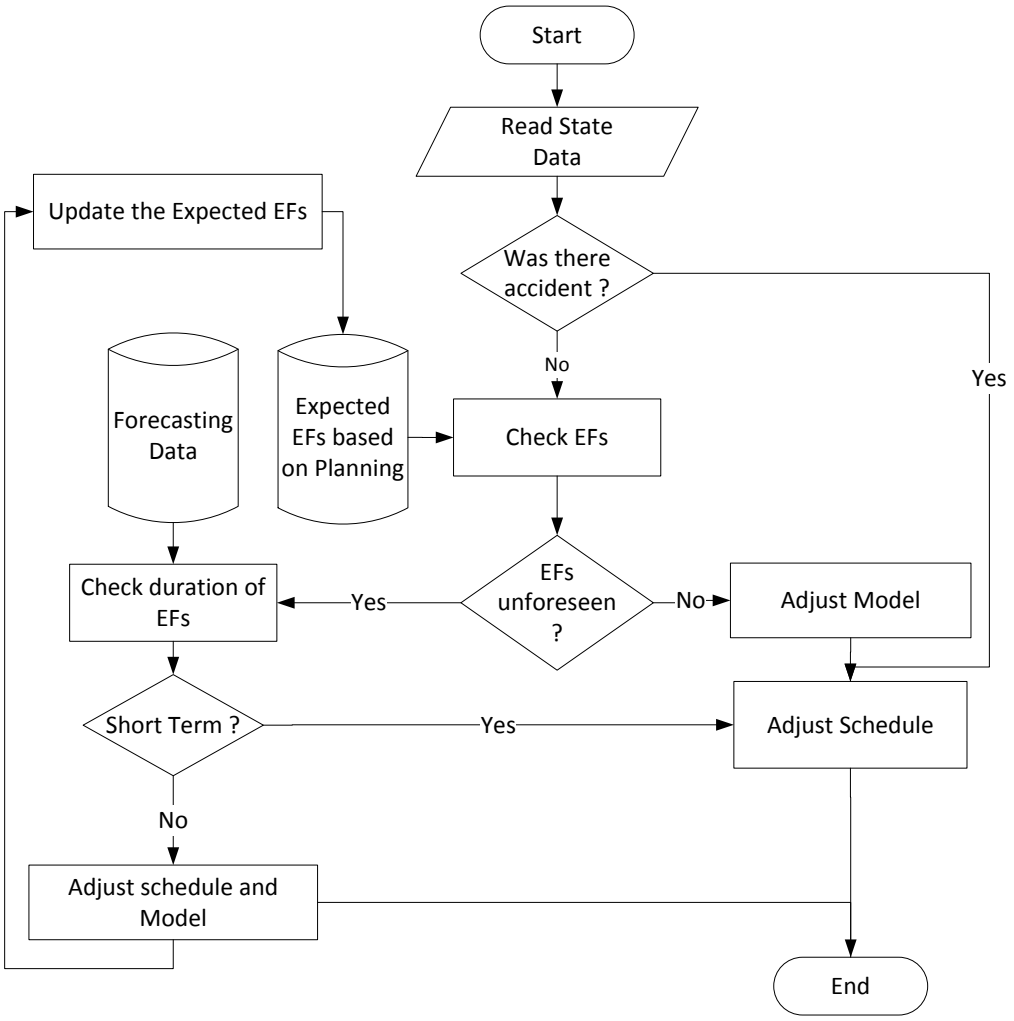


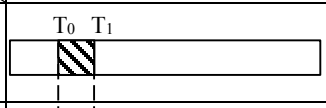
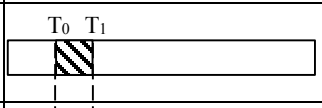
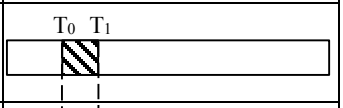
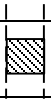
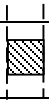
Figure 5-6: Flowchart of Information Filter

For this component to be able to distinguish refinement-requiring discrepancies from negligible data fluctuation, three types of data are required, namely: (1) expected conditions at the planning phase, (2) data capture from the site in near real-time, and (3) the most recent weather forecasting data. These data can be captured from on-site data capturing tools, e.g. thermometers and hygrometer, and a wide range of public weather stations. The parameters that may carry transient implications on the process include, but are not limited to, temperature, humidity, precipitations, operators' skill, equipment condition, working hours, time of the day, dynamic changes to the site layout, accidents, etc. For the brevity, these parameters are termed

Environmental Factors (*EF*s) in this research. While any of the *EF*s could affect the progress of the project, it is conceivable that their impacts are temporary. Accordingly, every machine state and the pertinent simulation parameters need to be processed and analyzed in line with a set of corresponding *EF*s.

The *Information Filter* is triggered when state data are generated by the rule-based system and a mismatch is identified between the current assumptions in the model and the most recently captured data from the site. Needless to say, if no discrepancies are identified, no action is taken, otherwise the *Information Filter* checks if there have been any accidents. The *Information Filter* adjusts the schedule based on the amount of delay if an accident had happened. It is worth mentioning that the case of broken equipment is also considered as a type of accident. If a piece of equipment has been broken only during the last updating interval (T_0~T_1), only the impact is reflected in updating the schedule, as shown in Figure 5-6. However, if the equipment is known to be unavailable in future, this change has to be performed manually by the manager via adjusting the simulation model. Table 5-3 shows three scenarios that may follow when discrepancies are identified and no accidents were detected. In this figure, T_0 and T_1 represent the start and end of the update interval, respectively. First the current EFs are checked against the expected values at the planning time to identify any unforeseen environmental conditions. If all the EFs are according to the expectations, the model estimates need to be updated (Scenario 1 in Table 5-3); else the duration of the unexpected EFs should be checked based on the forecasting data. In case they are short-term changes, only the schedule is updated based on the occurred delay but the model is left unmodified, (Scenario 2 in Table 3). If the unexpected EFs are found to be long-term, both the model and schedule need to be adjusted, (Scenario 3 in Table 5-3), and also the expected EFs at the planning phase need to be updated based on the most recent data.

Table 5-3: Different Scenarios Resulting from Information Filter based on EFs When No Accidents Are Detected

Scenario / Source of EFs	Scenario 1: Matched site and planning EFs	Scenario 2: Short-term unmatched site and planning EFs	Scenario 3: Long-term unmatched site and planning EFs
Planning Phase			
Site between T_0 and T_1			
Forecasting at T_1			
Decision	Model Update	Schedule Update	Model & Schedule Update

5.3.3 MODEL ANALYZER

The *Model Analyzer* interprets the initial model and translates the simulation's internal logic to a formalized and modifiable format. Then, it disintegrates the model into the constituent components and formalizes the causal relationships, activities and resources so that they can be individually modified and re-assembled by the *Model Refiner* for further execution.

5.3.4 MODEL REFINER

The inputs for this component are (1) the change request coming from the *Information filter*, which represents a major deviation from the existing values in the model, (2) the state of the equipment resulting from the rule-based system, and (3) the parsed model coming from the *Model Analyzer*. The *Model Refiner* identifies the parameters that need to be adjusted based on the change request placed by the *Information Filter*, and updates the model accordingly. The Model refiner also performs the statistical distribution fitting to identify the best statistical distribution that represents the captured data.

5.3.5 SCENARIO ANALYZER

The scenario analyzer mainly performs sensitivity analysis over a range of resource quantities to help optimize the configuration of the fleet. For example, if a unit runs out of service, this component would allow the management to explore the possible options they can choose from. This component includes an engine for the sensitivity analysis, which generates a range of resource combinations, and a module to opt out the best solution. Sensitivity analysis refers to the process of running the simulation model over the range of options that can be considered through varying the values of resources. It allows performing a comparative study of various scenarios in terms of productivity and costs. The selection of the best scenario relies on the subjective opinion of the manager or can be performed using a range of methods, from complex optimization method, e.g. generic algorithm, to simple ranking methods.

5.3.6 SIMULATION ENGINE

The simulation engine is a platform on which the process model is run and the subsequent report is generated. This component goes through the model simulation and flows the resources amid the network of activities to perform the time and productivity evaluation of the model.

5.4 IMPLEMENTATION

To demonstrate the applicability of the proposed approach, a prototype, including two loosely-coupled applications was developed and tested in this research. Both applications are developed using *VBA* in Microsoft Excel. The first application encompasses the simplified correction method and the Rule-based System and is designed to receive the raw location data as input and process them through the steps elaborated in Section 4.2.2. Once the location data are corrected,

the rule-based system considers many different location parameters to infer the state of different pieces of equipment, as explained in Section 5.2, and presents them in a different sheet.

The second application comprises the *Model Analyzer* and *Model Refiner*. Stroboscope (Martinez 1996) is chosen as the simulation engine due to its palpable strengths in extensibility, robustness, availability, compatibility with simulation of construction operations and ease of use.

This implementation is generic in the sense that it can accommodate any type of operations that can be represented by a simulation model.

5.4.1 DATA CORRECTION AND RULE-BASED SYSTEM APPLICATION

This application requires the user to input the file containing the location data generated by *RTLS*. At the present state of the prototype, it can be applied to a *UWB* system (Ubisense 2015), which is based on local Cartesian location data, and the *GPS* system, which is based on latitude, longitude and altitude.

In an *UWB RTLS* system, a sensor cell is composed of several connected sensors which capture the location of tracked objects. The synchronization of sensors by the master sensor is done using timing signals from each sensor. Location data are registered from *UWB* tags that are attached to objects of interest. The tag data registry allocated by the sensors is done using *TDMA* method that splits the signal into several time slots and allocates each slot to one tag. Tags data are received in quick succession at their own time slots. The time slot allocation is optimized to ensure that while enough space is maintained for the registration of new tags, each tag is receiving the attention in line with the expected service quality (Zhang 2010; Zhang et al. 2012).

Once the data is imported to the application, the data are categorized according to the tags' names, sorted from the earliest to the latest reading and stored separately in different sheets. At

the initial run of a new set of data, since the synchronization is required for the pairwise comparison of tags, correction and averaging over the tags, the user is restricted to perform averaging over time before any further processing. The data can be averaged over time at the time span desired by the user. The correction module executes the simplified correction method explained in Section 4.2.2. The averaging over tags module is intended to provide the possibility for averaging several tags into one representative point using the averaging of the original tags' coordinates.

The state-identification module subjects the data to the rules explained in Section 5.2 and infers the state of the equipment associated with every entry and further process them using pattern analysis and cycle logic check. At the present version of the application, only rules for a simple hauling-dumping operation that considers the states of trucks and excavators are embedded. For this purpose, the user is asked to choose the tags that represent the truck, the excavator body and the excavator bucket. Also, the user can decide to use fixed zones, for which the coordinates of the geo-fences are required, or dynamic zones (Pradhananga and Teizer 2013).

With regard to the excavator's state identification, given that the combination of the low speed of the bucket motion and its associated errors may hinder the precise identification of the beginning and end of swing action, a data smoothening technique needs to be used. In this application, Matlab curve-fitting tool is integrated with the *VBA* code to perform Sum of Sine method for curve smoothening (Mathworks 2015). However, as part of the future work, the implementation of more robust techniques such as Kalman filtering is being considered.

A module is designed to provide several types of visualization. The first type of visualization is the chart of the movement tracks of the tags, which represents how a particular tag moved on the site, as shown in Figure 5-7(a). These charts are animation-enabled, meaning that the user can

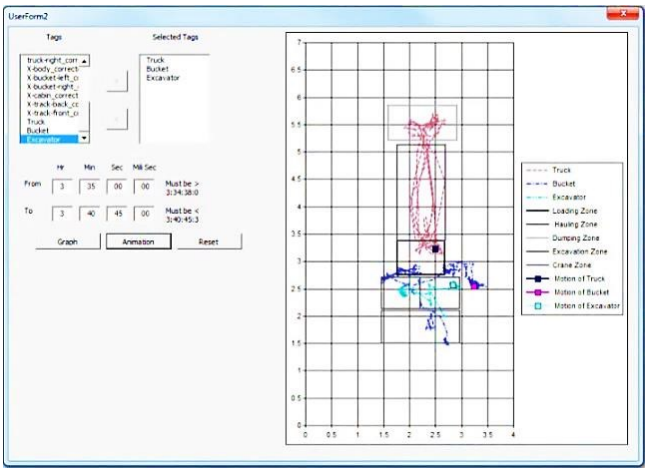
chose to see the movement of tags over time. The second type of visualization pertains to the state-identification, as shown in Figure 5-7(b). This type, which is also animation-enabled, helps the user pinpoint the state of different pieces of equipment in comparison to the location of different tags. The third type of visualization allows the user to load the video of the operation through the panel shown in Figure 5-7(c) and compare the results of state-identification and movement tracks with the video for the purpose of validation.

Furthermore, the user is presented with a table of the averages and standard deviations of the durations each piece of equipment spent in every state, as shown in Figure 5-7(d). This information is obtained through averaging the duration of all the similar states over the period of analysis, excluding the states where violation of threshold is identified. Figure 5-7(e) shows the loaded video at exact same time as the location of tags shown with small squares in Figure 5-7(a) and (b).

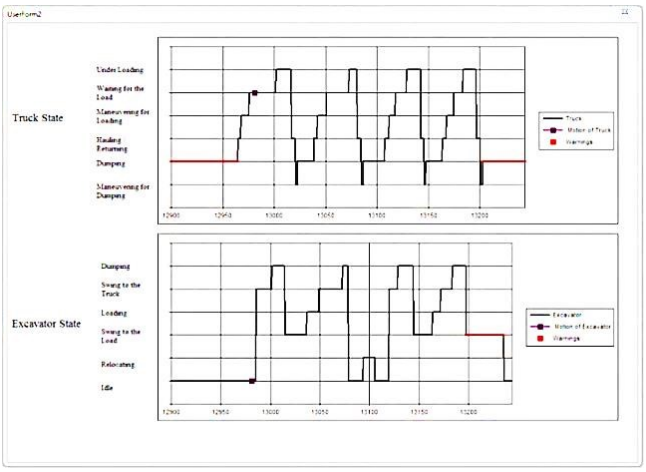
5.4.2 NEAR REAL-TIME SIMULATION APPLICATION

In the proposed approach, once the raw location data are processed in the rule-based system by *OA*s, and thus transformed to information about the machine states, and then filtered by the *Information Filter*, they are sent to the *Model Refiner*. However, at the current stage of the application, the assumption is that the input state data coming from the rules are already classified by the *Information Filter* as changes requiring the refinement of the simulation model.

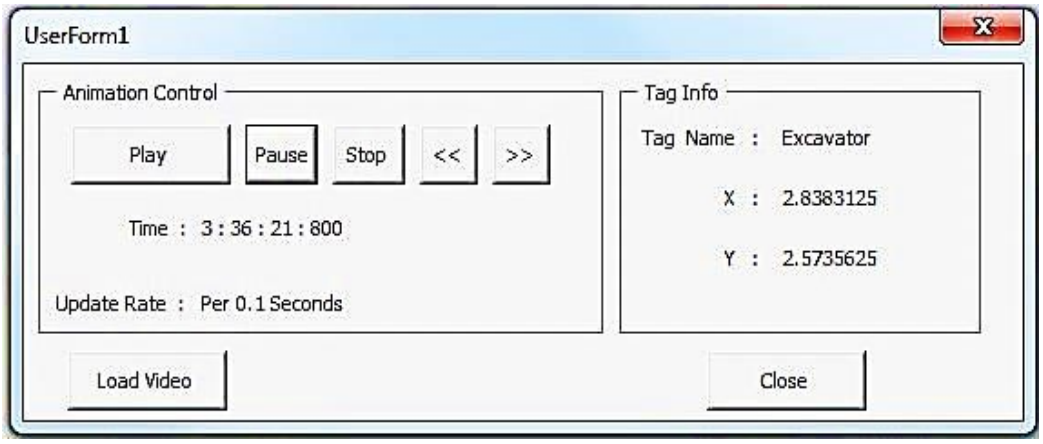
The *NRTS* module, as shown in Figure 5-8, begins with receiving the initial simulation model as an input to the application. The *Model Analyzer* then parses the model and creates the list of durations and features that are used in the model. At this point, the user is asked to choose the variables and parameters that are used in the simulation model, e.g. hauling speed, and correlate the chosen variables/parameters with the relevant state data coming from the *Information Filter*.



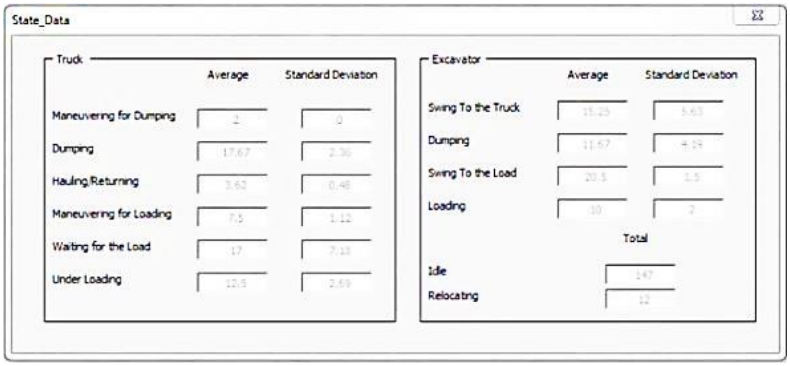
(a) Tags' Location



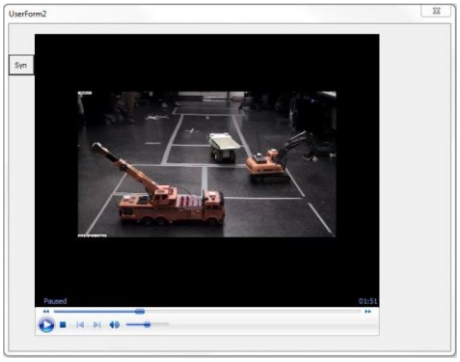
(b) Equipment states



(c) Video control panel



(d) Summary of cycle information



(e) Synchronized video

Figure 5-7: Interfaces for the Presentation of the Results of Simplified Correction Method and Rule-Based System

For instance, the user determines that the actual speed of truck 1 in the hauling state is the state data appearing in the column *A* in an Excel sheet. To establish this correlation, the user chooses the hauling speed in the interface of the *NRTS* module and defines the new value as *A*. Once the

correlations are made, the application monitors the flow of state data and updates the model according to the incoming state information. Every time an update is made, the model is re-run in the simulation engine and results are published in the Excel sheet. The interfaces of this application are shown in Figure 5-9.

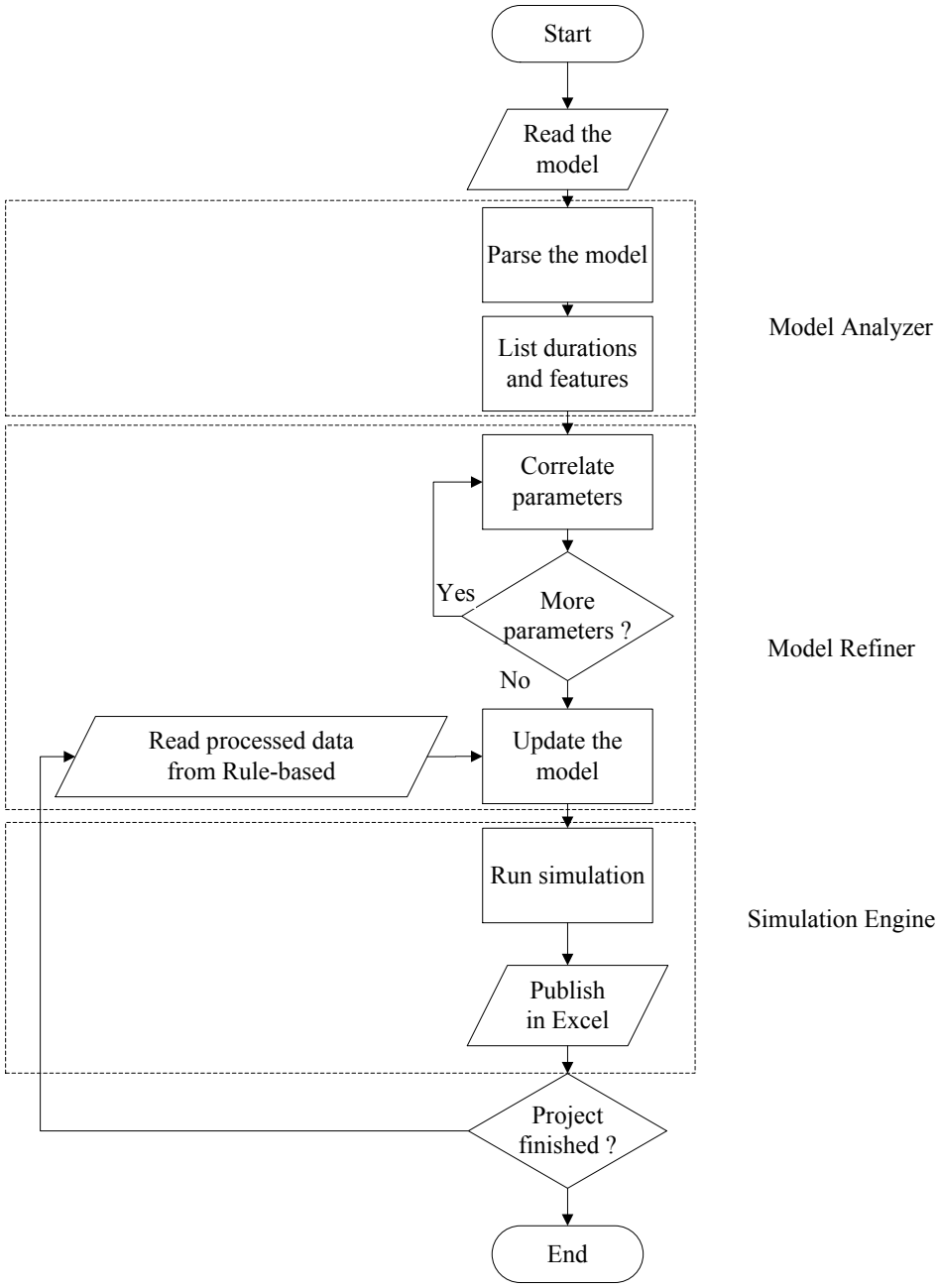
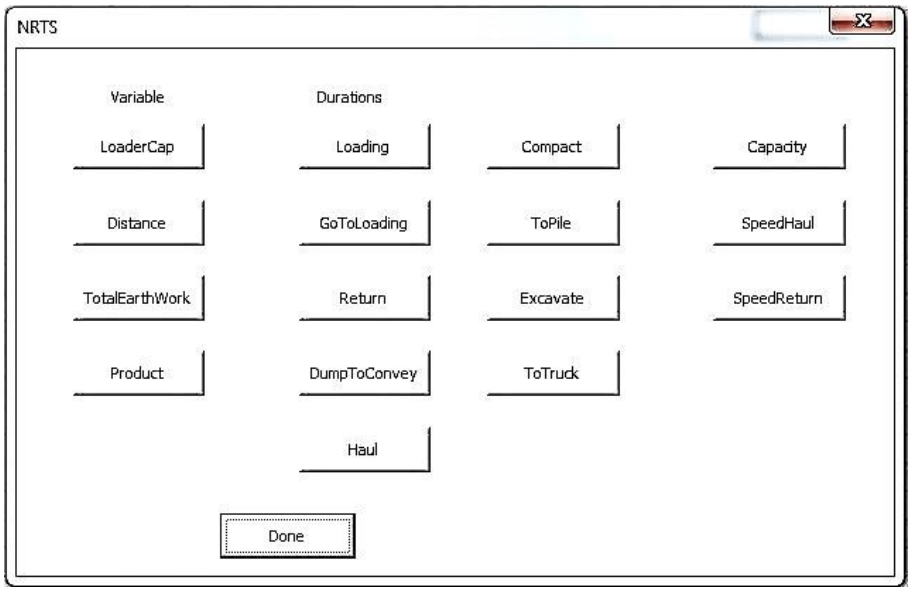
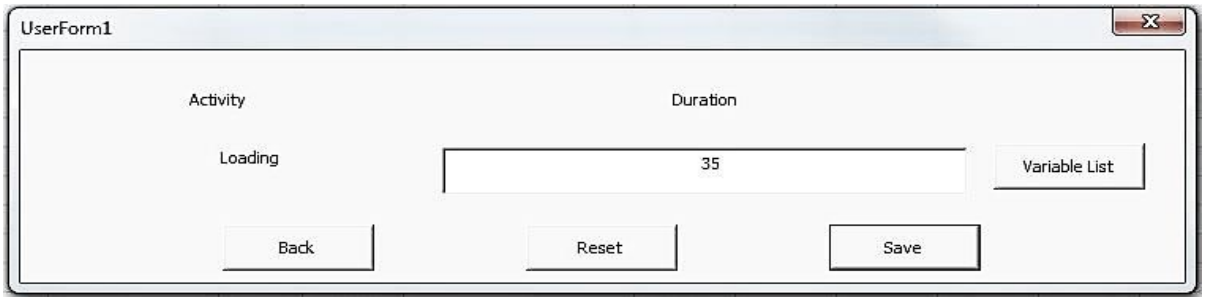


Figure 5-8: Flowchart of the NRTS Module

(a) Main menu



(b) Selection of Parameters to be Correlated to Captured Data



(c) Correlating the Captured data with the Parameters

Figure 5-9: Interfaces of NRTS Application

5.5 CASE STUDIES

The following case studies aim to validate how the state of the equipment can be derived from the location and sensory data and further used for the decision making and NRTS. Although the implementation of a fully integrated system is beyond the scope of the present research, the developed prototype is used to test the proposed approach.

The case studies demonstrate the tasks related to the following components of the approach: (1) Simplified Correction Method: correcting the location data so that errors are minimized; (2) Rule-based system: inferring the state information from the location data using the rules from Table 5-1 and Table 5-2; and (3) Model Refiner: updating the simulation model using the developed application. In the first case study, the proposed method was applied on the lab test presented in Section 4.4.2. The second case study, on the other hand, is designed to demonstrate the applicability of rules in Table 5-2 for an *AMG* excavator in a mining site.

5.5.1 RESULTS OF THE FIRST CASE STUDY

The module for simplified correction method, explained in Section 5.4.1, is used to process the location data. In order, averaging over time, correction and averaging over tags are applied to the data. Figure 5-10 provides an example of how the applied correction helps improve the quality of the data. Also, the improvements of state-identification by the pattern analysis and cycle logic check are shown in Figure 5-11.

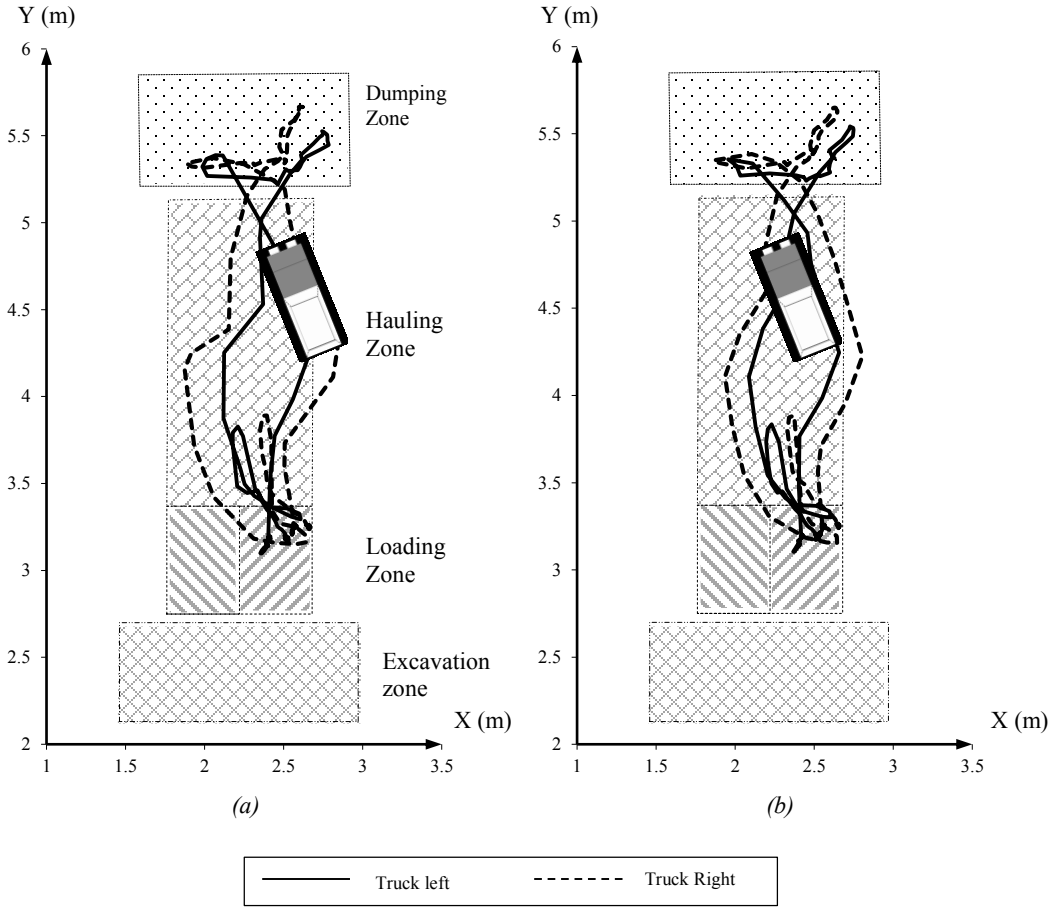


Figure 5-10: Movement of Track (a) Before and (b) After Correction

In these charts, the horizontal axis represents time and the vertical axis indicates the states. The solid line implies the presence in a particular state. It is evident that these steps helped in eradicating the minor errors that suggest break in the logic of the cycle.

Comparison of the results suggests that, in conformity with the expectations, if the cycle logic test is not performed, the result of state-identification is more ruffled for the sub-second timespans, where a less number of entries are averaged, and thus more readings need to be modified by the pattern analysis and cycle logic check. This is demonstrated in Figure 5-12, where the results of state identification of truck before cycle logic check is presented for 0.5, 1 and 2 seconds.

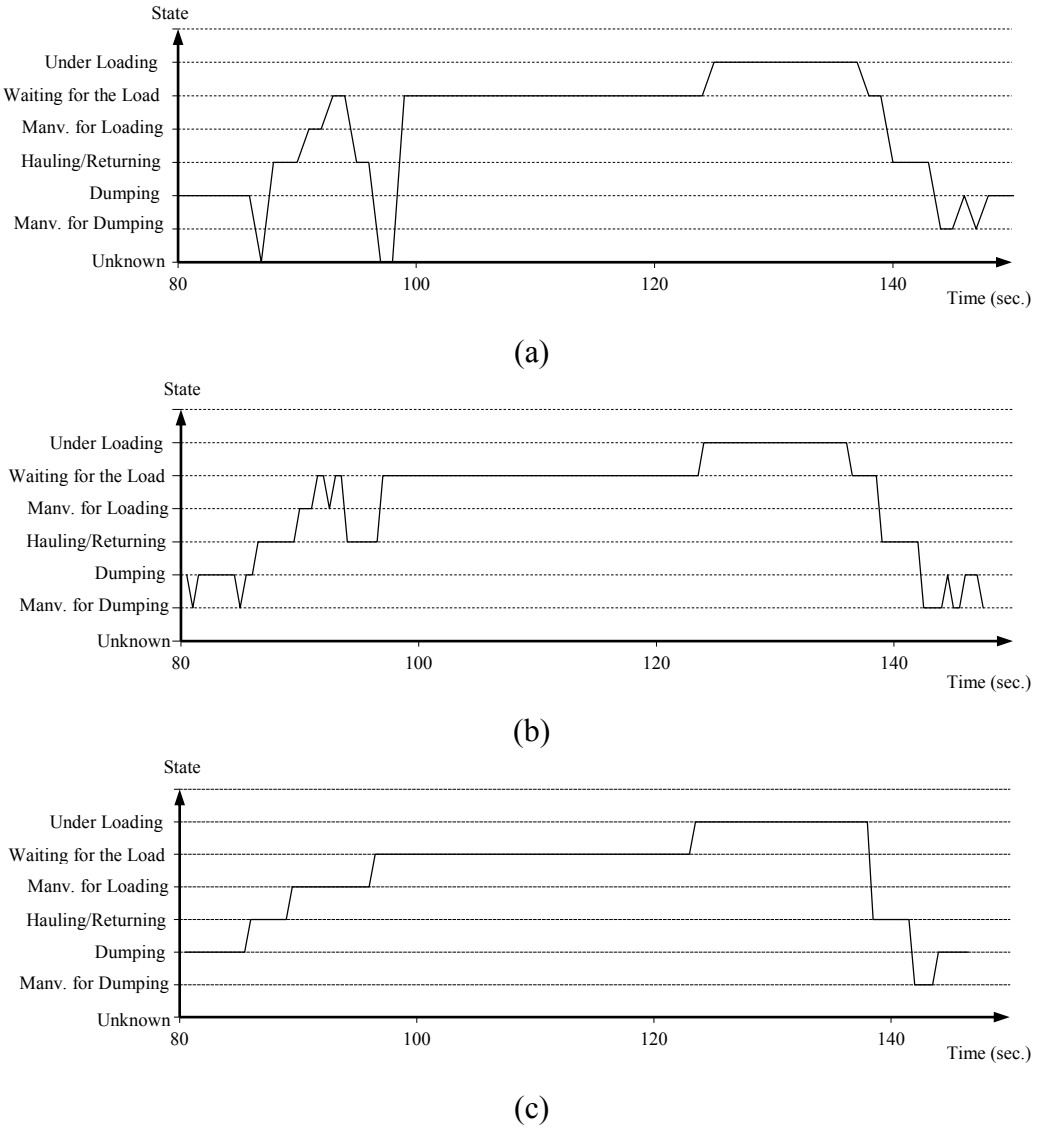


Figure 5-11: State Identification for the Truck with: (a) No Processing, (b) Pattern Analysis and (c) Pattern Analysis and Cycle Logic Check (0.5 sec. averaging timespan)

In order to identify the optimum averaging timespan, the data preparation was performed for three different timespans, namely, 0.5, 1 and 2 seconds, with the assumption that zones are fixed. The state-identification outcomes are shown in Figure 5-13. The upper and lower charts show the transition and duration of states for the truck and excavator, respectively. As can be seen in Figure 5-13, the cycle logic check uses the user-defined threshold to mark the initial dumping as too long. Also, it can be inferred that with the increased timespan, along with the increased

smoothness, there exists the loss of detail, which in the case of 2-second averaging timespan results in the total omission of the excavator's relocation. The juxtaposition of state-identification with the video, suggests that 1-second averaging timespan yields the best result, in terms of smoothness and accuracy, as it is the point after which the increased smoothness causes the omission of some states, e.g. maneuvering for loading in 2-second timespan.

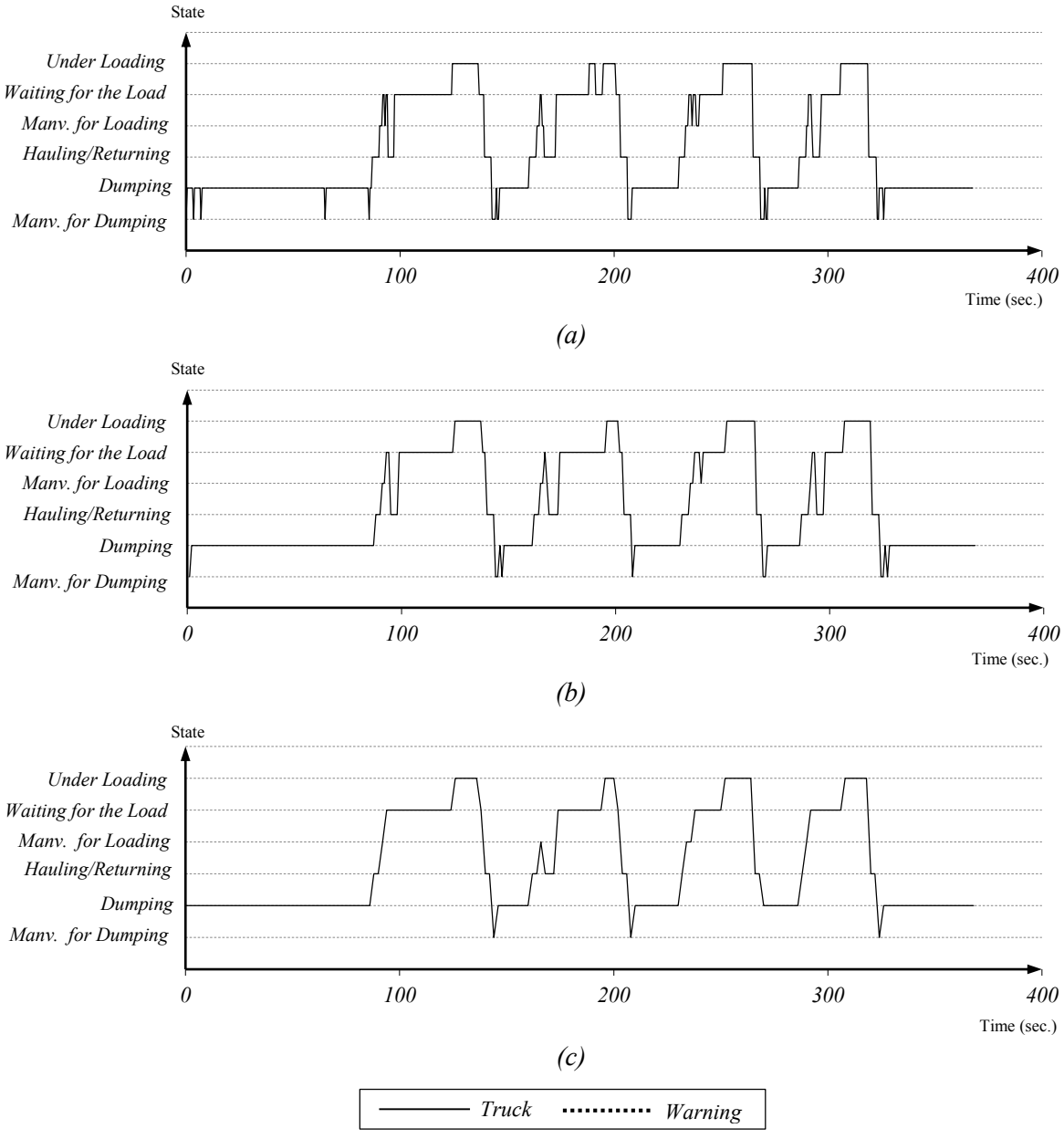


Figure 5-12:Results of State-Identification for Truck for: (a) 0.5 Sec., (b) 1 sec., and (c) 2 sec. Averaging Time Span before Cycle Logic Check

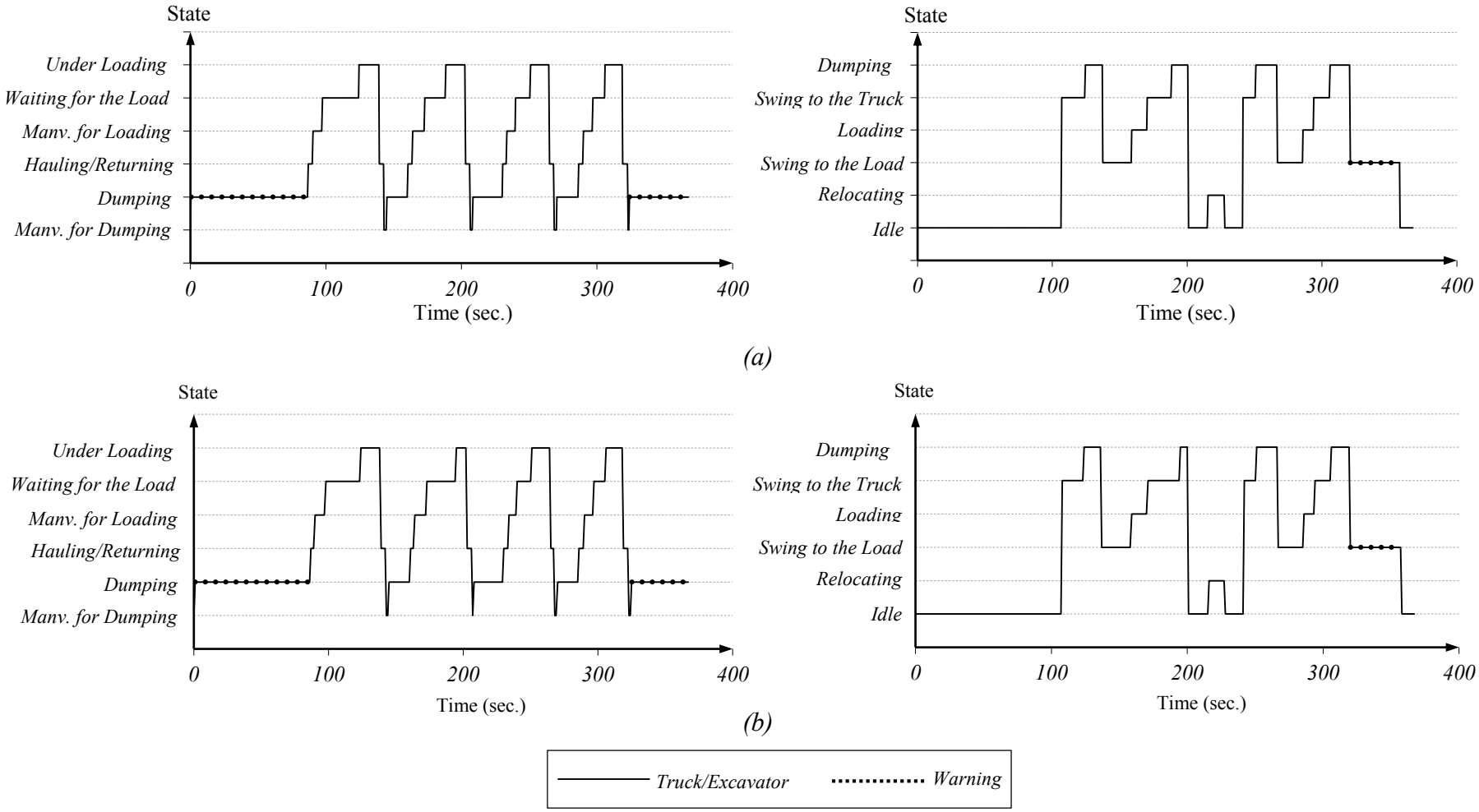


Figure 5-13: Results of State-Identification for: (a) 0.5 Sec., (b) 1 sec., and (c) 2 sec. Averaging Time Span

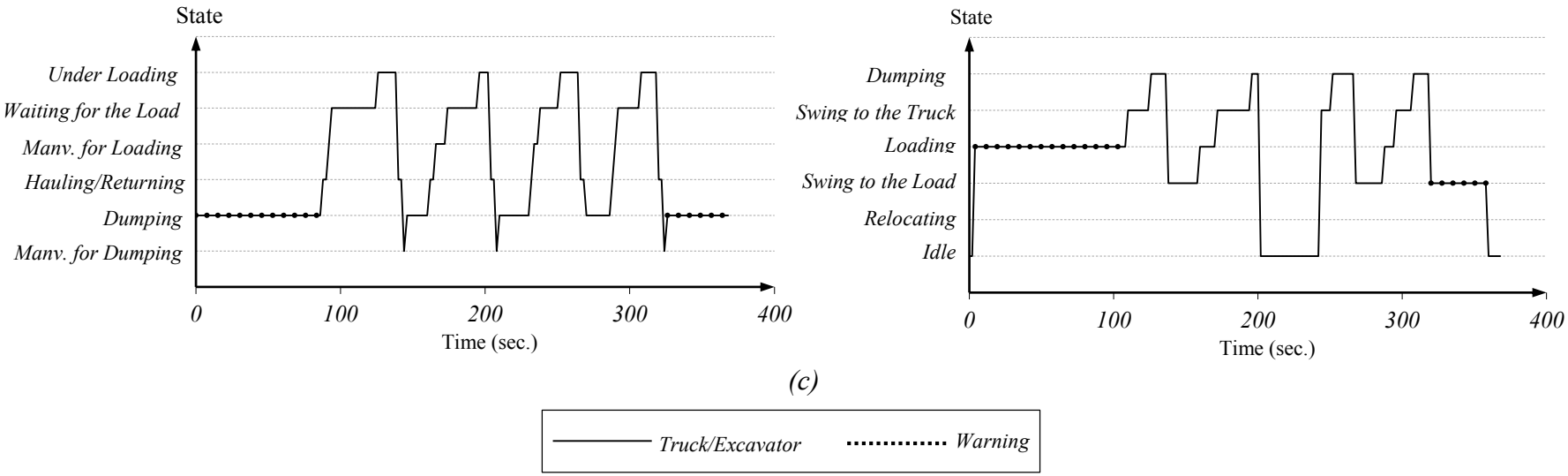


Figure 5-13: Results of State-Identification for: (a) 0.5 Sec., (b) 1 sec., and (c) 2 sec. Averaging Time Span (cont.)

The unrealistic ratio of different states in relation to one another, e.g. the hauling time to waiting time for the loading, can be ascribed to the combined effects of the scale of the equipment and site together with the equipment's operational speed. Also, given that only one truck was utilized in this test, states that involve waiting in queues are not pertinent to this case study. Additionally, with the hauling and return path being identical in this test, no discrimination is made between the two states.

Another observation is that when 2-second timespan is used, the initial state of the excavator is identified as loading. This inaccuracy is caused by the loss of detail. Dotted lines depict the states where the equipment exceeded the user-defined threshold. These warnings can help the manager further investigate the cause of the prolonged presence in a state. However, in the present case study, all the warnings happened at either the inception or the conclusion of the test, implying that the cause of the prolonged presence is the mismatch between the starting/ending point of the test and data registry.

The data was also tested for the accuracy of the rules without the assumption of the fixed zones for the averaging timespan of 1 second. Figure 5-14 shows the result of the state-identification for this scenario. Since the state-identification of the excavator does not require zone-detection, the excavator's states are identical to those shown in Figure 12, and thus not repeated in Figure 5-14. As evidently shown in Figure 5-14, dynamic zones cause the loss of information regarding maneuvering for dumping because in this test no tag was used to represent the conveyor belt.

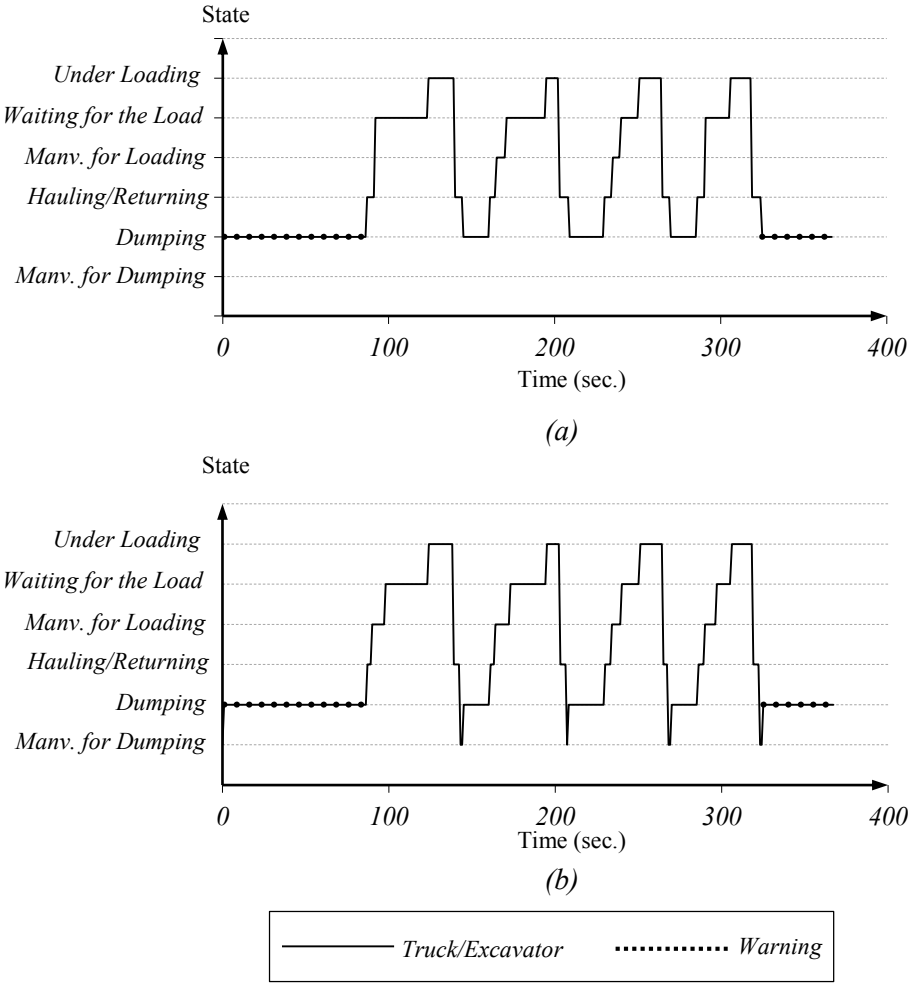


Figure 5-14: Results of State-Identification for Truck for 1 sec. Averaging Time Span (a) without Fixed Zones and (b) With Fixed Zones

The output of the state-identification is presented as shown in the column "*after*" in Table 5-4. For the brevity, only the results of 1-second averaging time-span with the fixed zones are presented. For instance, according to the Table 5-4, the dumping of the truck took 17.67 sec. in average with the standard deviation of 2.36 sec. over four cycles.

As shown in Figure 5-15, the simulation model of the operation was developed using Stroboscope. The data are, then, passed to *NRTS* application for the update of the model.

Table 5-4: The Results of Simulation Before and After Update

Equipment	Simulation Parameters	Duration (sec.)		Productivity (Truck/min)		Total Duration (sec.)	
		Before	After	Before	After	Before	After
Excavator	Relocation	Norm[3, 0.5]	Norm[12, 0]	1.12	0.96	199	249.32
	Swing to Load	Norm[14, 2]	Norm[20.5, 1.5]				
	Loading	Norm[8, 1.75]	Norm[10, 2]				
	Swing to Truck	Norm[14, 2]	Norm[15.25, 5.63]				
	Dumping	Norm[12, 2]	Norm[11.67, 4.19]				
Truck	Maneuvering for Dumping	Norm[3, 0.5]	Norm[2, 0]				
	Dumping	Norm[6, 1.5]	Norm[17.67, 2.36]				
	Returning	Norm[4, 1]	Norm[3.62, 0.48]				
	Maneuvering for Loading	Norm[3, 0.5]	Norm[7.5, 1.12]				
	Hauling	Norm[5, 1]	Norm[3.62, 0.48]				

In the *NRTS* application, the user defines how the state information relates to the parameter defined in the simulation model. The application then runs the simulation model with the new input and presents the results. The improvement in the simulation model, when the estimated durations are replaced with the actual durations, is demonstrated in Table 5-4. With the initial estimates the entire operation should have lasted 199 seconds, while the experiment actually lasted for 285 seconds. However, the update of the simulation suggested that the operation will take 249.32 seconds which represents more than 17% improvement of simulation accuracy. It is noteworthy that the durations of the initial simulation model are usually based on the historic data and expert's estimation. Accordingly, in this case study, too, the assumed values presented in Table 5-4 are based on the specification of the used equipment, and the designed site layout.

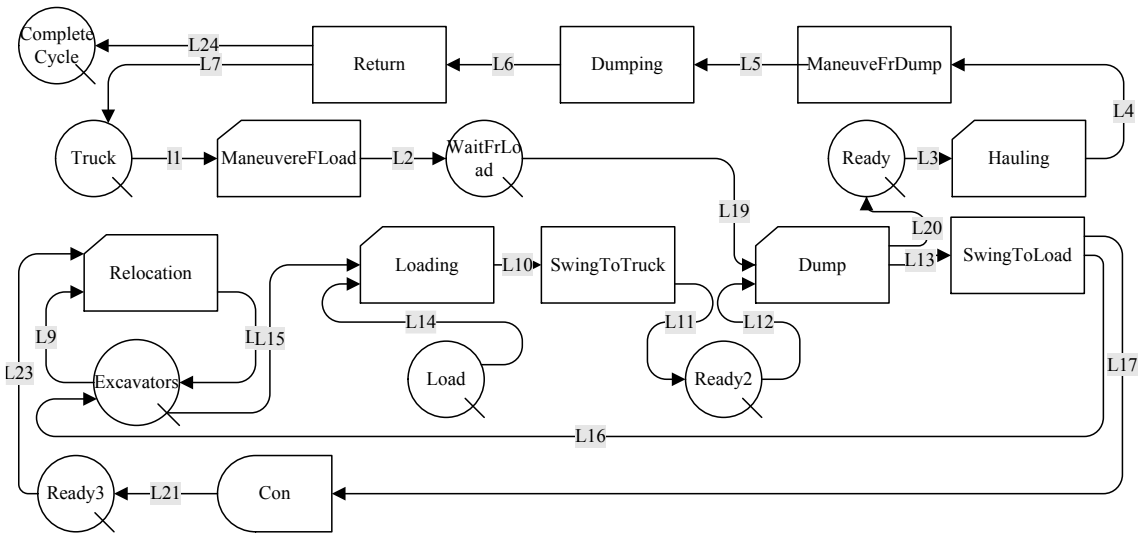
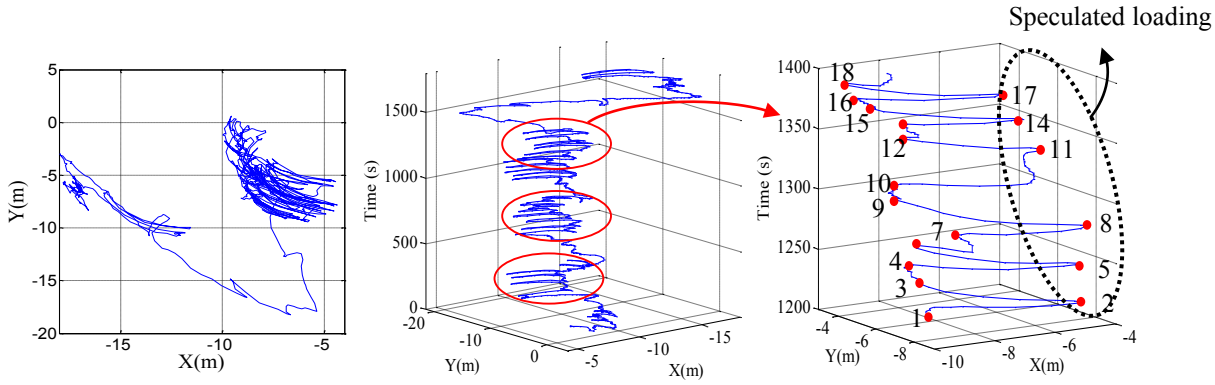


Figure 5-15: Simulation Model for the First Case Study

5.5.2 SECOND CASE STUDY

The second case study was performed to validate the accuracy of the proposed state-identification rules for an actual project where an *AMG* excavator was used. A 30-min log of *GPS* data were obtained from an *AMG*-enabled Hitachi EX-3600 hydraulic excavator operating at the Obed Mine in Alberta. It is noteworthy that the obtained data came with no elevation data but had a limited sensory data that indicate the tram mode of the equipment, i.e. relocation. The 2D and 3D plots where time is used as the third dimension of data are shown in Figure 5-16.



(a)2D View *(b)2D and Time* *(c)Magnified View*

Figure 5-16: GPS Data of the AMG Enabled Excavator

As can be seen in Figure 5-16(a) and (b), some swinging activities can be discerned in the marked areas. Figure 5-16(c) provides a magnified view of a number of swinging actions. The starting and ending points of swinging are marked with enumerated points in Figure 5-16(c), which are identified through visual detection of sharp edges. With the coordinates of these points known, the elaborated status of the excavator can be induced as shown in Table 5-5. For instance, it can be inferred from Table 5-5 that during loading activities (from point 3 to 4) the coordinates of the *GPS* do not change greatly, i.e. small length of move, and thus the speed is low, while during swinging (from point 4 to 5) the moving speed is much higher. A closer scrutiny of Figure 5-16(c) suggests that subsequent to the dumping represented by point 8 the excavator relocates to a new loading point.

Table 5-5: The Details of One Cycle of Excavator Based on Visual Analysis and State-Identification Rules

State	Duration based on Visual Analysis (s)	Duration based on state-identification rules (s)
Swinging to Load	7	5
Loading	17	19
Swinging to Truck	8	5
Dumping	N/A	3
Total	32	32

The state-identification rules were applied to the isolated section shown in Figure 5-16(c), using the 2D location data and sensory data that determine the tram mode. However, with the information about the position of trucks missing, it was required to speculate a loading zone based on Figure 5-16. A loading zone was defined in an area that encompasses edges on the

right-hand side of Figure 5-16(c). The results, shown in Figure 5-17, present a good account of the cyclic motion of the excavator.

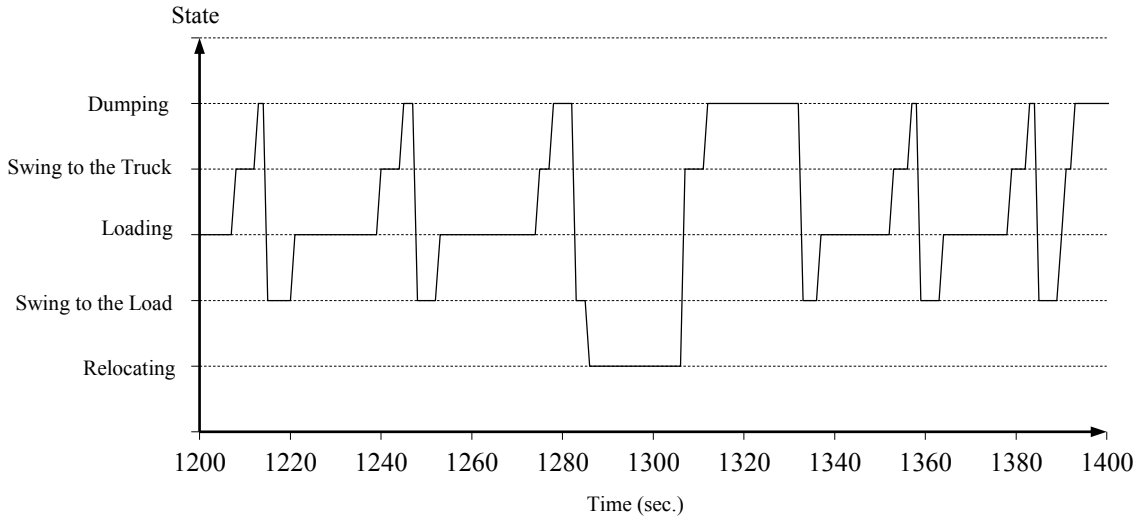


Figure 5-17: Results of State-Identification for the Excavator

For a better visualization, Figure 5-17 depicts the result for only an excerpt, i.e. 6 cycles, of the overall data, which encompasses 24 cycles. The comparison of the results of visual analysis and state-identification rules for one cycle are presented in Table 5-5. Despite minor discrepancies at the level of state durations, which is partly due to the fact that dumping state is embedded in two swing motions in the case of visual analysis, the total durations of the cycle match perfectly. This validates the accuracy of the state-identification rules, suggesting that the rules apply to actual operations on site as well as to the controlled laboratory test. The overall results of state identification for the entire 30 min of the operation are presented in Table 5-6.

Table 5-6: The Result of State-Identification for the Entire 30 Min of the Operation

Equipment	Simulation Parameters	Average (sec.)	Standard deviation (sec.)
Excavator	Relocation	50.25	32.37
	Swing to Load	4.08	1.13
	Loading	24.68	7.51
	Swing to Truck	3.63	1.25
	Dumping	2.95	1.82

5.6 DISCUSSIONS

In the context of *NRTS*, it is of crucial importance to determine the update interval of the simulation model. The update interval has to be selected in such manner that enough number of cycles' data is captured for the analysis and update. While a less-than optimal interval results in a frequent change of schedule and model, and thus is impractical, a greater-than optimal interval leads to a more restrictive leeway for corrective measures. In addition to the sufficient number of cycle times required for simulation update, the determination of the update interval is also influenced by the level of flexibility in the project schedule and the type and level of control severity the manager wants to practice and the duration of typical cycle-times. Basically, the less flexible the schedule and the more risk-averse the manager, the shorter the update interval. Nevertheless, the analysis of cycle data can be performed in much shorter intervals for purposes other than simulation update, e.g. daily site management and safety management, as will be discussed in Chapter 7. The warning capacity embedded in the cycle logic check enables the monitoring of the operation in hourly basis, so as to identify the temporary cause of irregular durations and react to it. On the other hand, the combination of real-time tracking technologies

and near real-time rule-based system can be used in much shorter intervals, e.g. minutes or seconds, to capture the potential safety hazards, e.g. possible collision of two pieces of equipment.

Additionally, as discussed in the chapter, although it is possible to identify the equipment states even without the knowledge of the static zones where different states are likely to occur, the zones offer the ability to capture a more detailed account of equipment states. On this premise, it is more favorable to use a wider area for a zone, so that all the possible relocations are encompassed, than to apply rules for dynamic zones.

It should be highlighted that the notion of *NRTS* is more applicable to operations with a highly repetitive nature, which can be represented by a simulation model. For this reason, heavy earthmoving operations that take place in open mine settings can also benefit from this approach.

Although the proposed approach is primarily built on the vision of future construction sites, where large and medium scale earthmoving projects are dominated by *LGS* equipment and an immense amount of location and sensory data are made readily available, is design is flexible enough to accommodate the application of the cheaper and more readily available *RTLS* technologies, e.g. *UWB*, to enable the existing *AMC/G*-incompatible equipment to benefit from NRTS with minimal retrofitting.

5.7 SUMMARY AND CONCLUSIONS

This chapter presented the proposed methods for the state identification and the NRTS. The proposed method for the state identification is expected to be embedded in every *OA* in order to enable them to provide the required information for the execution of *NRTS* by the responsible *TCA*. Intended mainly to only demonstrate the feasibility and applicability of the proposed

functionalities, the implementation of these methods were presented outside the context of *MAS* structure at this stage of the research. The proposed *NRTS* approach enables the application of various types of tracking technologies for fine-tuning the initial assumptions based on the captured data from the actual operation. To validate the proposed approach, a prototype was developed and implemented. Two case studies, one in the lab environment and with *UWB* technology and another based on the data from the actual project where *AMC/G* enabled equipment was used, were conducted to demonstrate the strength and flexibility of the proposed approach. In the light of the results this research, and the shortcomings of the existing proposed *NRTS* methods, mentioned in Section 2.2.3, the following conclusions can be made: (1) The proposed overarching *NRTS* approach provides a tracking-technology-independent method for processing, analyzing, filtering and visualizing the equipment states that can work with various types of *RTLS* technologies and under the availability of different levels of sensory data; (2) The developed rule set is able to capture the finest details of truck and excavator motions through the concurrent consideration of the fleet as a whole, and for several scenarios where various range of sensory and location data are individually or collectively deployed. Furthermore, pattern analysis and cycle logic test further improve the results of the state identification; and (3) The case study demonstrated the feasibility of the proposed approach.

The developed *NRTS* approach, as presented in this chapter, does not perform distribution fitting to the newly captured data from the site and assumes that the statistical nature of data remains intact in the course of the project. However, this functionality can be added to the Model Refiner component so as to enable capturing the variation in the statistical characteristics of various elements of the cycle-time.

CHAPTER 6: DYNAMIC EQUIPMENT WORKSPACE

6.1 INTRODUCTION

As stated in Section 2.5, the safety application of *LGS*-enabled equipment is not sufficiently addressed in the literature. The existing methods for the provision of safety supports for the operators of earthwork equipment tend not to take a full advantage of the combination of valuable pose, state, geometry, and speed characteristics of the equipment to accurately estimate the shape of *DEW*s, as explained in Section 2.5.2. Consequently, the present chapter aims to leverage a set of information regarding the geometry, pose, state, and speed characteristics of the equipment, which is the output of methods presented in Chapters 4 and 5, to determine the shape and size of the workspace based on the required stoppage time of the equipment so as to secure the early identification of potential collisions while making a more economic use of space.

The structure of the chapter is as follows. First, the *DEW* generation method is elaborated. Next, a case study is elucidated as a means to validate the proposed method. Finally, the conclusions are presented.

6.2 PROPOSED METHOD

*DEW*s aim to use the pose, state, geometry, and speed characteristics of the equipment to generate a space around the equipment that would allow the prevention of immediate collisions with workers on-foot, other pieces of equipment or obstacles on site, considering Equipment Stoppage Time (t_s). t_s can be used to determine how much of the space in the moving direction of the equipment is unsafe after the operator becomes aware of a potential collision. According to

the definition presented by AASHTO (Fambro et al. 1997), t_s has two main components, namely reaction time and braking time. The reaction time pertains to the human factors and denotes the time from the perception of the warning to the application of the break, during which the equipment continues travelling with its current speed and acceleration/deceleration. The research of reaction time is a long-standing trend for urban vehicles in traffic engineering (Baerwald et al. 1965; Taoka T. 1989; Barfield and Dingus 1998). For instance, Gazis et al. (1960) and Wortman et al. (1985) specify 1.14 s and 1.30 s, respectively, as the mean reaction time for an unalerted driver. The breaking time, on the other hand, is the time required for the equipment to come to a complete halt after the breaks are applied, during which time the speed of the equipment declines to zero from its current value. This component of t_s is more pertinent to the mechanics of the equipment and road conditions. Nevertheless, the research about the stoppage time is scarce for construction equipment where multiple *DOF*s should be considered. Guenther and Salow (2012) suggested a value between 0.7 s to 1.5 s as a suitable stoppage time for mining excavation equipment. In another study, Stentz et al. (1999) proposed a value between 2 s to 3 s as a value that is empirically found suitable for t_s for excavators. In this research, the value of t_s is considered to be 2 s.

Although according to the definition of t_s, this value includes a period of moving with the current speed and a period of deceleration, in order to simplify the calculation process, this research conservatively assumes that the equipment continues to travel with its current speed and acceleration. Another assumption of the proposed method is that all pieces of equipment are equipped with an *RTLS* so that their poses and states can be calculated accurately. The update rate of *DEW*s is equal to the update rate of the corrected pose data (Δt) coming from the *RTLS* used in the equipment. A rule-based system is used to identify the states of different equipment

with a high accuracy by leveraging a set of equipment proximity and motion rules that determine the states of the equipment, as explained in Chapter 5. Also, a robust optimization-based method that uses geometric and operational characteristics of the equipment is used to improve the quality of the pose estimation, as explained in Chapter 5.

Furthermore, in addition to the *DEW*s of equipment, workers on-foot and semi-dynamic obstacles (e.g., temporary structures), also need to be tracked by means of *RTLS* and to be represented by their own corresponding safety zones to enable effective collision avoidance.

In this research, two types of equipment, namely excavators and trucks, are used as the main types of equipment that are typically used in earthwork operations. Excavators represent equipment that has articulated mechanism with rotational *DOF*s, whereas other types of equipment, e.g., trucks and rollers, are represented by trucks. However, at the abstract level of the discussion, *DEW*s of excavators are used as an example because they are more complicated than those of trucks due to their large number of *DOF*s.

Figure 6-1 shows instances of a workspace that, unlike the *DEW*s shown in Figure 2-20, consider the pose, state, geometry, and speed characteristics of the equipment. In such a workspace, depending on the state of the excavator, different shapes are used. Figure 6-1(a) shows the DEW of an excavator in the swinging state that uses the magnitude of the rotation speed to determine the angle (α). However, ignoring the moving direction of the equipment may result in an uneconomic usage of the space that can be very valuable in a congested site. Therefore, it is recommended to consider the swinging direction of the equipment to differentiate between parts of the space to which the equipment is approaching and parts from which it is moving away. Figure 6-1(b) shows the proposed workspace of an excavator in the swinging state where the direction of rotation is considered to differentiate between the angles of

workspace on the rotation direction (α) and the opposite direction (β). The rationale behind the asymmetric shape of the workspace is that the risk of collision along the direction of movement is much greater than along the opposite direction. Thus, a greater accent should be placed upon the space at the moving direction of the equipment. This research proposes asymmetric workspaces that consider the moving direction of a piece of equipment in each state as explained in the following sections. This arrangement better captures the potentially hazardous space around the equipment while using the space frugally, rendering this type of workspace very suitable for congested sites.

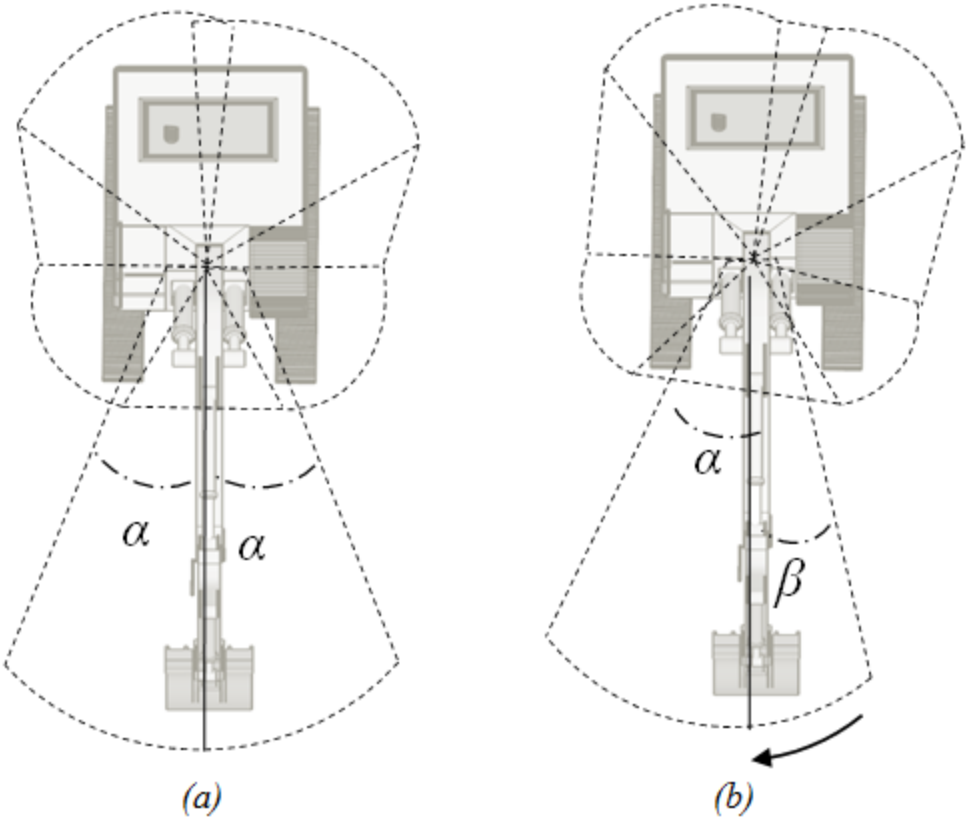
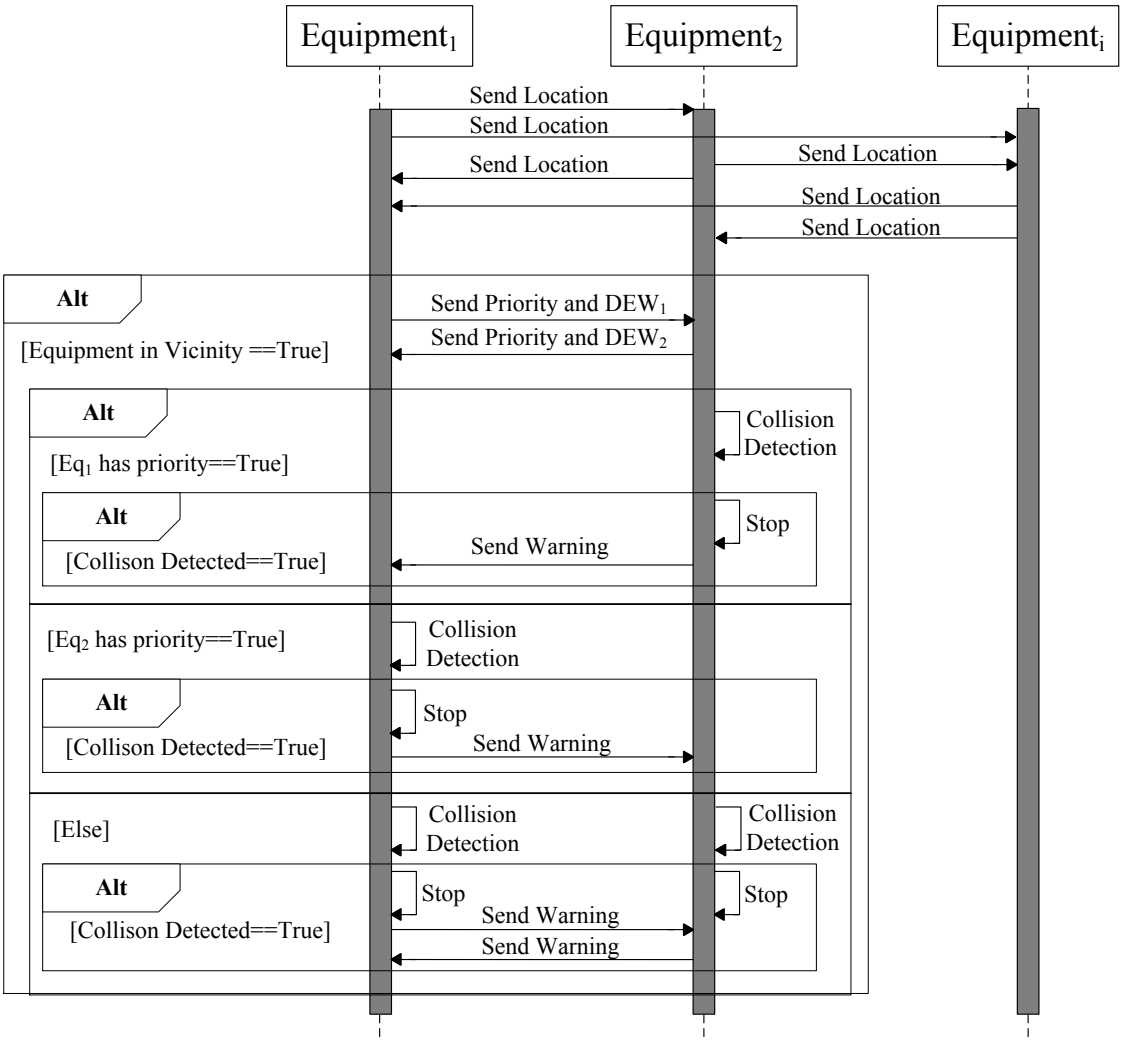
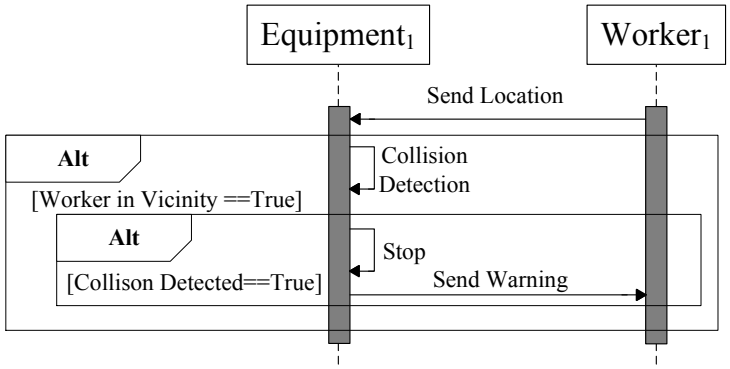


Figure 6-1: (a) Symmetric Workspace, and (b) Proposed Workspace



(a)



(b)

Figure 6-2: Sequence Diagram of Communication between (a) Several Pieces of Equipment, and (b) a Piece of Equipment and a Worker Using DEWs for Safety Control

For the *DEW*s to be effectively used for the purpose of collision detection and avoidance, every piece of equipment needs to be able to generate its own *DEW* and have near-real-time information about the *DEW*s of other equipment and the workers' workspaces. Figure 6-2(a) shows the sequence diagram (Booch et al. 1999) of the communication between different pieces of equipment that enables the near-real-time exchange of *DEW*s and the subsequent collision detection. To avoid redundant computation, the equipment can perform pairwise comparisons only with the pieces of equipment that are in its vicinity. To determine the equipment in vicinity, the multi-layer workspace concept (Chae 2009, Luo et al. 2014, Wang and Razavi 2015) can be applied. In this method the distance between every two pieces of equipment is calculated and if the distance is less than a specific threshold, then the collision detection between their *DEW*s is performed. In order to further reduce the computation efforts and avoid redundant calculations, the priorities of the different equipment can be used to delegate the calculation to the equipment with the lower priority. If a collision is detected between two pieces of equipment, the equipment with the lower priority will stop and send a warning to the other equipment. If both pieces of equipment have the same priority, then both should perform the collision detection and if a collision is detected they should both stop.

Similarly, Figure 6-2(b) shows the sequence diagram of the communication between a piece of equipment and a worker on-foot. It should be clarified that given the workers vulnerability, they always have a higher priority over the equipment. As shown in Figure 6-2(b), every piece of equipment receives the location of the worker and checks for the potential collision between its *DEW* and the cylindrical workspace of the worker. If a collision is detected, the equipment stops and sends a warning to the workers to clear out the dangerous zone. While it is indispensable to account for workers on-foot in addition to the equipment and semi-dynamic structures for

effective collision avoidance on a construction site, the current research focuses only on the collision between equipment. This is because, given the size of the equipment and their inherently more complex kinematics, the interaction between equipment is more complex and more difficult to monitor. Nevertheless, the proposed method can be easily applied to consider the workers as simple cylindrical workspace to avoid all types of collisions on a site.

Figure 6-3(a) and (b) show the flowchart for the generation of the proposed *DEW*s for excavator and trucks. In both flowcharts, with the 3D model of the equipment and its pose and state information available, the method proceeds to determine the linear and angular speeds of the equipment. For instance, as shown in Figure 6-4(a), an excavator can travel on its tracks with the linear speed of $\vec{v}$, move its bucket with the linear speed of $\overrightarrow{v_b}$, or swing with the angular speed of $\overrightarrow{\omega_1}$. Figure 6-4(b), shows the speeds corresponding to the controllable *DOF*s of a truck. The calculation of different speed and acceleration elements of the equipment is based on considering the changes in the pose of the equipment over two consecutive pose estimation data. It is previously shown in Chapter 4 how the pose estimation method can be used to accurately determine the 3D pose of the equipment, which includes the corrected location of the equipment and the orientation of the multiple parts of the equipment. The linear speed elements can be calculated using the difference in the equipment location data over the time between two updates of the pose data (Δt). Similarly, the angular speed can be calculated using the difference in the orientations of the different parts of the equipment over Δt. The relevant acceleration elements can also be derived from these speed elements.

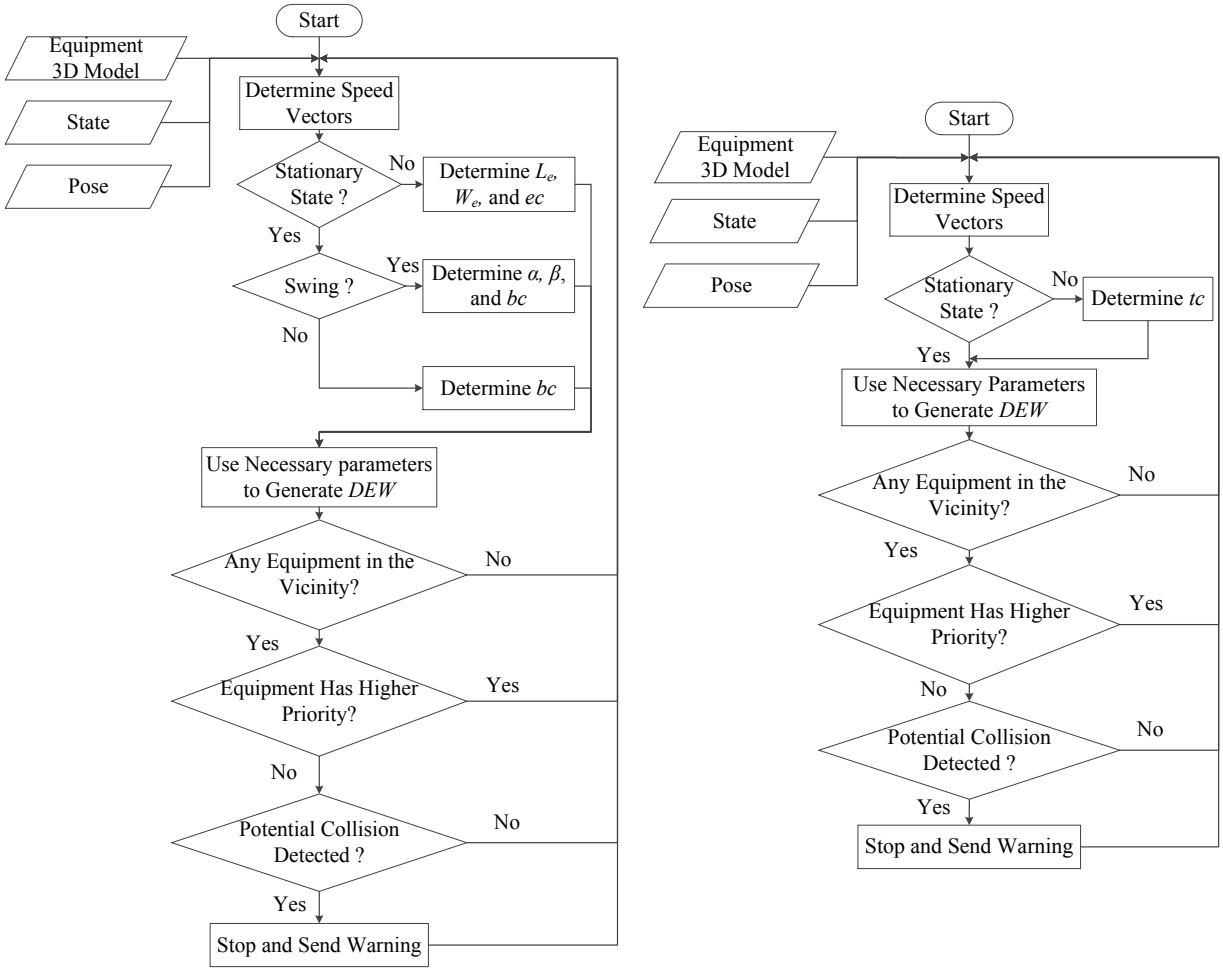
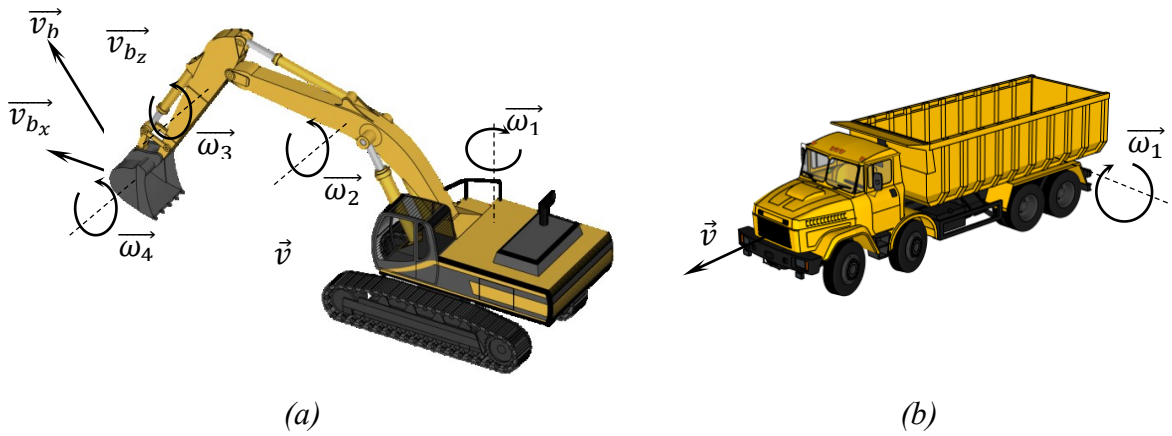


Figure 6-3: Flowchart of the Generation of DEW for (a) Excavator, and (b) Truck

It should be emphasized that the pose estimation method applies the required corrections to the location data to remove the *RTLS* error up to an acceptable level (e.g., 20cm), but there is always a certain degree of residual error in the estimated pose that will propagate through the speed and acceleration calculation. However, as long as the amount of the residual error is within the acceptable level, the calculated speed and acceleration elements are considered reliable.



(a) *(b)*

Figure 6-4: Speed Vectors Corresponding to Controllable DOFs for (a) an Excavator, and (b) a Truck (Models of Truck and Excavator Are Obtained from Google 3D Warehouse (2014))

Upon the determination of the speed vectors, the *DEW* can be generated based on the type of the equipment and the equipment state as explained in the following sections. It should be emphasized that this method determines the shape of *DEW*s based on the assumption that the equipment is going to remain in its present state. Accordingly, the boundary situations, where the equipment is transiting between one state to another are not considered. However, this is tolerable in view of the high update rate of the *DEW*s. The types of the *DEW*s and the parameters that determine their shape are introduced in the following sections. Then, as explained above, the equipment with the lower priority (or both pieces of equipment if they have similar priorities) applies the collisions detection between *DEW*s of equipment in the neighborhood. If the collision is detected, warning is sent to the operators of the involved equipment with a lower or similar priority to stop.

It is also noteworthy that the generation of *DEW*s can be simplified by first calculating its projected shape in the x-y plane and then extruding it along the vertical axis so that the entire range of movement of the *DOF*s along the vertical axis is covered. However, a full 3D *DEW* will be investigated in the future.

6.2.1 DEW OF EXCAVATORS

Two distinct types of states can be identified for an excavator, namely stationary states (swinging, loading, dumping, and waiting) and traversal states (relocating, maneuvering). Usually an excavator can only engage in one of the two types at a time.

As shown in Figure 6-4(a), a typical excavator can be controlled through five controllable *DOF*s resulting in the speed vectors $\overrightarrow{\omega_1}$, $\overrightarrow{\omega_2}$, $\overrightarrow{\omega_3}$, $\overrightarrow{\omega_4}$, *a*nd $\vec{v}$. However, since the workspace calculation is done in the x-y plane, three of the above-mentioned *DOF*s *(*$\overrightarrow{\omega_2}$, $\overrightarrow{\omega_3}$, $\overrightarrow{\omega_4}$*)* can be combined at any point in time to generate the instantaneous linear speed vector at the tip of the bucket *(*$\overrightarrow{v_b}$*)*. This reduces the number of the speed vectors to three *(*$\overrightarrow{\omega_1}$,$\vec{v}$, $\overrightarrow{v_b}$*)*.

6.2.1.1 DEW OF EXCAVATOR IN STATIONARY STATES

When an excavator performs stationary operations, it only moves along either or both of $\overrightarrow{\omega_1}$ and $\overrightarrow{v_b}$. This is because a skillful operator is able to control multiple *DOF*s along $\overrightarrow{v_b}$ while swinging. The shape of the *DEW* is defined based on the identified current stationary state of the excavator (swinging, loading, dumping, and waiting). Additionally, since the tracks of the excavator are not moving during the stationary states, *DEW* is defined only for the upper body of the excavator in these states.

(1) Excavator in swinging state: As shown in Figure 6-5, if an excavator is identified to be in the swinging state with the angular speed of $\overrightarrow{\omega_1}$ and the linear speed of $\overrightarrow{v_b}$, the *DEW* is determined by the corresponding values of α, β and $\overrightarrow{v_{b_x}}$,where α represents the angle along the direction of rotation, β represents the angle in the opposite direction that is reserved for the possible change of swinging direction instigated by unforeseen circumstances, and $\overrightarrow{v_{b_x}}$ accounts for the combined

movements of the boom, stick, and bucket in the vertical plane containing the axes of the boom and the stick.

For the simplicity of the calculation, each part of the equipment can be represented by a tight-fitting bounding box, as shown in Figure 6-5(a). The *DEW* can be generated through the determination of the rotation radius for each bounding box (R_1, R_2, R_3, R_4, R_5), a buffer (b), t_s, the rotation angles α and β, and the bucket motion clearance (bc), as shown in Figure 6-5(b).

R_1 is a variable that is defined as the distance from the excavator's center of rotation to the furthest point on the boom, stick, and bucket axis in the x-y plane at the current time. R_2 to R_5 are fixed parameters that are dictated by the equipment geometry and correspond to the distances from the excavator's center of rotation to the front and rear corners of the upper body of the excavator. b is also a fixed parameter used in order to define *DEW* with a degree of conservativeness. b is proportionate to the size of equipment and can be defined as a percentage of the maximum dimension of the equipment, for example 1% of the length of the equipment, and is applied along the radii R_i. Other factors that may have impact on the value of b are the update rate and the accuracy of the applied *RTLS*. Another buffer (b) is added to the bounding box that contains the boom, stick and the bucket, as shown in Figure 6-5(b).

The angle β represents the amount of swing the excavator will do over t_s if the operator stops the swinging in its current direction and swings in the opposite direction for any reasons. With this definition, β is a function of ω_1, t_s, and the swinging acceleration/deceleration (τ_s), assuming that they are equal. τ_s is a predefined value due to the fact that it pertains to the acceleration and deceleration that are expected to happen in case of swing direction shift. Equation 6-1 can be derived from basic kinematic equations (Forshaw and Smith 2009) for the calculation of β.

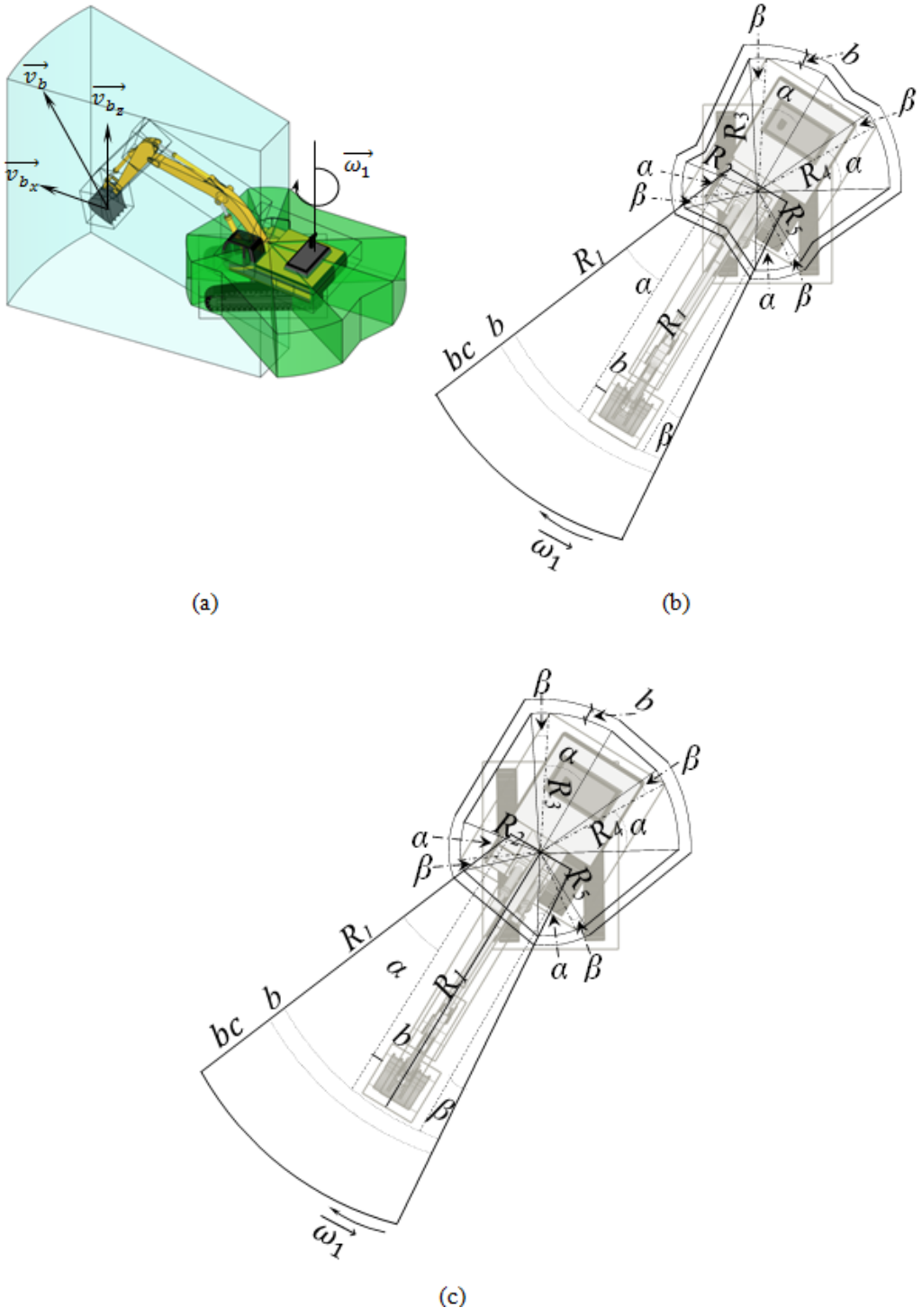


Figure 6-5: Schematic (a) 2D and (b) 3D and (c) simplified 2D Representations of DEW of an Excavator in Swinging State

$$\beta = \frac{1}{2}\,\tau_s(t_s - \frac{\omega_1}{\tau_s})^2 - \frac{{\omega_1}^2}{2\tau_s}$$ **Equation 6-1**

The angle α, on the other hand, denotes the amount of swing the excavator will be doing in the current direction provided it continues with its current speed ($\overrightarrow{\omega_1}$) and acceleration/deceleration (τ_a). Unlike τ_s, τ_a is a value measured in real time because it considers the current actual acceleration. Nevertheless, τ_a will be zero during most of the swinging operation since most excavators tend to reach to the steady state swinging speed quickly and then continue with that speed. Similarly, when the swinging is completed, the excavator decelerates to a complete halt quickly. Based on this definition, α is a function of τ_s, ω_1, and t_s, as shown in Equation 6-2.

$$\alpha = \frac{1}{2}\,\tau_a \times {t_s}^2 + \omega_1 \times t_s$$ **Equation 6-2**

bc represents the clearance buffer for the movement of the bucket along the boom, stick, and bucket axis when the skilled operator is combining the swinging motion with boom/stick/bucket movement away from the excavator (Rowe and Stentz 1997). It is determined by $\overrightarrow{v_{b_x}}$ (the projection of $\overrightarrow{v_b}$ on the horizontal plane), its corresponding measured acceleration (τ_{b_x}), and t_s, as given in Equation 6-3. To generate *DEW* conservatively, *bc* is defined based only on the outward tilting of the combination of bucket/stick/boom movement and it ignores the inward tilting.

$$bc = \frac{1}{2}\,\tau_{b_x} \times {t_s}^2 + v_{b_x} \times t_s$$ **Equation 6-3**

Although the accurate representation of the *DEW* for the swinging state is as shown in Figure 6-5(a) and (b), a conservative simplification can be made to the geometry of the *DEW* by connecting the corners of the pie shapes resulting from the rotations of each corner of the bounding boxes, as shown in Figure 6-5(c).

(2) Excavator in loading/dumping states: Figure 6-6 shows the workspace when the excavator is in the loading/dumping states. Since the excavator's upper body is not swinging in these states, it is enough to reserve space for the movement of the boom/stick/bucket using a buffer. The shape presented in Figure 6-6 is the natural result of the excavator workspace in the swinging state, shown in Figure 6-6(c), when ω_1 and τ_s are zero, and thus α and β are zero. Accordingly, the workspace in these states is determined mainly by *b* and *bc*, where the calculation of *bc* is done similar to the case of the swinging state through Equation 6-3.

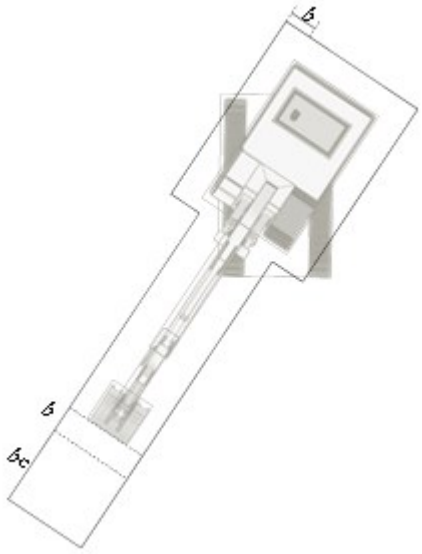


Figure 6-6: Schematic Representation of DEW of an Excavator in Loading and Dumping States

(3) *Excavators in waiting state:* The excavator workspace during the waiting state resembles that of Figure 6-6, with the difference that since the excavator is not engaged in any operations, the value of *bc* is zero.

6.2.1.2 DEW OF EXCAVATOR IN TRAVERSAL STATES

Whereas Figure 6-5 and Figure 6-6 illustrate the basic principle behind the generation of *DEW* for an excavator in stationary states, Figure 6-7 depicts the ruling parameters in forming the

DEW when the excavator is performing traversal operations (i.e., relocating or maneuvering along $\vec{v}$).

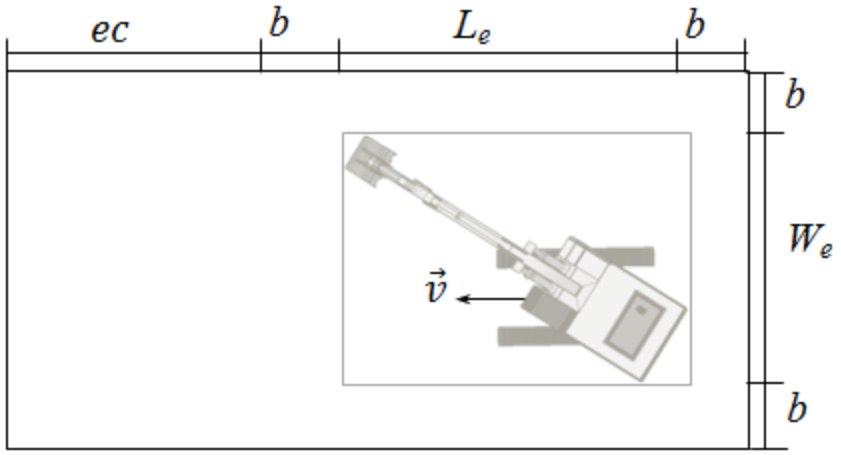


Figure 6-7: Schematic Representation of DEW of an Excavator in Traversal States

The workspace in this case is a box whose dimensions are regulated by (1) the dimensions of a bounding box representing the entire excavator (L_e, W_e) at a given pose, where L_e and W_e are the instantaneous length and width of the equipment, a buffer (b), and the excavator motion clearance (ec). Unlike the workspace in stationary states, where the tracks were disregarded from the *DEW*, in traversal states, the tracks need to be incorporated in the workspace. This is because the tracks are not stationary and can be a source of collision risks. Given that the *DEW* is an instantaneous workspace generated solely based on the pose and the speed characteristics of the equipment, it is defined linearly along $\vec{v}$, even if the equipment is actually moving on a curved path. However, if the construction site has a road network, then the location data of the equipment can be integrated with the road data to ensure that the workspace is following the road alignment. This integration is not currently considered in this research but will be addressed in the future.

The rationale behind b is similar to the one explained earlier in Section 6.2.1.1. *ec* represents the clearance buffer for the movement of the excavator when it moves on its tracks along the speed vector $\vec{v}$. It is therefore a function of v, measured τ_t, and t_s as given in Equation 6-4.

$$ec = \frac{1}{2}\,\tau_t \times {t_s}^2 + v \times t_s \qquad \textbf{Equation 6-4}$$

6.2.2 DEW OF TRUCKS

Similar to excavators, truck can be also engaged in two distinct types of states, namely stationary states (loading, dumping, and waiting.) and traversal states (hauling, returning, and maneuvering).

6.2.2.1 DEW OF TRUCK IN STATIONARY STATES

The stationary states for the truck are loading, dumping and waiting. Due to the simple geometry of trucks, the *DEW* for the truck is basically represented by a box whose dimensions follow the dimensions of the equipment (L_t, W_t) with an additional buffer (b_t). Figure 6-8(a) represents the *DEW* for the truck in stationary states. Note that in the dumping state, an additional buffer should be added to the rear of the truck based on the speed at which the material is being spread, which is a function of the type of soil and the rotating angle of the bed of the truck. However, the present research does not cover the details of this case.

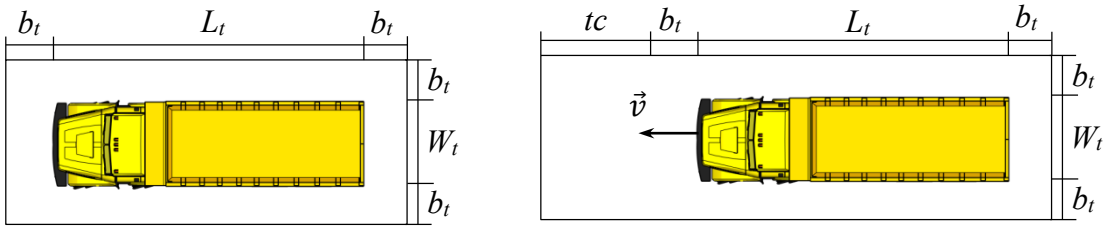


Figure 6-8: Schematic Representations of DEW of a Truck in (a) Stationary States, and (b) Traversal States

6.2.2.2 DEW OF TRUCK IN TRAVERSAL STATES

The truck is in a traversal state when it is hauling the material to the dumping area, returning to the excavation area, or maneuvering at the loading or dumping areas. The *DEW* of a truck in these states is determined by the equipment dimensions (L_t, W_t), a buffer (b_t) and a truck motion clearance (*tc*). *tc* is calculated using an equation similar to Equation 6-4. Figure 6-8(b) represents the *DEW* for the truck in traversal states. It should be noted that when the truck is maneuvering, it may be moving backward. In such scenarios, *tc* is applied at the rear of the equipment.

Once the *DEWs* of all pieces of equipment are generated, it is possible to identify the potential conflicts between them and take the required corrective measures, e.g., alerting the operator or stopping the equipment.

6.2.3 AN EXAMPLE OF THE APPLICATION OF DEW

Figure 6-9 shows a scenario in which a truck enters a site, maneuvers to the loading area where an excavator is digging a trench, gets loaded by the excavator and departs from the site. As shown in Figure 6-9, the orientation and size of the truck's *DEW* change according to its varying direction and magnitude of its speed vectors ($\vec{v}_i$), respectively. It is worth mentioning that not all types of collisions among workspaces and safety zones are actually a safety threat. For instance, as shown in Figure 6-9, while the collision between the workspaces of the truck and the excavator may lead to a safety hazard, an overlap between their workspaces is inevitable as part of the regular excavation work cycle. This limitation is because of the simplified shape of *DEW* along the vertical axis resulting in seemingly overlapping workspaces. However, in future a full 3D *DEW* that accounts for the geometry and kinematics of excavators along the vertical axis will be developed to address this shortcoming.

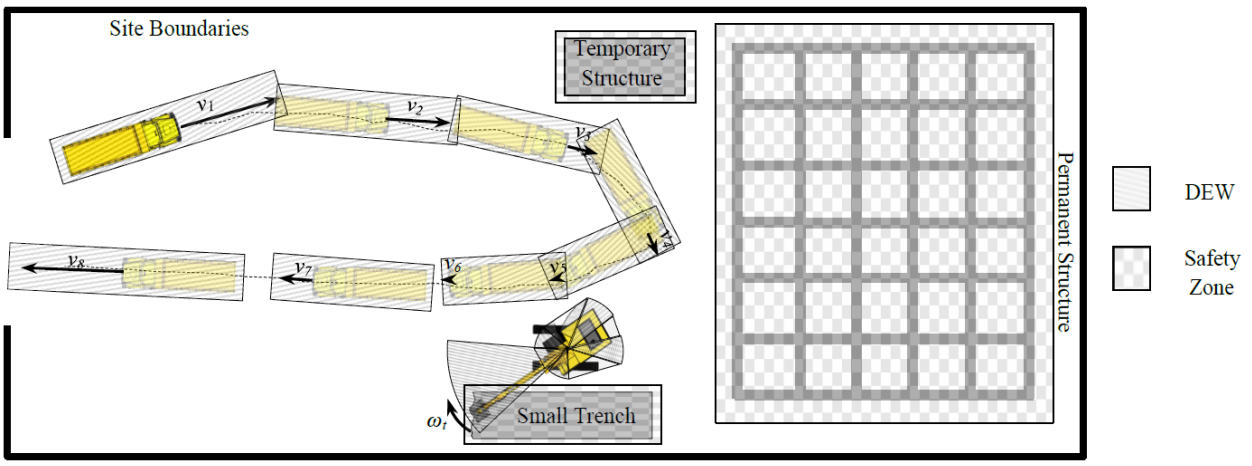


Figure 6-9: Schematic Representation of Safety Analysis based on DEWs

6.2.4 ANALYSIS OF CONGESTION LEVEL

In addition to benefiting the safety of site by preempting the potential collisions between different pieces of equipment, *DEWs* can also serve to calculate an index of the congestion level on the site in preparing site reports. Two approaches are presented in the literature for the quantification of congestion level on the site. Dawood and Malasti (2006) calculated the congestion level through an index named space criticality, which is the result of dividing the summation of *DEWs* sizes by the size of the site. Nevertheless, this index does not capture the temporal aspect of the used space. In another approached proposed by Andayesh and Sadeghpour (2014), the congestion level is represented through the space requirement index, which is the result of the summation of the multiplication of *DEWs* sizes by their corresponding durations. Unlike the previous approach, this approach ignores the volume of the site. Therefore, in this research it is proposed to integrate the two approaches to capture both the temporal dimension of the *DEW*s and the space availability, i.e., the size of the site. For this purpose, if a precise record of *DEWs* for different pieces of equipment and the number of hours they have been working on the site are available, the multiplication of the average volume of *DEW* (V_i) and the equipment working hours (H_i) would indicate how much space the equipment required to perform its work

over its working hours. The summation of these values for all the equipment divided by the multiplication of the site area (A_s) and the overall working hours for which the congestion is being calculated (H_0) would present an index that indicates how much space has been used on the site and for how long. The congestion index (*CI*) can be calculated as shown in Equation 6-5. The greater the congestion index, the more space has been used over the analysis duration.

$$CI = \frac{\sum_{i=1}^{n} H_i \times \mathrm{V}_i}{A_s \times H_0} \qquad \textbf{Equation 6-5}$$

Where *n* is the number of equipment.

6.3 IMPLEMENTATION AND CASE STUDY

A case study was conducted to verify and validate the proposed method for generating *DEWs*. The data from a the lab test presented in Section 4.4.2 were used to demonstrate the generation of *DEW* and its ability to effectively preempt potential collisions between equipment.

The proposed method for the generation of *DEWs* was implemented using Microsoft Excel. The implementation at the present stage does not incorporate the equipment communication structure explained in Section 6.2, and generates the *DEW*s and controls the collisions using a centralized method, where the central platform performs all the computations. The recorded *UWB* coordinates, the corrected pose, and the states of different pieces of equipment are imported into an Excel sheet as the input. The governing equations that generate *DEWs* were developed in Excel, as explained in Section 6.2.1. At every time step, the relevant speeds and accelerations/decelerations of the equipment are measured, and the corresponding *DEW* is generated. For the collision detection between *DEW*s, an automated method was used based on the line segment intersection algorithm (De Berg et al. 2000). In this method, all edges of the two

DEWs are checked against one another for potential intersection using many-to-many relationship. The arc part of the *DEW* is approximated by two segments. *A* collision is detected when a pair of edges is found intersecting. In this case study, which was implemented using a personal computer with an Intel Core i7-2600 CPU (3.40 GHz), the calculation time for the generation of the *DEW*s and the collision detection were measured. The average calculation time and its standard deviation were found to be 23.10 milliseconds and 1.2 milliseconds, respectively. In the generation of *DEW*, the values of t_s, *b, and* τ_s were set to 2 s, 5 *cm,* and *2* $\frac{deg}{s^2}$, respectively. The parameters used for different pieces of equipment are shown in Table 6-1.

Table 6-1: List of Different Parameters Used for the Generation of DEW

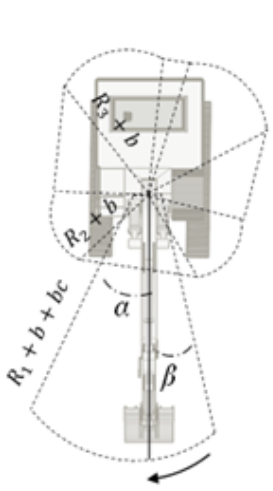
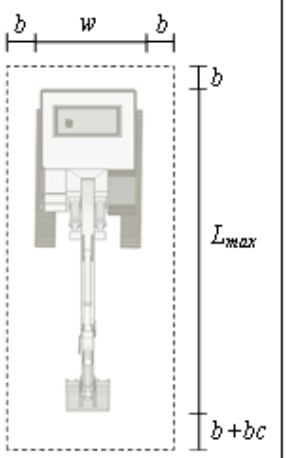
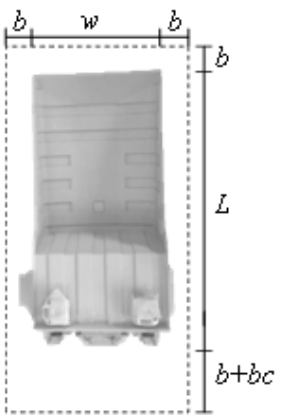
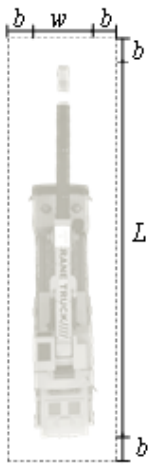
Excavator-Swing		Excavator-Relocation		Truck-Traversal State		Crane-Stationary State	
							
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
R_1	*Variable*	*w*	*20 cm*	*w*	*20 cm*	*w*	*14 cm*
R_2, R_4	*20 cm*	L_{max}	*68 cm*	*L*	*42 cm*	*L*	*91cm*
R3, R5	*25 cm*						

Figure 6-10 illustrates several snapshots of the generated *DEWs* at different stages of the simulated operation. Figure 6-10(a) shows the equipment at the inception of the operation. The locations of attached tags on the equipment are indicated by the cross symbols. The *DEW*s of

different pieces of equipment are shown using the dotted lines surrounding the equipment. The front of the truck is distinguished by the locations of the tags attached to the front of the bed.

Figure 6-10(b) shows the hauling truck and its corresponding DEW. The length of the DEW ahead of the truck is determined by the instantaneous speed of the truck at that point in time. Figure 6-10(c) depicts the excavator at the beginning of the swinging state. The excavator DEW in the relocation state is shown in Figure 6-10(d). Figure 6-10(e) shows a part of the operation when the truck was moving backward to adjust itself for loading. In this case, the extension of the DEW takes place at the rear of the equipment, representing the potential area of collision. Finally, Figure 6-10(f) shows the last phase of the operation where the excavator was intentionally steered towards a collision with the crane. As shown in this figure, the DEWs could be used to successfully identify and warning against the impending collision 4 s before the actual collision. In order to demonstrate the effectiveness of the DEWs, a comparison was made between the proposed method and the alternative types of workspaces shown in

Figure 2-20 and Figure 6-1. The R for the cylindrical workspace and b for the buffer workspace were set to 50 cm and 5 cm, respectively. Table 6-2 shows the results of the comparison, where different methods were analyzed in terms of the average size (the area reserved by the generated workspace in different states), the number of triggered warnings, the number of false warnings (false positive), the number of missed warnings (false negative), and the average time between the warning and the actual collision (collision detection clearance time).

The workspace area is measured in terms of the averaged area and its standard deviation, based on the simplification that the height is the same for all the workspaces. The space saving was calculated through comparing the averaged areas of every method with the area of cylindrical workspace, which is the worst case in terms of the space economy. A collision is defined as any

instances where the distance between the pair of equipment was less than 5 cm. The false positive warning is defined as any warnings that did not entail actual collisions within the next 5 s. The false negative warning, on the other hand, is the count of unwarned collisions. Finally, the collision detection clearance time is computed by finding the earliest warnings prior to a collision.

Table 6-2: Comparison of Different Types of Workspaces

Parameter \ Method	Proposed Workspace	Symmetric Workspace	Buffer Workspace	Cylindrical Workspace
Swing workspace area [μ and σ] (m^2)	[0.22 , 0.03]	[0.25 , 0.06]	N/A	N/A
Loading/dumping workspace area [μ and σ] (m^2)	[0.24 , 0.02]	[0.24 , 0.02]	N/A	N/A
Relocation workspace area [μ and σ] (m^2)	[0.49 , 0.06]	[0.49 , 0.06]	N/A	N/A
Overall averaged workspace area [μ] (m^2)	0.25	0.27	0.22	0.78
Space savings (Compared to cylindrical workspace) (%)	67.95	65.38	71.80	0
False positive warnings (%)	24.53	25.00	29.24	68.81
False negative warnings (%)	0	0	0	0
Collision detection clearance time [μ and σ] (s)	[4.28, 1.16]	[4.43, 1.05]	[3.28 , 1.58]	[5.0 , 0]

It can be discerned that the proposed workspace takes less space than the symmetric workspace, with the space saving of 67.95%, and the least false positive, i.e., 24.53%, which represents the reliability of the workspace. Another interesting observation is that both the proposed and the symmetric workspaces perform efficiently by successfully warning against every collision within the average of more than 4 s, only 1 s less than the best case that belongs to the cylindrical workspace. In the observed clearance times, no instance with a clearance time less than the stoppage time was recorded for the proposed, symmetric and cylindrical workspaces. Only in the case of the buffer workspace, one instance with the clearance time of 1 s was observed. Therefore, the proposed workspace always provided enough clearance for collision avoidance.

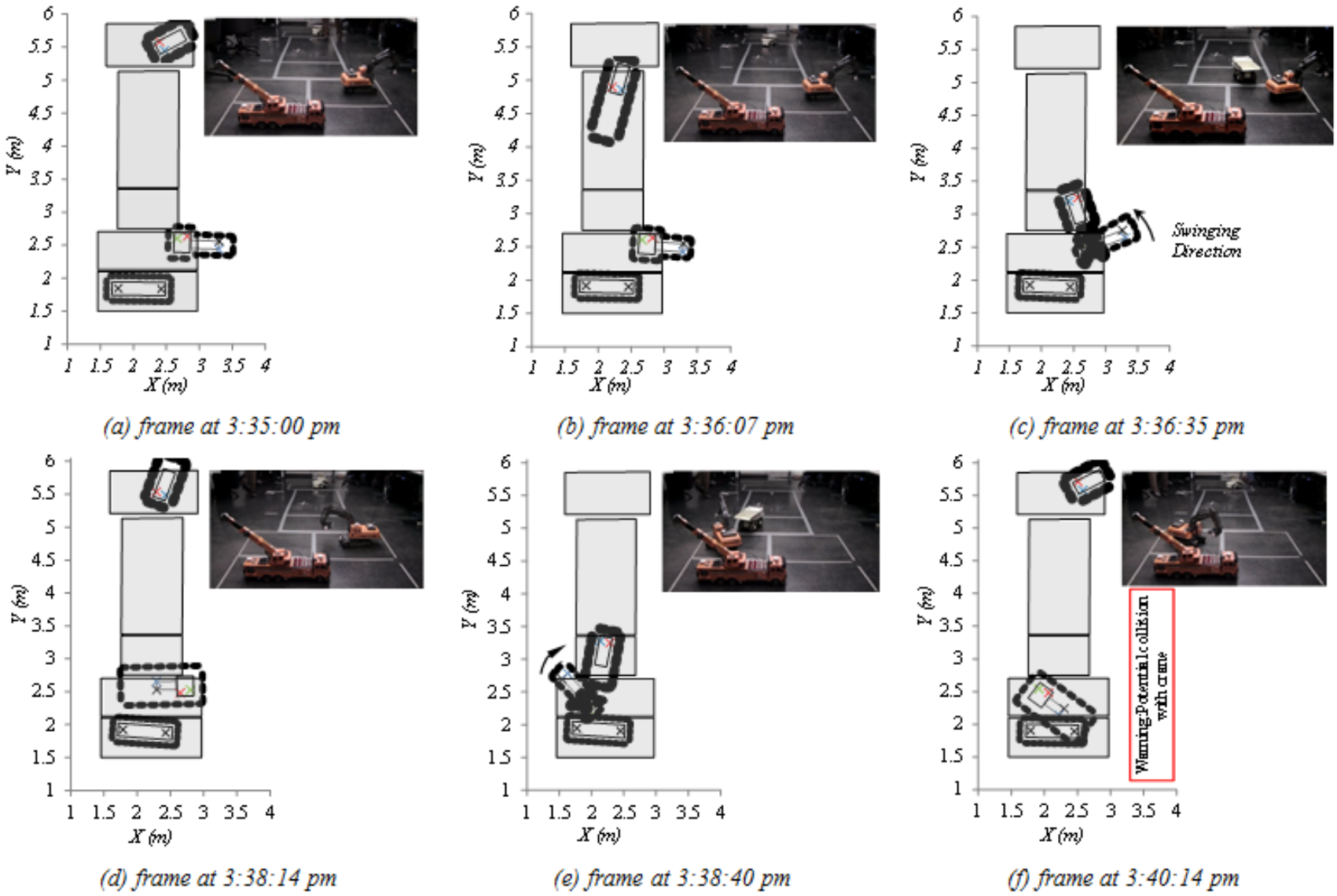


Figure 6-10: Results of Generated DEWs of the Case Study

While the cylindrical workspace provides the best collision detection clearance, it occupies more than 3 times space than that of the proposed workspace and tends to trigger a considerable number of false warnings. The buffer workspace outperforms other types of workspaces in terms of the space economy, but has the least clearance time. Although the improvement of the space saving in the proposed workspace is not significantly more than the symmetric workspace, the difference is determined mainly by the length of the stoppage time. In this case study, the stoppage time was set at 2 s, but should the stoppage time be increased, the difference in the space saving is expected to rise noticeably.

In all types of workspaces, no false negative is observed. This phenomenon can be explained in the view of the nature of the workspaces, which is to create a safety buffer around the equipment. As such, the collision between workspaces always happens prior to the actual collision. However, false negatives can happen if the communication network between the equipment is disrupted or the update rate of the corrected pose data is less than the stoppage time. In the case study, none of these cases happened, and the proposed method assumes that the robust infrastructure is available for the generation of *DEW*s.

Given that the operation took nearly 6 min, the overall area of the site was 18 m^2, and all pieces of the equipment have been present on the site during this time, the required data for the calculation of the congestion index can be collected as shown in Table 3. Using Equation 6-5, the overall space usage and the congestion index are calculated as 0.034 ($m^3 \times hours$) and 0.019 ($\frac{m^3 \times hours}{m^2 \times hours}$), respectively.

Table 6-3: Values Used for the Calculation of the Congestion Index

Parameter \ Equipment	Excavator	Truck	Crane
Working hours (hour)	0.10	0.10	0.10
Average volume of *DEW* (m^3)	0.125	0.053	0.162
Overall Occupied Space (m^3×hour)	0.013	0.005	0.016

Although the presented case study is a good indication of the efficiency of the proposed method in terms of reducing the number of false alarms and more effective use of space compared to other types of the existing workspaces, the scale of the equipment used in the case study posed some limitations. First, the effect of the residual error on the accurate pose estimation of the scaled equipment is relatively larger than the case of the actual equipment, due to the small size of the equipment. Second, although it was tried to introduce some noise to the *UWB* tracking system in the laboratory by partially obstructing the direct line of sight between sensors and tags, there is a need to test the tracking system under the conditions of actual construction sites. However, applying such a test will require careful considerations of some technical and logistical difficulties that might influence the performance of the *UWB* system, including the proper calibration of the *UWB* system under the pressure of the actual construction work, the setting of the *UWB* cables, and the smooth collaboration with the contractors.

6.4 DISCUSSION

The main contribution of the presented method for the generation of *DEWs* is the use of equipment's state, speed, geometry and pose data to economically mark the safety workspace around different pieces of equipment. It is shown in the case study that the proposed method is capable of warning against all potentially hazardous proximity without using the space over-conservatively or generating too many false alarms. Nevertheless, there is a positive correlation

between the level of congestion and the rate of false positive alarms generated by the *DEW*s. However, this is coming from the nature of the problem rather than the characteristics of the proposed method. Although the rate of false positive rises with the increase in the congestion level, the rate would still remain lower than the rate of false positive alarms generated by the conventional methods, e.g., cylindrical workspace. Also, it should be highlighted that generally with the increase in the congestion level, the average speed by which the equipment travels on site will also decreases, resulting in a smaller average *DEW* area, which in turn results in a smaller chance of false alarms.

6.5 SUMMARY AND CONCLUSIONS

This chapter proposed a novel method for the generation of real-time dynamic equipment workspaces considering the pose, state, geometry, and the speed characteristics of the equipment. This method is built on the results of previous chapters, where robust methods for the calculation of pose and state of different pieces of equipment based on *RTLS* data were presented. The present method considers the required operator stoppage time to determine how much space needs to be reserved in order to ensure that the equipment will not collide with other pieces of equipment in the immediate future. Excavators and trucks were used as the representatives of different types of equipment used in an earthwork project. The appropriate *DEWs* and their calculation process for all possible states of the equipment were presented. Finally, the application of *DEW*s for the calculation of congestion index was discussed.

In view of the results of the case study, it can be concluded that: (1) the proposed method is providing a balance between economic use of space and the ability to warn against potential collisions in an effective manner using the pose, state, geometry, and speed characteristics of the equipment, (2) the flexibility of the method in using more than one speed vector in the

calculation of *DEWs* enabled effective capturing of the operation of skilled operators where multiple *DOF*s can be used simultaneously.

Finally, some false warnings resulted from capturing the movement along various *DOF*s only in 2D. Therefore, the future efforts can be dedicated to avoiding this problem by considering the details of the movement in the third dimension in the generation of *DEW*s.

CHAPTER 7: LOOK-AHEAD EQUIPMENT WORKSPACE

7.1 INTRODUCTION

As stated in Section 2.5.3, although *DEWs* are an adequate means to preempt potential collisions in a proactive manner, their real-time nature renders them useful only to trigger warnings or immediately stop the equipment. On this premise, they do not provide the predictive characteristics to foresee the equipment motions for a long enough period to enable path re-planning of the equipment.

Accordingly, the objectives of the present chapter are: (1) Developing a novel method to generate equipment risk maps based on the integration of the proximity-based risks and visibility-based risks using the pose and state data of the equipment and the *NRTS*; and (2) Generating *LAEW* based on the equipment risk maps so that the resulting workspaces can be used to perform path re-planning when a potential collision is identified.

The structure of the chapter is as follows. First, the proposed workspace generation method is elaborated, followed by the explanation of the implementation and a case study. Finally, the conclusions and future work are presented.

7.2 PROPOSED METHOD

LAEWs are generated for the purpose of look-ahead re-planning of equipment motions and are updated in near-real-time with an interval of δt. The update interval is a function of the available computational power and the extent to which the future states of the equipment can be reliably predicted. Generally, the larger the value of δt, the greater the chance of the potential changes in

the predicted conditions, and thus the less the reliability of the generated *LAEW*s. In order to put the sensible value of δt into perspective, it is envisioned that it is most effective in a range between 10 s to 1 min. While a value less than 10 s has the risk of being impractical for being too short for the planning of future motions, a value greater than 1 min reduces the reliability of the generated risk maps.

The flowchart of the proposed method for the generation of the *LAEW* of one piece of equipment (equipment q) is shown in Figure 7-1. As shown in this figure, the input of this method comprises the sensory data, the equipment specifications and their accurate 3D models, the current pose and state data generated by the *OA* of the equipment q (OA_q), and future state data coming from the *NRTS* that is performed by the *TCA*. Additionally, the updated 3D model of the site, and the project detailed plan (including the location of different scheduled tasks, their time frame, and the site layout) are available through the *Information Agent*. The updated 3D model of the site includes the *DTM* and the 3D models of buildings and other permanent and temporary structures, which can be easily imported into the virtual model from a *BIM* tool (e.g. Autodesk Revit). Finally, given that a parametric motion planning is proposed to be used for the determination of the future motions of different pieces of equipment (discussed in Section 7.2.1), a set of heuristic rules that define the operation of a skilled operator is also required to be available to each *OA*.

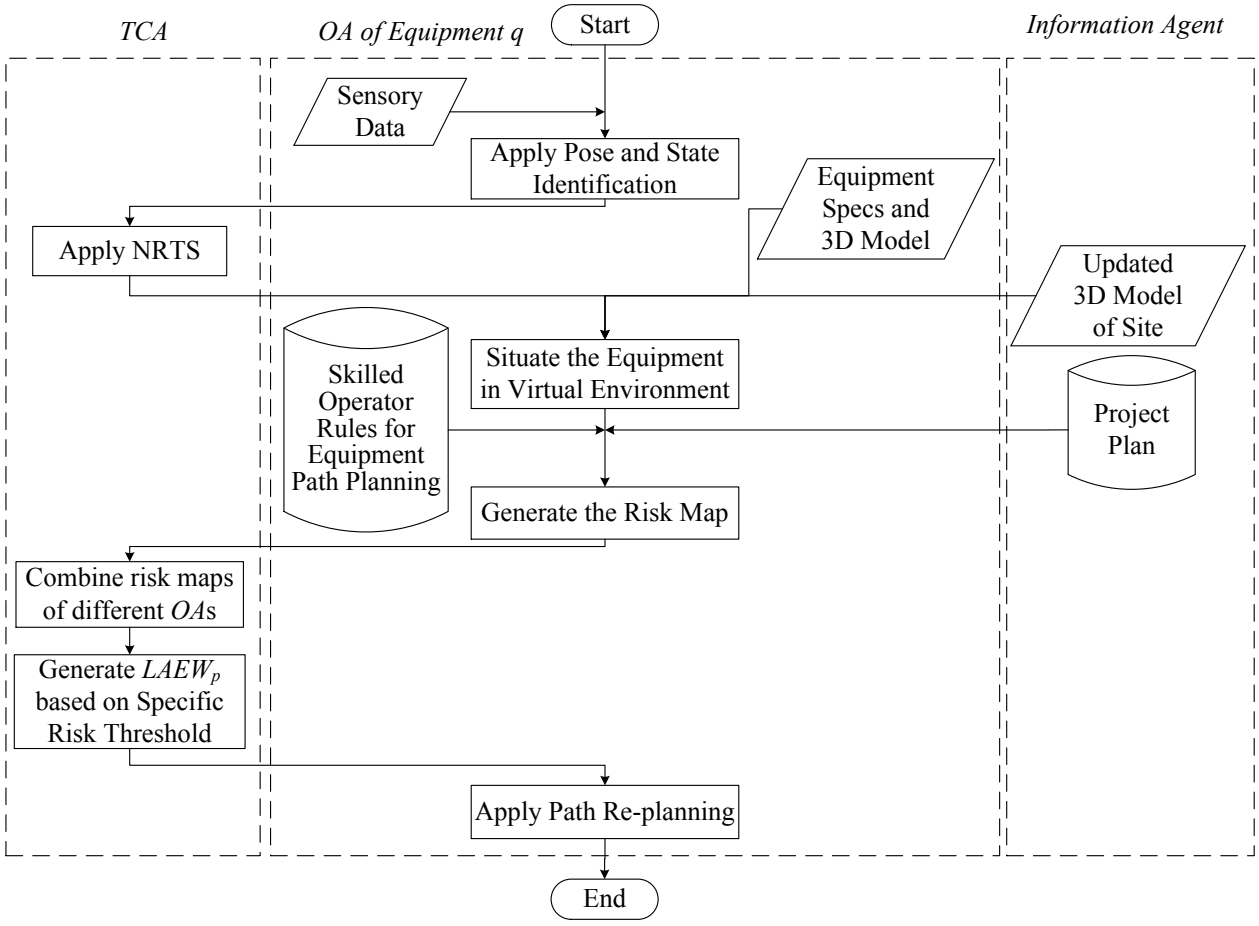


Figure 7-1: Flowchart for the Generation of LAEW

The generation of *LAEW* is based on the discretization of the entire site space into cells, and then calculating the risk associated with each cell given the future expected states of different pieces of equipment, which is performed by each *OA*. As shown in Figure 7-1, the pose data are used to identify the current state, which is then passed on to the *TCA* to perform the *NRTS* in order to generate the operational pattern of each *OA*. These data are then communicated with the OA_q who will first integrate the equipment pose with its 3D model and the updated 3D model of site to situate the equipment in the virtual environment. Then, the OA will use the project plan, and the rules that govern the operation of the machine by a skilled operator to generate the risk map of the equipment, as explained in Section 7.2.1. Finally, the *OAs* transfer their individual risk

maps to the *TCA*s who will first combine these risk maps and then use the tolerable risk level of each *OA* to generate the *LAEW*. The reason why the *LAEW* generation is assigned to *TCA*s instead of individual *OA*s is that given the intensive computational effort required for this process, it is more effective to equip only the *TCA* with high computational resources rather than doing it for all the *OA*s. It should be highlighted that $LAEW_p$ for equipment p is generated based on the combination of the risk maps from all pieces of equipment surrounding equipment p, excluding equipment p itself. $LAEW_p$ can be used by the OA_p to perform path re-planning, if required. Similarly, the path-replanning performed by the OA_q at the end of the flowchart shown in Figure 7-1 is realized through $LAEW_q$.

Chapters 4 and 5 elaborated on the method to obtain near-real time pose and state data. Therefore, the scope of the present chapter begins from the point where all the input data are transferred to the virtual environment of the *MAS* for the determination of the motion path of the equipment. In the following sections, the *LAEWs* of excavators are used to explain the proposed approach.

7.2.1 EQUIPMENT RISK MAP

Figure 7-2 shows the detailed flowchart for the generation of equipment risk maps. The first step is to identify the motion path of the equipment over the next δt (discussed in Section 7.2.1.1). Next, the OA_q uses the generated path to identify the space that could be potentially impacted by the operation of the equipment in the next δt. This space is referred to as the Analysis Space of equipment q (S_q) and is discussed in Section 7.2.1.2. Furthermore, the cells that fall within S_q are identified and reserved for the calculation of the risk pertinent to the operation of the equipment. Upon the completion of this step, three risk indices, namely Shortest Distance to Equipment q ($SDE_{i,q}$), Time to Shortest Distance to Equipment q ($TSD_{i,q}$) and Visibility Index of equipment q

($VI_{i,q}$) are calculated for each cell C_i in S_q. Subsequently, the OA_q combines these risks factors to generate the overall risk map in S_q (discussed in Section 7.2.1.3).

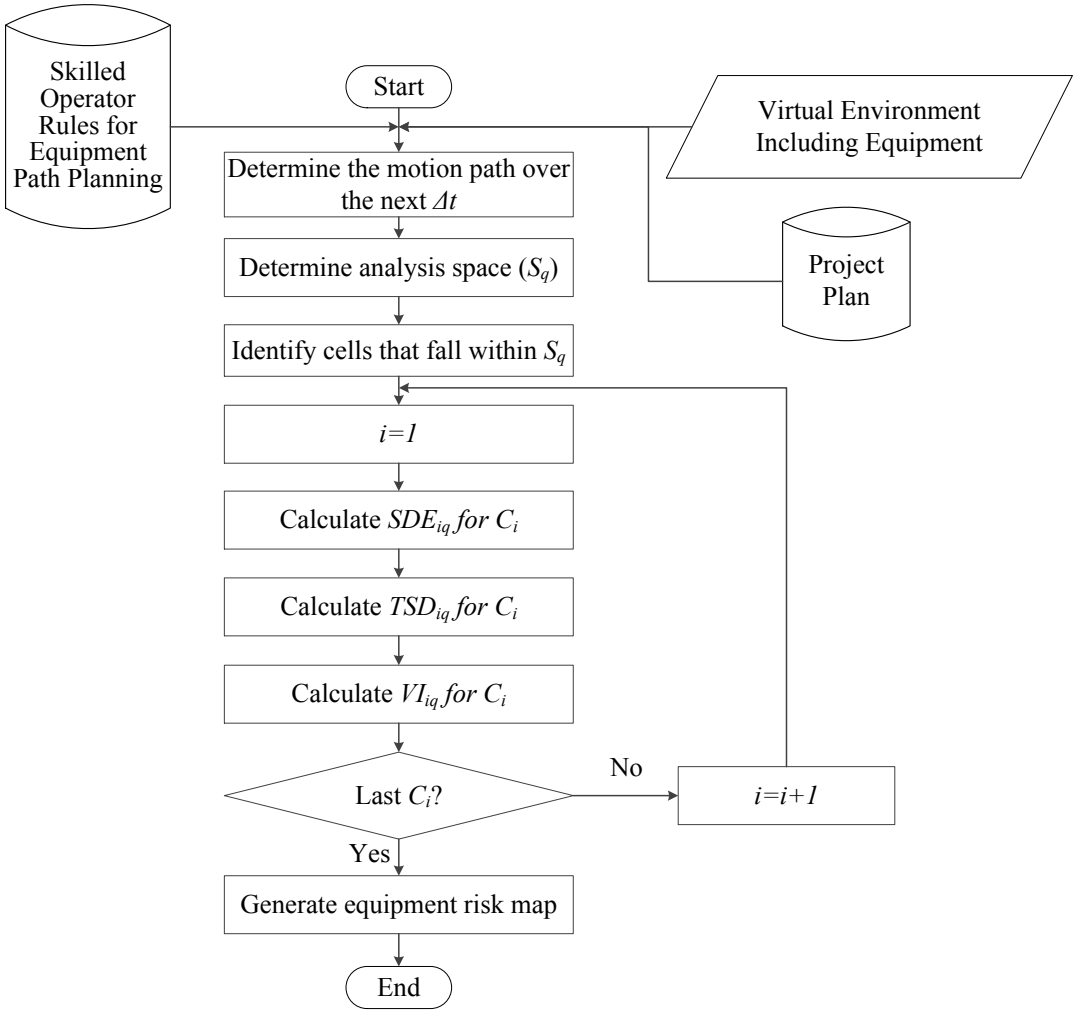


Figure 7-2: Flowchart for the Generation of Equipment Risk Map

7.2.1.1DETERMINING THE MOTION PATH OVER δt

The first step in the generation of the equipment risk map is to determine the path of the equipment over the period of δt. This step intends to combine the temporal data about the future states of equipment, coming from *NRTS,* with their corresponding spatial data, coming from a motion planning algorithm, to simulate the likely future paths of the equipment over δt in the virtual environment. It is crucial to highlight that the path planned at this step of the method is

done based on limited information about the surrounding of each piece of equipment. However, this path will be evaluated based on the consideration of the risk maps of all the surrounding equipment (as will be explained in Section 7.2.1.3) and readjusted if a safety threat is identified. This research adopts the parametric motion planning method (Rowe 1999; Kamat and Martinez 2005; Sarata et al. 2006) that combines inverse kinematic and rules from the motions executed by a skilled operator to find a smooth and realistic path for different pieces of equipment. Although this method is adopted from the abovementioned research, it is elaborated in detail in this section for the completeness of the overall proposed method.

Figure 7-3 illustrates the flowchart for the integration of the *NRTS* data with the parametric motion planning. The inputs of this step are: (1) the dimensions of the equipment coming from the equipment specifications and 3D model, (2) the expected location of the equipment at different states (e.g., loading points and dumping points), (3) the sequence of upcoming states and their expected durations (t_k) for different pieces of equipment based on the results of *NRTS*, (4) the current pose of the equipment, and (5) skilled operator rules that define a natural transition between different poses of the equipment.

The first phase of this step is to identify the key poses that express the equipment at the start of different states. For this purpose, the results of *NRTS* are used to determine the future states of different equipment and the time passed from the beginning of the present state (t'_0), the time to be spent in the present state (t_0), and the expected durations of future states ($t_1 \sim t_k$), as shown in Figure 7-4. As shown in this figure, the updated simulation model is able to provide a reasonable prediction of the future states of different equipment. For instance, the result of *NRTS* for an excavator shown in Figure 7-4 indicates that the equipment is currently in the loading state and it will perform swinging to the truck and dumping over the next δt. Next, the expected location of

the equipment at the determined states and the dimensions of equipment can be used to perform the inverse kinematic to identify these poses as proposed by Rowe (1999), Kamat and Martinez (2005), and Sarata et al. (2006). As shown in Figure 7-5, the pose of the equipment at the start of state k is manifested through the combination of the values of different DOF_j of the equipment $(\theta_{j,k})$.

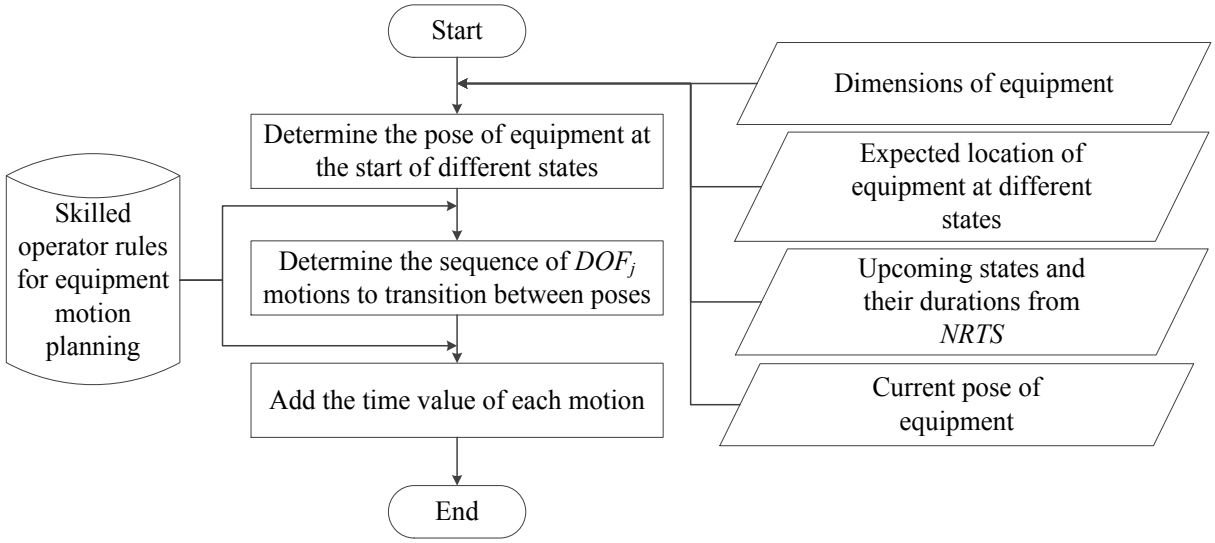


Figure 7-3: Flowchart for the Integration of NRTS data with Parametric Motion Planning

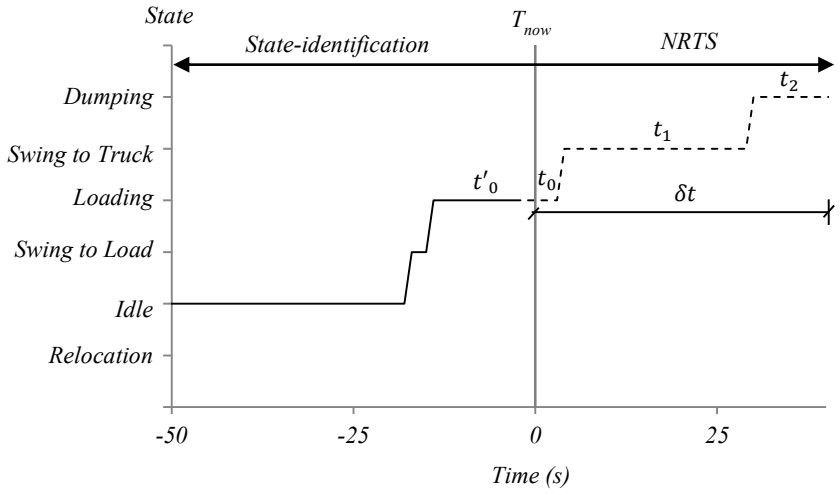


Figure 7-4: Integration of NRTS and State-identification Data

As shown in Figure 7-5 and Figure 7-6, with the poses of the equipment at the start of the upcoming states determined from the previous phase, it can be determined which *DOFs* of the equipment change during each state. The values of these changes can be calculated by taking the difference between $\theta_{j,k}$ and $\theta_{j,k-1}$. For example, the excavator represented in Figure 7-5 would change the following *DOF*s while it is in different states: (1) while loading, it curls the bucket inwards to dig the soil (rotates along θ_1 for $\theta_{1,1}$-$\theta_{1,0}$) and pitches the stick inwards toward the body (rotates along θ_2 for $\theta_{2,1}$-$\theta_{2,0}$), as shown in Figure 7-5(b); (2) while swinging to the truck, it swings the upper body (rotates along θ_4 for $\theta_{4,2}$-$\theta_{4,1}$), pitches the stick outward toward the truck (rotates along θ_2 for $\theta_{2,2}$-$\theta_{2,1}$), and pitches the boom outward toward the truck (along θ_3 for $\theta_{3,2}$-$\theta_{3,1}$), as shown in Figure 7-5(c); and (3) while dumping, it curls the bucket outwards to dump the soil (rotates along θ_1 for $\theta_{1,3}$-$\theta_{1,2}$), as shown in Figure 7-5(d). It is worth mentioning that these steps are subject to change depending on the nature of the operation and the site-specific characteristics.

However, the sequence of these changes and the start time and the duration of changes in each DOF_j are not known from the previous phase. Depending on the skill level of the operator, the changes in $\theta_{j,k}$ can happen with or without overlaps. This is because while a skilled operator is able to control multiple *DOF*s at the same time, a novice operator most likely controls only one *DOF* at a time.

Through observing the skilled operator, the time at which a partial movement of DOF_j starts, represented by the percentage ($a_{j,k}$) of the total duration of state k (t_k) after which the partial movement starts, and the duration of each partial movement, represented by the percentage ($c_{j,k}$) of the total duration, can be determined.

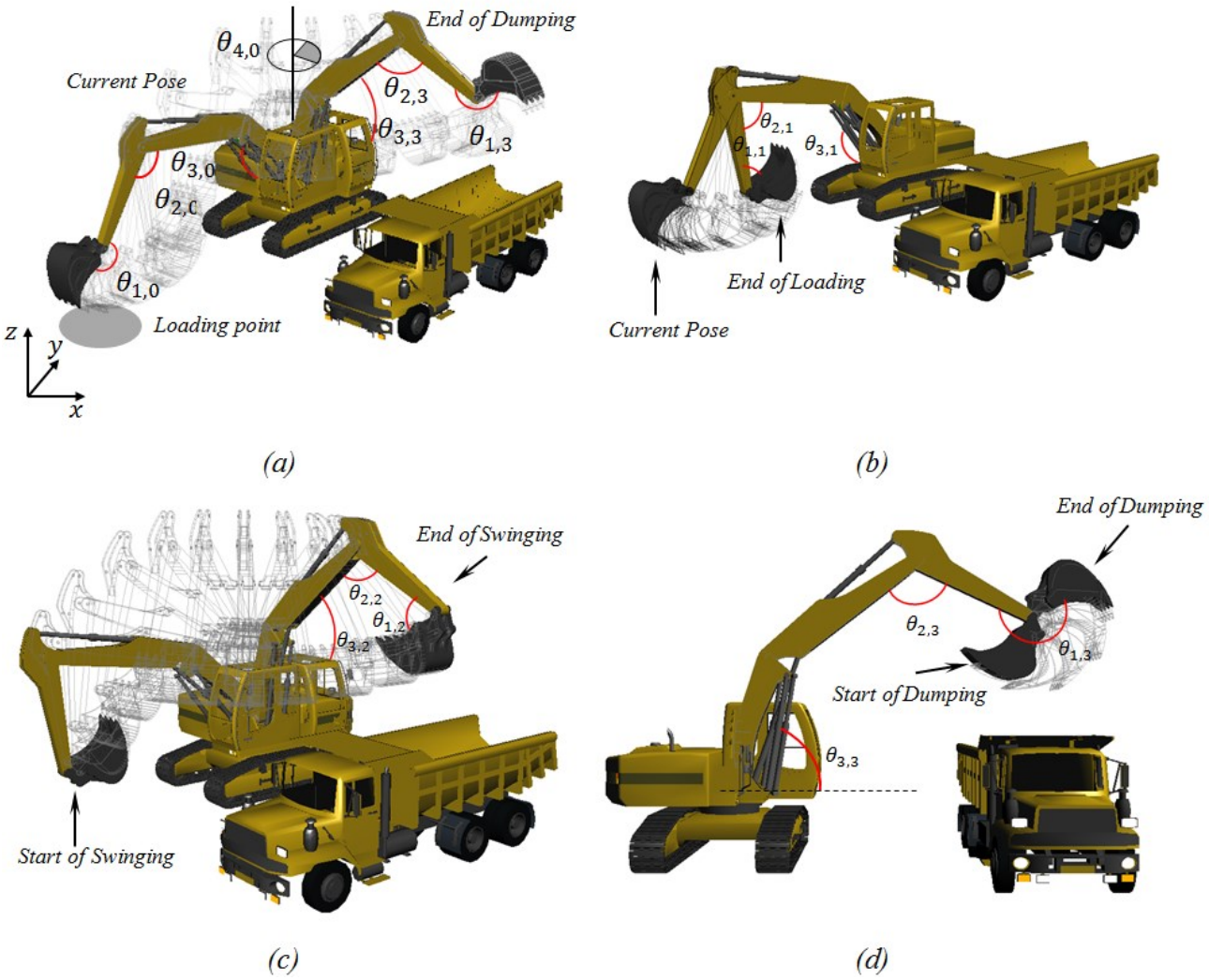


Figure 7-5: (a) Current and Final Poses of the Equipment over δt, and the Motion Paths During (b) Loading, (c) Swinging, and (d) Dumping (Equipment Models Were Imported from Google 3D Warehouse (2015))

Figure 7-6 shows the detailed representation of motion paths generated through the integration of parametric motion planning and *NRTS* for the scenario shown in Figure 7-5 and Figure 7-6. The vertical axis shows the values of the DOF_j and the horizontal axis represents the time. The horizontal lines underneath the graph shows the starting times and the durations of changes in $\theta_{j,k}$ in different states in terms of $a_{j,k} \times t_k$ and $c_{j,k} \times t_k$, respectively. With regard to the first state, a correction needs to be applied in cases where the beginning of δt does not match the beginning of the state. In these cases, the starting time and duration of changes in DOF_j are adjusted

according to Equation 7-1 and Equation 7-2. Notice that the pattern shown in Figure 7-6 is subject to change depending on the skill level of the operator, the depth of the dig, the location of the truck and the conditions of the terrain.

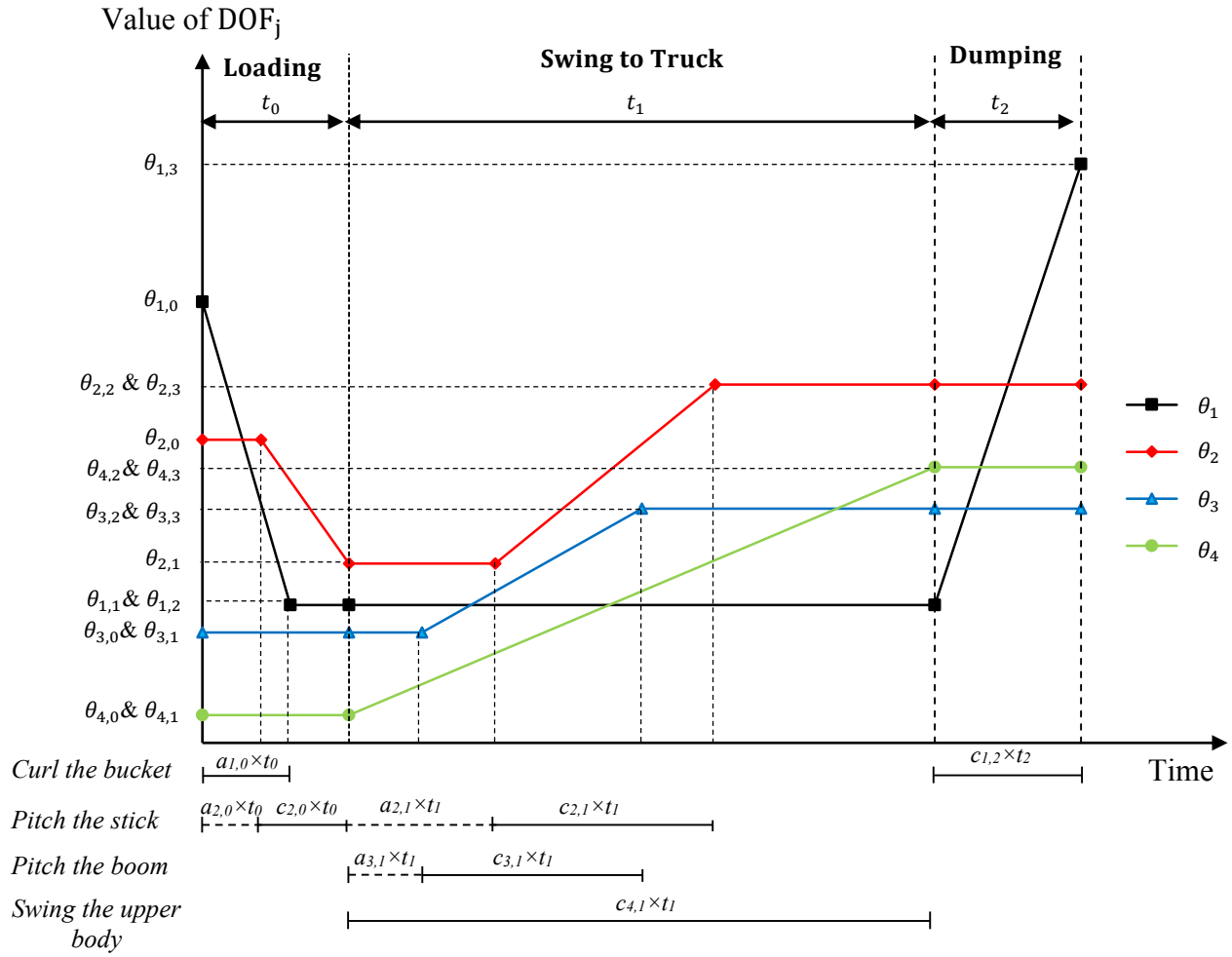


Figure 7-6: Detailed Representation of the Motion Paths Over δt

$$Start\ time = a_{j,0} \times (t_0 + t_0') - t_0' \qquad \textbf{Equation 7-1}$$

$$Duraction = c_{j,0} \times (t_0 + t_0') - t_0' \qquad \textbf{Equation 7-2}$$

7.2.1.2 DETERMINING THE ANALYSIS SPACE (S)

The size of *S* around the equipment is determined by the equipment dimensions, the equipment motion path, the equipment boundaries covering the entire space that can be occupied by the equipment through moving along any of its *DOF*s when performing stationary states (e.g.,

swinging or loading), and a buffer (b). The rationale behind the dependency of the analysis space on the motion path is that if the equipment is completely or partially in a traversal state (e.g., relocation) the analysis space needs to consider the space that is travelled by the equipment over δt in addition to the space determined by the equipment boundaries in stationary states. Figure 7-7 schematically represents this concept for an excavator.

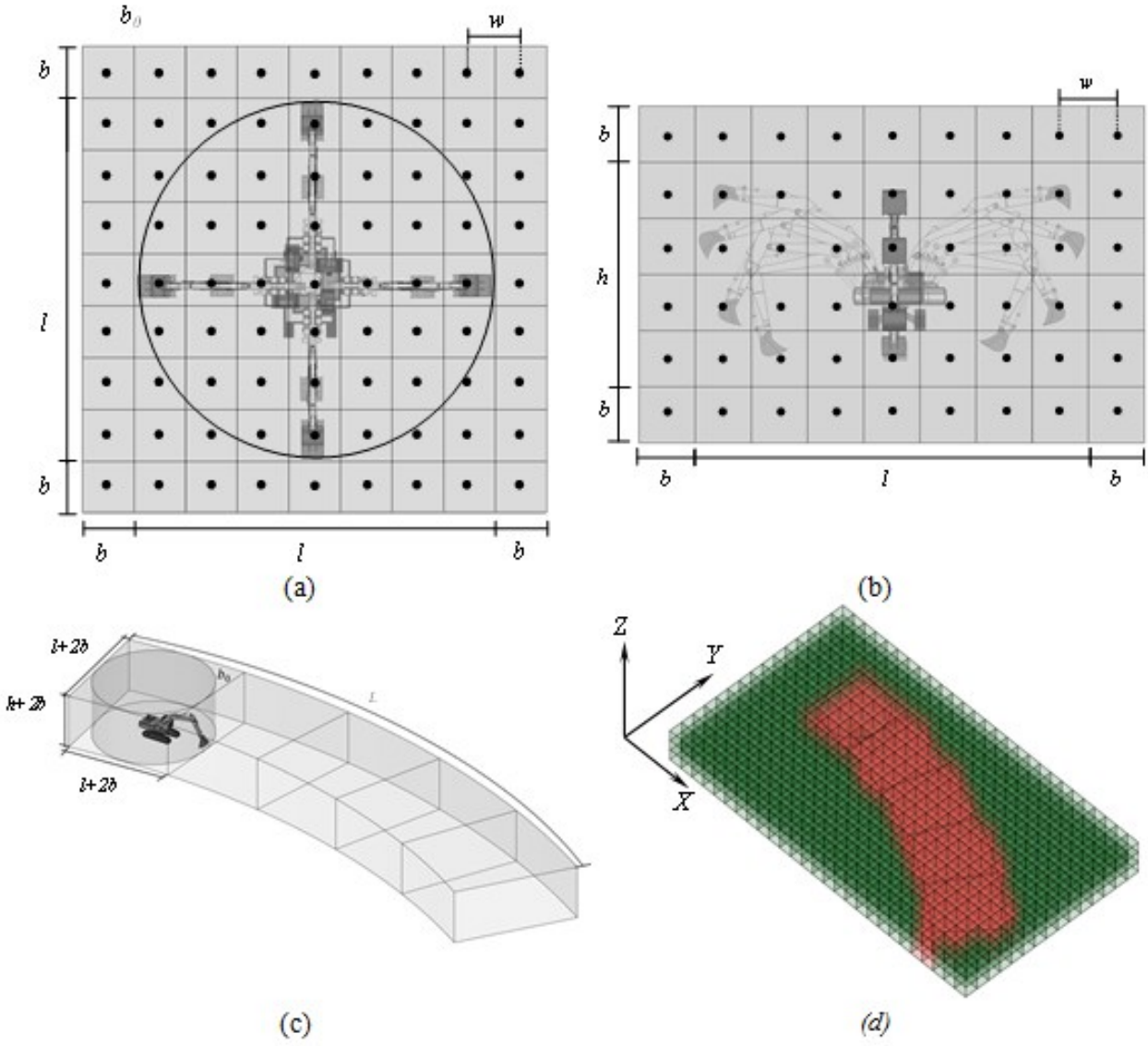


Figure 7-7: (a) Top View of S for Excavator in Stationary State, (b) Side View of *S* for Excavator in Stationary State, (c) *S* for Excavator in Traversal State, and (d) Cells that Fall Within *S*

Figure 7-7(a) and (b) show the top and side views of an excavator, respectively, when it is moving along the extremes of its *DOF*s while it is stationary. The result of the shown boundaries for a stationary excavator is a cylinder, as shown in Figure 7-7(a). However, for the simplicity of calculation and easier discretization of *S*, this space can be conservatively considered as a box (b_0), as shown in Figure 7-7(a) and (c). When the equipment is partially or completely in traversal states, *S* can elongate along the path of the equipment generated in the previous step. Therefore, as shown in Figure 7-7(c), if the equipment is expected to move along a curved path, *S* follows this path. It should be emphasized that depending on the operation-specific characteristics, the equipment can be only in stationary states, only in traversal states, or in a combination of both during δt. *S* represents the conservatively defined potential dangerous area that requires further analysis.

As shown in Figure 7-7(d), once the boundaries of *S* are determined from the previous step, the cells from the global grid that are included in *S* have to be determined. This approach facilitates the integration of analysis spaces of different equipment for path re-planning purposes.

As shown in Figure 7-7(a) and (b), the size of a cell is selected based on the required spatial resolution while considering the available computational capacity. The accuracy of the analysis increases with the decrease in the size of the cells. However, the limited computational capacity available for the near-real-time analysis necessitates the investigation of the optimum size for the cell that yields an acceptable level of accuracy without imposing too much computational efforts on the analysis. The subsequent steps of the proposed method are dedicated to calculating the risk level of cells within *S*.

7.2.1.3 DETERMINING THE RISK INDICES FOR EACH CELL IN S

Once the analysis space is determined, the indices that determine the level of risk for each cell in S can be calculated. SDE_i, TSD_i, and VI_i are considered in this research.

$SDE_{i,q}$ is defined as the distance between the center of the cell C_i and the part of the equipment q that moves closest to the cell during δt. In other words, $SDE_{i,q}$ for C_i is the smallest of the distances between motion paths of each part of equipment q (shown in Figure 7-5) and the center of C_i. Figure 7-8 schematically depicts the concept of $SDE_{i,q}$, $TSD_{i,q}$, and $VI_{i,q}$ using the portion of the excavator operation shown in Figure 7-5(a). Figure 7-8(b) shows the different values of $SDE_{i,q}$ using color coding where red and green represent the extremes of the range, corresponding to low and high values of $SDE_{i,q}$, respectively. As shown in Figure 7-8(c), $TSD_{i,q}$ is defined as the temporal distance from T_{now} to the time when the $SDE_{i,q}$ happens for cell C_i. The rationale is that the shorter the time to $SDE_{i,q}$, the greater the risk for the cell.

Finally, $VI_{i,q}$ is defined as an index representing the visibility of cell C_i over the period of δt in relation to the blind spots of equipment q, considering static obstacles and the equipment q itself. More precisely, $VI_{i,q}$ for cell C_i can be defined as the total time the cell has not been in a blind spot of equipment q over the period of δt. A blind spot is defined as a portion of the equipment's surrounding space that remains invisible to the operator even if the operator has omni-directional view at his eyes' position. The assumption is that the skilled operator knows how to adjust his line of sight inside the cabin not to leave any visible space around the equipment overlooked. Also, since the calculation of risk indices are done by each OA in near real time, the locations of the other pieces of equipment are not considered in the determination of VI_i of a cell. However, this is not an issue since the blind spot of equipment q behind equipment p is only hazardous if

parts of equipment q are planned to reach behind equipment p over δt, which is a very unlikely scenario.

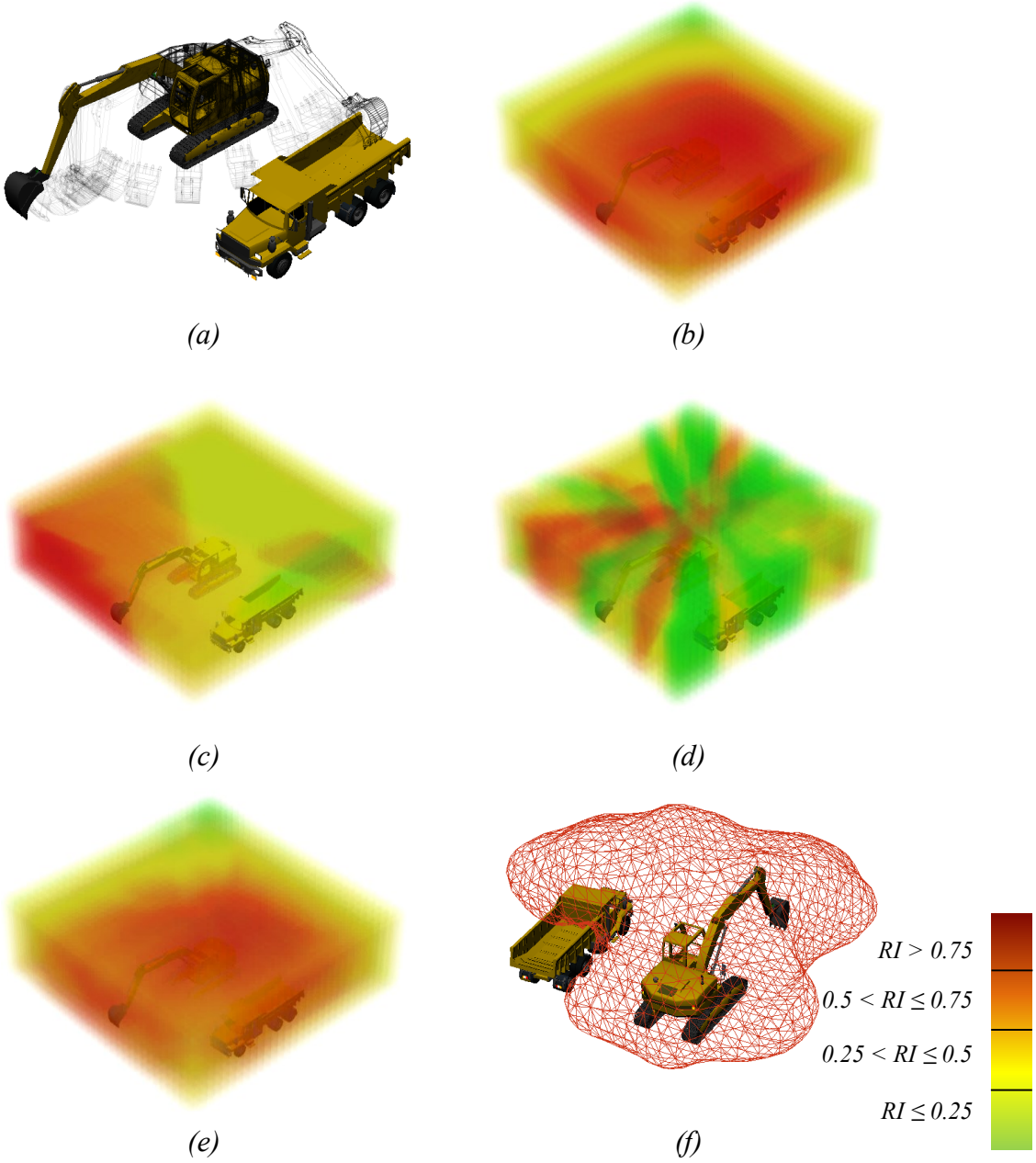


Figure 7-8: Schematic Representation of (a) Portion of Swinging and the Corresponding Indices in S for (b) $SDE_{i,q}$, (c) $TSD_{i,q}$, (d) $VI_{i,q}$, (e) Combined Risk Map of the Excavator, and (f) $LAEW_{truck}$

Figure 7-9 shows the flowchart for the determination of VI_i. The process starts from time T_{now} and from the first cell. Ray tracing is used to check if the cell is visible from the Point of View

(*POV*) of the operator considering the terrain conditions and possible obstacles that might lie between the cell and the *POV* of the operator, including the components of the same equipment. Upon the completion of this step, if the cell is found to be visible to the operator, VI_i is incremented by one unit of time (*dt*). This process is then repeated for all cells before the time is advanced for one step (*dt*), until the calculation is completed for δt. The final outcome of this process is the values of VI_i for every cell in the analysis space.

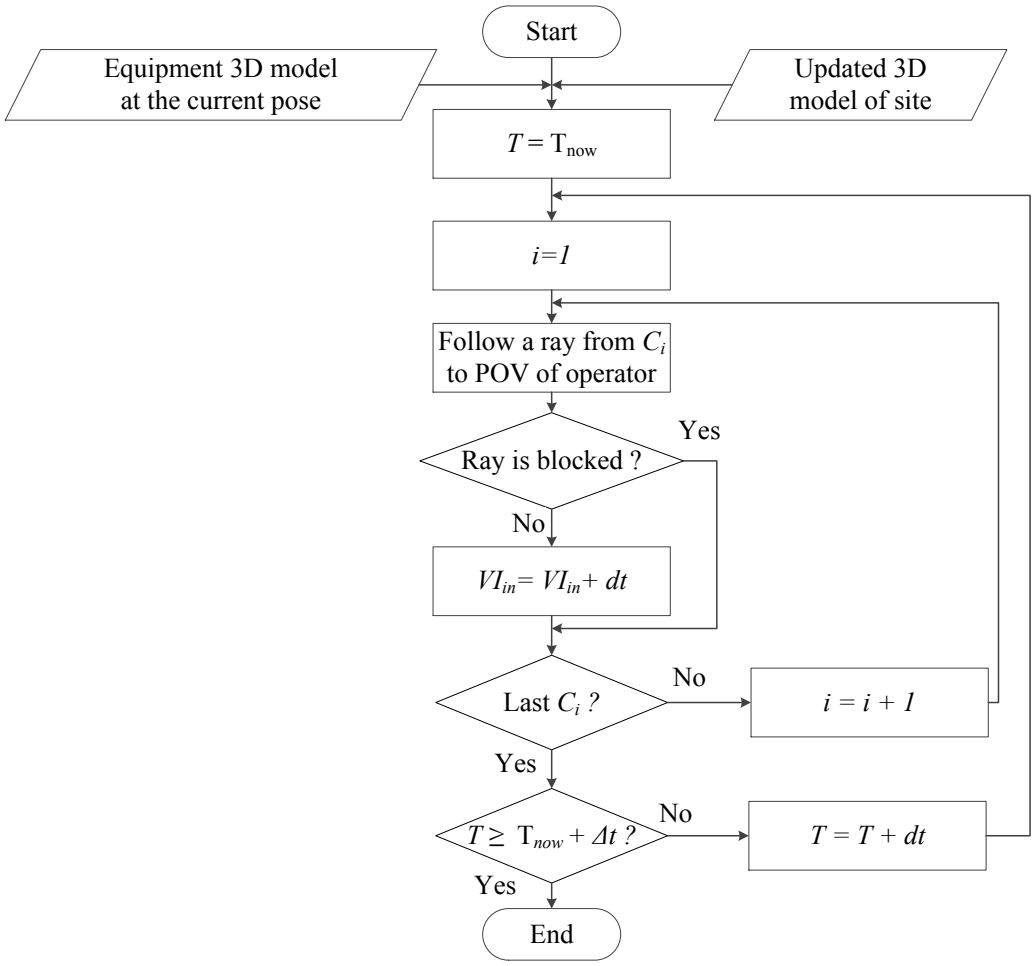


Figure 7-9: Flowchart for the Calculation of Visibility Index

Figure 7-8(d) illustrates an example of VI_i calculation. It is noteworthy that other climatic factors that may impact the visibility, e.g., fog, can be also taken into account in the calculation of VI_i.

However, these factors affect the space in a uniform manner, meaning that their impact reverberate equally into all cells. Thus, their impact can be added to the calculated VI_i as a post-processing step.

7.2.1.4 GENERATION OF EQUIPMENT RISK MAP

Once all the indices are calculated, they need to be normalized so that the overall risk index associated with each cell can be measured in a range between zero and one; the value of zero representing a risk-free cell and the value of one representing a cell that is already colliding with the equipment and that remains invisible for the entire period of δt, for example the cell that contains a part of the excavator superstructure during a swinging operation. Table 7-1 summarizes the three indices and their dimensions, initial ranges, and normalization method. In order to normalize SDE_i, it is divided by the largest value of SDE_i within δt and then the result is subtracted from 1, i.e., inverted so that the value of 1 represents the highest risk. Also to normalize TSD_i and VI_i, their initially calculated values are divided by δt and the result is subtracted from 1.

Table 7-1: List of Different Indices and Their Normalization Method

Index	Dimension	Range	Normalization	Risk Index
SDE	Distance	[0~SDE_{max}]	$1 - \frac{SDE_i}{SDE_{max}}$	[0~1] (low risk ~ high risk)
TSD	Time	[0~Δt]	$1 - \frac{TSD_i}{\Delta t}$	[0~1] (low risk ~ high risk)
VI	Time	[0~ Δt]	$1 - \frac{VI_i}{\Delta t}$	[0~1] (low risk ~ high risk)

Equation 7-3 represents the method to combine all the risk indices into one value considering different weights (w_1, w_2, w_3) for these indices. Although the weights are defined according to the risk attitude of the safety manager, it is recommended that SDE_i and VI_i are given the highest and lowest weights, respectively. This is due to the fact that while the consideration of SDE_i is

coming from the expected movements of the equipment, VI_i is considered for the cases where the equipment may deviate from its expected movement and thus might pose a danger to crews and equipment in its blind spot.

$$RI_{i,q} = 1 - \frac{w_1 \times \frac{SDE_{i,q}}{SDE_{q,max}} + w_2 \times \frac{TSD_{i,q}}{\delta t} + w_3 \times \frac{VI_{i,q}}{\delta t}}{w_1 + w_2 + w_3} \quad \textbf{Equation 7-3}$$

Where:

$RI_{i,q}$*: Overall risk index for cell* C_i of the risk map of equipment q

w_1, w_2, w_3*: Weights associated with* $SDE_{i,q}$, $TSD_{i,q}$ *and* $VI_{i,q}$, *respectively*

$SDE_{q,max}$*: Maximum* $SDE_{i,q}$ *in* S_q *over* δt

δt*: Duration for the LAEW analysis*

The outcome of this step is the risk map of the equipment over the period of δt. Figure 7-8(e) shows an example of the excavator risk map based on the weight distribution of (0.8, 0.1, 0.1).

7.2.2 GENERATION AND APPLICATION OF LAEW$_P$

Once the risk maps of all pieces of equipment are generated by the *OAs*, the *TCA* uses them to generate the $LAEW_p$ for equipment p for a specified risk level. As described in Section 7.2, it is important to re-emphasize that the risk maps are generated by the *OAs* of each piece of equipment (e.g. equipment q), and the *TCA* considers all the equipment that impact equipment p to generate $LAEW_p$.

A *TCA* does not generate $LAEW_p$ solely based on the team of the equipment it is coordinating, given that there might be situations where a piece of equipment from one team is working in proximity to another team. Therefore, the *TCA* considers not only their subordinate *OAs* for the generation of $LAEW_p$, but also they communicate with other *TCA*s to identify *OA*s of other teams

which are working in their proximity to incorporate their risk maps into $LAEW_p$. Nevertheless, the communication scheme between *TCA*s is out of the scope of the present chapter and will be discussed in the future work.

The next step in the integration of the $LAEW_p$ is the readjustment of values of risk indices of multiple risk maps in view of the nature of interactions between various pieces of equipment. This is important because a piece of equipment does not pose the same level of risk to all other pieces of equipment. For instance, a truck that is being served by an excavator is more tolerant to the risks posed by that excavator compared to the risks posed by a loader from another team of equipment, given that the interaction and proximity between the truck and excavator is expected as part of the nature of operation they are performing. To consider this issue, the values of risk indices of the risk map of equipment q that have to be considered to generate the $LAEW_p$ will be adjusted using $w_{q,p}$, which is a weight representing the level of significance of equipment q for equipment p, as shown in Equation 7-4. The smaller the value of $w_{q,p}$, the more tolerant is equipment p to the risks posed by equipment q.

$$ARI_{i,q,p} = w_{q,p} \times RI_{i,q}$$ **Equation 7-4**

Where:

$ARI_{i,q,p}$*: Adjusted overall risk index for cell* C_i *of the risk map of equipment* q for the *LAEW of equipment p*

$w_{q,p}$*: The weight associated with the level of significance of equipment q for equipment p*

$RI_{i,q}$*: Overall risk index for cell* C_i of the risk map of equipment q

After the determination of *Adjusted overall Risk Indexes* (*ARI*) for all the equipment that affect equipment p, they are combined to generate $LAEW_p$. $LAEW_p$ is generated by considering the

$ARI_{i,q,p}$ for the cells that belong only to one *S,* and the maximum of $ARI_{i,q,p}$ for the cells that belong to multiple *S*. Given that at the current phase of the research the initial path of the equipment is forecasted based on parametric motion planning in a deterministic manner, if a piece of equipment makes a decision based on the maximum overall risk indices of a space shared by multiple risk maps, it has already adapted itself to the worst case scenario. However, in future research, the initial path of the equipment used in calculating the risk will consider the probability of various potential paths and a more sophisticated method can be applied for a shared space to aggregate various risks posed by different pieces of equipment with different levels of probabilities.

Finally, $LAEW_p$ is generated as the envelope of all the cells with the $ARI_{i,q,p}$ equal to or greater than a specific risk level. Figure 7-8(f) shows the $LEAW_{truck}$ generated from the risk map of the excavator shown in Figure 7-8(e) with the risk level of 0.8. The $LAEW_{truck}$ can be used by the truck to analyze its initial future path in terms of whether or not it trespasses a high-risk spaces. The decision of which equipment needs to apply the path re-planning can be made based on the priorities of the operations of different pieces of equipment. However, once one piece of equipment performed path re-planning, steps explained in Sections 7.2.1.3 and 7.2.1.4 have to be repeated for that equipment and then the *TCA* should generate new *LAEWs* for all the pieces of equipment that have lower priority than that equipment.

The generation of *LAEW* from the risk map allows different equipment and crews to define the tolerable risk level based on their vulnerabilities. For example, while a worker on foot may be specified to avoid $LAEW_{worker}$ with a risk level of 0.5, a truck might be designed to avoid $LAEW_{truck}$ with a risk level of 0.8.

With regards to the application for path re-planning, Figure 7-10 shows how the *LAEW*s are used by different *OA*s within a period of δt for the purpose of path re-planning. As shown in this figure, every *OA* controls if the equipment is on the planned path. If the equipment path is found to have deviated from the planned path, the risk map of the equipment is generated and sent to the *TCA*. Then, the *TCA* generates the *LAEW* and sends it back to the *OA*. Next, if required, the *OA* uses the *LAEW* to re-plan it path.

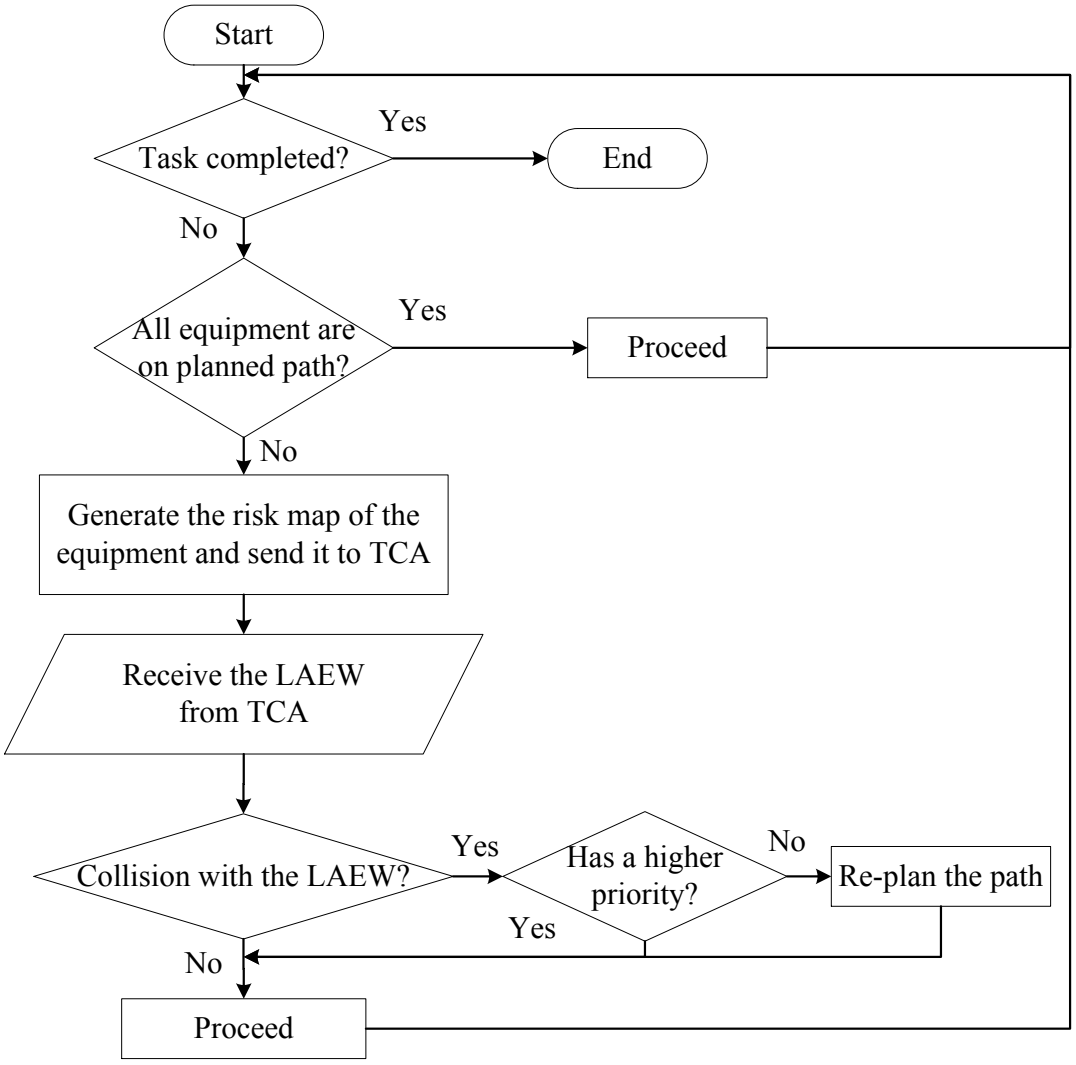


Figure 7-10: Algorithm Representing the Safety Monitoring Exercised by the OAs Using LAEWs

7.3IMPLEMENTATION AND CASE STUDY

7.3.1 IMPLEMENTATION

A case study was conducted to verify and validate the proposed method for generating *LAEWs*. The data from the lab test presented in Section 4.4.2 is used to demonstrate the generation of risk maps and *LAEWs* and their capacity to be used for equipment path re-planning. It is assumed that the generation of *LAEWs* is happening 2.5 minutes after the beginning of the test. The state and pose identification methods proposed in Chapters 4 and 5 have been deployed to provide the required input data. The pose correction method of *UWB* results in an update rate of is 1 Hz.

In order to demonstrate the applicability of *LAEW*s, a prototype system was developed as shown in Figure 7-11, where Autodesk Softimage (AutoDesk 2014) was used as the virtual environment. It should be noted that the current state of the implementation focuses on generating the individual risk maps and combining them to generate the *LAEW*s. The 3D models of the equipment similar to those used in the case study were imported from Google's 3D warehouse (Google 3D Warehouse 2015). The required modifications, including the placement of the operator's point of view in the cabin, which is used for the calculation of *VI*, and defining the tight-fitting bounding boxes, were done manually inside Softimage. The pose of the equipment was imported from the corrected *UWB* data using the pose estimation method, and the results of *NRTS* were used to extract the temporal values of the upcoming states within δt. The value of δt was set to 5.5 s, which is a reasonable timeframe, given the scale of equipment and their corresponding speeds.

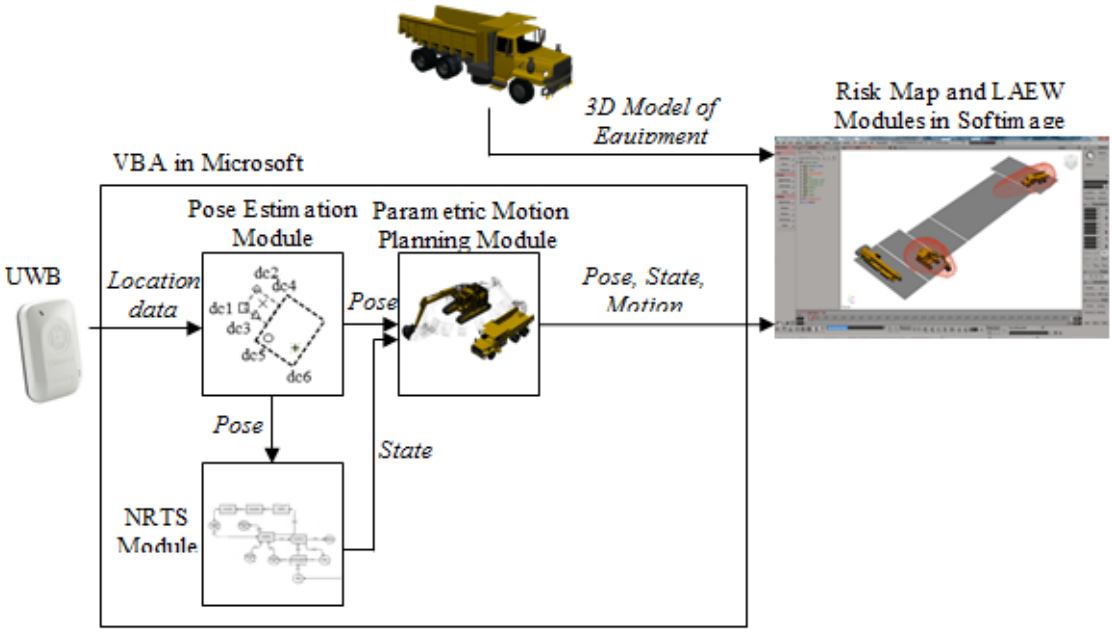


Figure 7-11: Architecture of the Implementation

A module is developed inside Softimage that enables creating the simulated operations based on the input for defining: (1) the current poses of the equipment, (2) the durations of different future states coming from *NRTS*, and (3) the values of $\theta_{j,k}$, $a_{j,k}$ and $c_{j,k}$ used for determining the motion path of different equipment. The visual programming embedded in Softimage is used to develop the tool for generating the global grid of the site and calculating the values of *SDE*, *TSD*, and *VI*. The movements of different parts of the truck and excavator in different states were modeled based on the operation of skilled operators as explained in Section 7.2.1.1. The initial paths generated for the truck and excavator based on the proposed method are shown in Figure 7-12.

Once the different pieces of equipment are situated in the virtual environment and their respective movements over the next δt are simulated, the cells that fall within the analysis space of each equipment are determined based on its movement, as explained in Section 7.2.1.2. Then,

while the operations of the equipment are simulated, the distances to the equipment and the visibility of each cell are calculated and recorded along with the time of the simulation. Upon the completion of the simulation, the aggregated values of the abovementioned parameters are used to calculate SDE_i, TSD_i, and VI_i for the different pieces of the equipment, as explained in Section 7.2.1.3. These values are then normalized and combined through Equation 7-3, using the user-defined weights, to generate the risk map of each piece of equipment. Finally, the risk levels assigned by the user are used to generate the *LAEW*s of each piece of equipment.

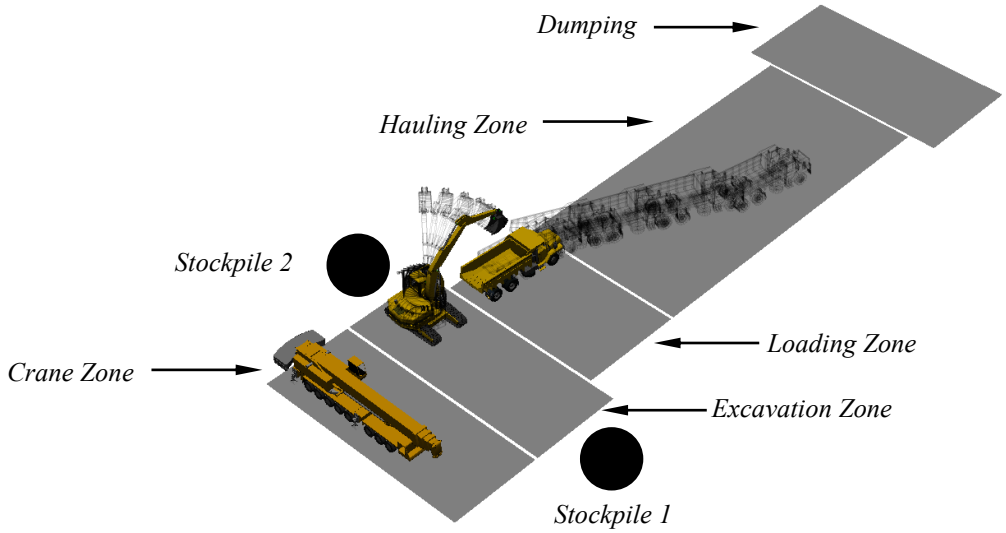


Figure 7-12: The Initial Paths Generated for the Truck and Excavator

As shown in Figure 7-11, at the current stage of the implementation, the pose estimation, *NRTS* and parametric motion planning modules are developed outside Softimage. Together, these modules collect raw location data, generate pose and state data, generate the motion parameters, and present these results to Softimage. The modules responsible for the generation of risk maps (explained in Section 7.2.1) and *LAEW*s (explained in Section 7.2.2) are completely developed inside Softimage. In order to reduce the complexity of the overall system, the pose estimation and *NRTS* modules can be developed in the future inside Softimage, using the software's *API*. In

this way, the entire system shown in Figure 7-1 can be developed in Softimage, and only the results of the location data need to be imported.

7.3.2 CASE STUDY

Figure 7-13 shows the risk maps of different pieces of equipment, where only the cells with risk indices greater than zero are shown. The range of color from green to dark red shows the range of the risk index from zero to one. To show the impact of changes in the weights shown in Equation 7-3 four different distribution of weights for the parameters are considered, namely (0.8, 0.1, 0.1), (0.1, 0.8, 0.1), (0.1, 0.1, 0.8), (0.33, 0.33, 0.33) for *SDE*, *TSD*, and *VI*, respectively. The first three distributions are selected so that one parameter has the dominant weight (i.e., 0.8) and the other two parameters share the remaining weights equally (i.e., 0.1 each). This pattern enables the analysis of the effect of the dominant parameter as well as the pairwise comparison of risk maps where one parameter is fixed and the other two parameters are changed. Figure 7-13(a) to (d) show the risk maps of the excavator for the different weights distributions. Figures 13(e) to (h) show the risk maps for the truck, and Figure 7-13(i) to (l) present the risk maps when the truck and excavator are considered together. It can be observed from the risk maps with the distribution of (0.8, 0.1, 0.1) in Figure 7-13(a), (e), and (i) that the cells close to the motion path of the equipment have higher risk indices than those farther. However, the comparison of Figure 7-13(e) and (f) suggests that with the decrease in the weight of *SDE* and increase in the weight of *TSD*, the risk indices of cells far from the current pose of the equipment decreases. The scrutiny of Figure 7-13(b) suggests that the high weight for *TSD* results in high risk indices for cells with the relatively high *SDE* and low *TSD*. Nevertheless, this is not desirable since the risk of short distances at later time is underestimated. The relative

importance of spatial proximity against the temporal proximity should be used to determine the ratio of the weight of *SDE* to the weight of *TSD*.

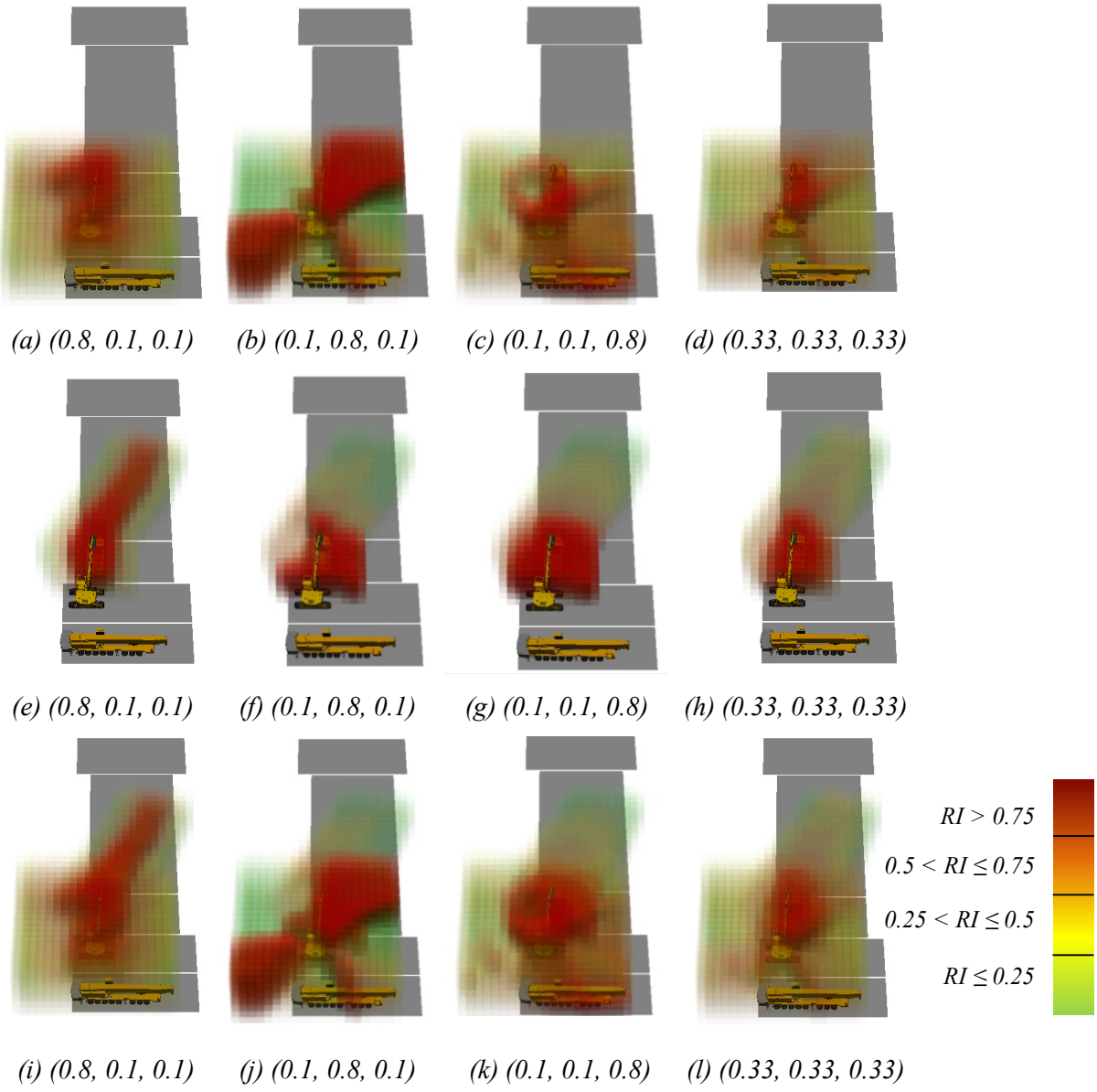


Figure 7-13: Risk Maps for (a) to (d) Excavator, (e) to (h) Truck, and (i) to (l) Combination of the Excavator and Truck for Different Weight Distribution

Additionally, based on the comparison of Figure 7-13(a) and (b), where the sum of the weights of *SDE* and *TSD* is 0.9, with Figure 7-13(c), where the sum of the weights of *SDE* and *TSD* is 0.2, it can be inferred that the lower the sum of the weights of *SDE* and *TSD*, the less sensitive

the cells to the proximity-based risks. However, since Figure 7-13(c) attributes greater risks to cells with low visibility but far from the equipment than cells close to the equipment but with high visibility, it is not effective. Therefore, the proximity-based risks should be given the higher importance over the visibility-based risks. This importance can be reflected in the ratio of the summation of the weights of *SDE* and *TSD* to the weight of *VI*. As stated earlier in Section 7.2.1.4, given the significance of spatial proximity, it is recommended that *SDE* and *VI* are given the highest and the lowest priority, respectively. Through several trials and errors, it is found that the ranges of (0.6~0.8), (0.1~0.2), and (0.1~0.2) for *SDE*, *TSD*, and *VI,* respectively, can effectively capture the risks induced by the future motions of the equipment to the surrounding environment.

The inputs from Figure 7-13 are used to generate the *LAEWs* corresponding to different risk levels. Figure 7-14 shows $LAEW_{truck}$, $LAEW_{excavator}$, $LAEW_{crane}$ for the weights distribution of (0.8, 0.1, 0.1). Figure 7-14(a) and (b) show $LAEW_{truck}$ corresponding to the risk level of 0.9 and 0.8, respectively. Similar results are presented for the $LAEW_{excavator}$ and $LAEW_{crane}$ in Figure 7-14(c) to (f). It can be seen that as the risk level decreases, the *LAEW* increases in size. As stated in Section 7.2.1.4, the $LAEW_{truck}$ can be used by the truck to check potential collisions with its initial path, and to perform path re-planning if necessary. On the other hand, the crane should use $LAEW_{crane}$ to plan a safe path that will not cause a collision with either the truck or the excavator. Figure 7-15 helps put the application of the generated *LAEWs* and the necessity of $ARI_{i,q,p}$ in perspective. Figure 7-15(a) and (b) show the side views of the workspaces of the truck $(LAEW_{truck})$ with the risk level of 0.9 and 0.8, respectively. As can be seen in Figure 7-15, at a high risk level, part of the excavator bucket is falling outside the $LAEW_{truck}$. These workspaces indicate that given the close nature of the collaboration between excavator and truck, it is

required to use a higher risk level for the part of $LAEW_{truck}$ coming from that excavator to avoid false alerts. Since the operation of the excavator can be underground, the analysis space for the excavator is intentionally extended to consider the areas that may fall under the ground. This arrangement can extend the application of *LAEW* for collision detection with underground utilities, as can be seen in Figure 7-15.

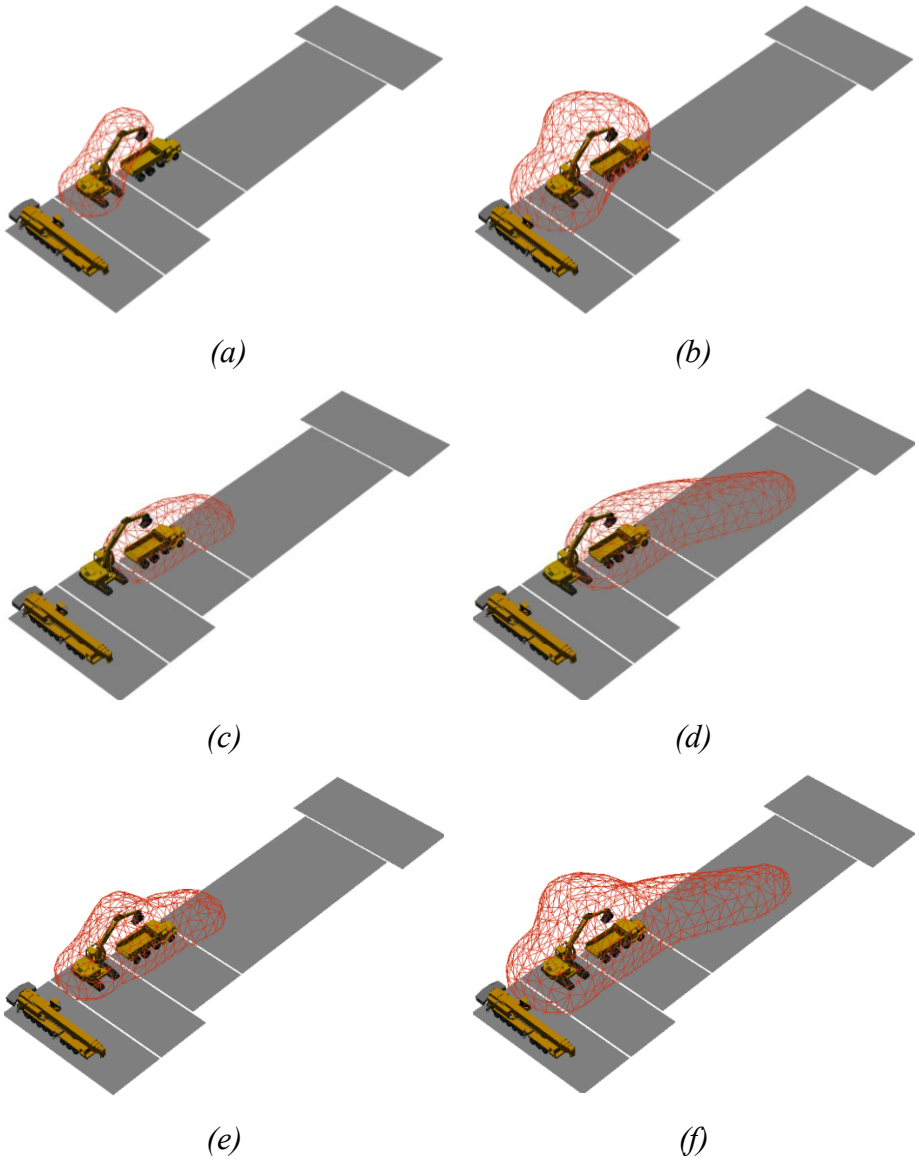
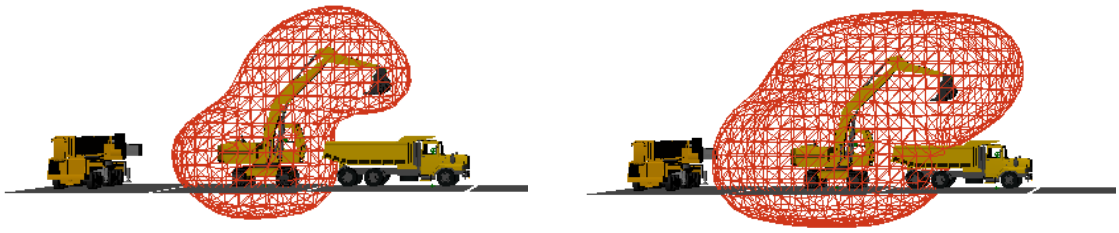


Figure 7-14: Considering the Weight Distribution of (0.8, 0.1, 0.1), (a) $LAEW_{truck}$ with Risk Level of 0.9, (b) $LAEW_{truck}$ with Risk Level of 0.8, (c) $LAEW_{excavator}$ with Risk Level of 0.9, (d) $LAEW_{excavator}$ with Risk Level of 0.8, (e) $LAEW_{crane}$ with Risk Level of 0.9, (f) $LAEW_{crane}$ with Risk Level of 0.8



(a) *(b)*

Figure 7-15: Side Views of $LAEW_{truck}$ with (a) Risk Level of 0.9, and (b) Risk Level of 0.8

Another interesting observation from this case study comes from the sensitivity analysis of the computation time for the generation of *LAEWs* with respect to cell size and δt in the developed implementation, which was running on a personal computer with Intel Core i7-2600 CPU (3.40 GHz). A range of seven cell sizes (3 *cm*, 5 *cm*, 10 *cm*, 15 *cm*, 20 *cm*, 25 *cm*, 30 *cm*), corresponding to a range of the number of cells (930,852, 196,000, 24,500, 7,896, 3,150, 1,568, 1,152) were considered. Similarly, δt was investigated in a range from 1.5 s to 5.5 s with the interval of one second. The computation time is presented in time per simulated second. Figure 7-16 shows that with the increase in the cell size from 3 cm to 10 cm, the computation time per simulated second decreases drastically (almost 82%). However, after the cell size of 10 cm (corresponding to 24,500 cells) the reduction in computation time becomes less significant. The calculation takes an average of 0.5 s (depending on the value of δt) for the cell size equal to 10 cm. This time is reasonable since the ratio of the calculation time to δt is approximately 16%, meaning that the calculation takes a small portion of the look-ahead window (i.e., δt) and therefore leaves enough time for further path re-planting and the execution. This ratio increases to 95% for the cell size equal to 5 cm, which basically means that the entire look-ahead window was used for the generation of *LAEW*s. Accordingly, it can be inferred that for the small look-ahead time to work efficiently, the cell size has to be increased.

The identification of the *SDE* and the ray tracing used for the calculation of *VI* are the first and second most time-consuming calculations. In order to put the intensity of the computation of these two parameters into perspective, the calculation of *SDE and VI* accounted for 14.5% and 14.3% of the total calculation time for the generation of *LAEW*s of the truck when the cell size is 10 cm and *δt is* 5.5 s.

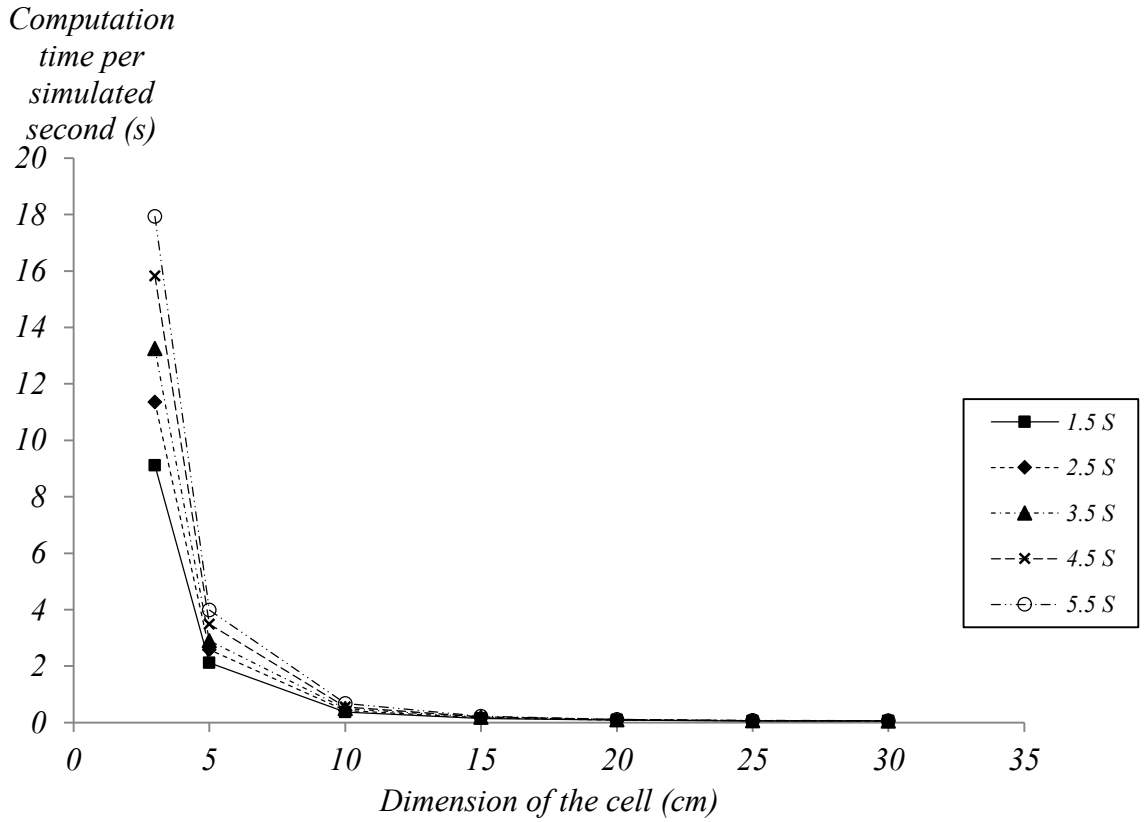


Figure 7-16: Analysis of the Computation Time based on the Cell Dimension and Δt

7.4 DISCUSSION

Given that the prediction of future movements of different pieces of equipment obtained through the integration of *NRTS* and parametric motion planning is an integral part of the proposed method, it should be highlighted that uncertainties are an indivisible part of these predictions. Consequently, the proposed method does no claim that it can eliminate all types of proximity-

based and visibility-based risks. Instead, it is committed to leveraging the most reliable data that are available about the imminent movements of equipment during a manageable time window to mitigate the risks as much as possible. Paths generated using the *LAEW*s are therefore the result of best efforts to use all the available information to avoid risks. However, the reliability of the *LAEW*s is directly correlated to the length of δt, since the shorter the look-ahead window, the fresher and more accurate the prediction of the motion planning and *NRTS*.

While one of the main merits of the proposed method of integrating the risk parameters is that it grants flexibility to the designer in terms of choosing appropriate risk weights, it has the limitation of relying on the shortest distance of cell to the equipment and the time associated with it for the calculation the proximity-based portion of the risk. This can result in potential negligence of another slightly larger distance that might have taken place earlier. Nevertheless, the magnitude of this limitation can be reduced by increasing the relative importance of *SDE* against *TSD*. By doing so, for the cell to have a higher risk at a time earlier than *TSD*, it is required to have a much closer distance to the shortest distance.

Moreover, the ratio of the calculation time to δt can be used to highlight that not the entire period of δt can be used to effectively avoid the collision risks. Accordingly, the *LAEW*s are most effective when coupled with the *DEW*s. Together, they act as a two-layer safety shield where the upper layer intends to ensure that the future paths of the equipment are collision-free and the lower layer aims to prevent immediate collisions if the generated paths fail to act as expected in face of uncertainties. *DEW*s can be used in the calculation portion of δt to avoid collisions.

Also, it is noteworthy that two levels of path planning were discussed in this chapter. The first level of initial path planning is a part of the proposed method and was explained in Section 7.2.1.1. On the other hand, the second level of path re-planning uses the output of the

proposed method to generate collision-free paths if necessary. The correlation between the two levels of path planning is that while the first level is devoid of full awareness of the surrounding environment, the second level uses the risk assessment represented by the *LAEW*s and a set of priority rules to identify the equipment that require path re-planning, and adjust the initial path, if required.

7.5 SUMMARY AND CONCLUSIONS

This chapter proposed a novel method for look-ahead equipment workspace for earthwork equipment that uses the predictive power of *NRTS* to evaluate the site safety based a number of parameters including *SDE*, *TSD*, and *VI*. This method enables different pieces of equipment to ensure that their initially planned paths are collision-free, or alternatively adjust their path planning to avoid potential collisions. The discretization of space is applied to decompose the earthwork site into a number of cells in a virtual environment. Next, *NRTS* and parametric motion planning are used to predict the future motions of different pieces of equipment over the period of δt. Using this information, the risk index of each cell is evaluated in terms of *SDE*, *TSD*, and *VI*. The generated risk map is then leveraged to determine the *LAEW*s of the equipment corresponding to a given risk level. A case study was conducted to validate the proposed method.

In light of the results of the case study, it can be concluded that: (1) the proposed method is providing a reliable basis for the generation of the risk maps of earthwork equipment, using the expected pose and state and considering the proximity-based and visibility-based risks; and (2) the risk maps can be combined to generate *LAEW*s with different risk levels that can be used by different equipment and crews based on the varying levels of risk they can tolerate to adjust their initial paths.

CHAPTER 8: FLEET-LEVEL AUTOMATED EQUIPMENT GUIDANCE IN MAS

8.1 INTRODUCTION

Chapter 3 presented the overview of the proposed *MAS*. To recap on the proposed *MAS*, this research proposed a multi-layer agent architecture in which multiple layers of agents provide support for the operators of earthwork equipment. The structure of the proposed *MAS* is shown in Figure 3-2. Additionally, Chapters 4 to 7 have introduced some of the main functionalities performed by the *MAS* in earthwork operations. This chapter provides an in-depth discussion of the interaction and communication scheme between multiple layers of agents and how different components of the MAS are integrated in a coherent system. Also, the task and operation management and safety management mechanisms in the proposed MAS are elaborated.

8.2 DESCRIPTION OF AGENTS IN THE PROPOSED MAS

8.2.1 OPERATOR AGENTS

Figure 8-1 shows the architecture of an *OA* in terms of the contents of its context-awareness and functionalities. The *OA* requires information about its surroundings, task and environment. This combined types of information is referred to as external information because they are provided by external sources. Surroundings information contains the poses, states, *DEW*s, and safety warning of other pieces of equipment. This information can be directly used by the *OA* to identify safety threats and take immediate actions, if required. Task information allows the *OA* to perform its (semi-) autonomous operations and contains safety warnings, *LAEW*s, a strategic plan, and *NRTS*-based schedule provided by the *TCA* and 3D design made available by the *DDA*.

The functionalities of an *OA* include: (a) Identifying the pose and state of the equipment; (b) Generating the tactical plan of the equipment (i.e., path (re-) planning); (c) Generating the *DEW*s to ensure the safety of the equipment and workers by stopping or re-routing the equipment if required or communicating with other *OA*s; (d) Ensuring the safety of the equipment and workers by stopping or re-routing the equipment if required or communicating with other *OA*s; (e) Requiring the equipment to execute certain tasks based on the task assignments by *TCA*s and the generated tactical plan; (f) Calculating the required parameters for the generation of the equipment risk map; (g) Collecting data about the cycle time of the machine and reporting to the respective *TCA*; (h) Monitoring the equipment conditions; (i) Detecting underground utilities based on the provided maps; and (j) Updating the *DTM* upon any changes to the topography of the site.

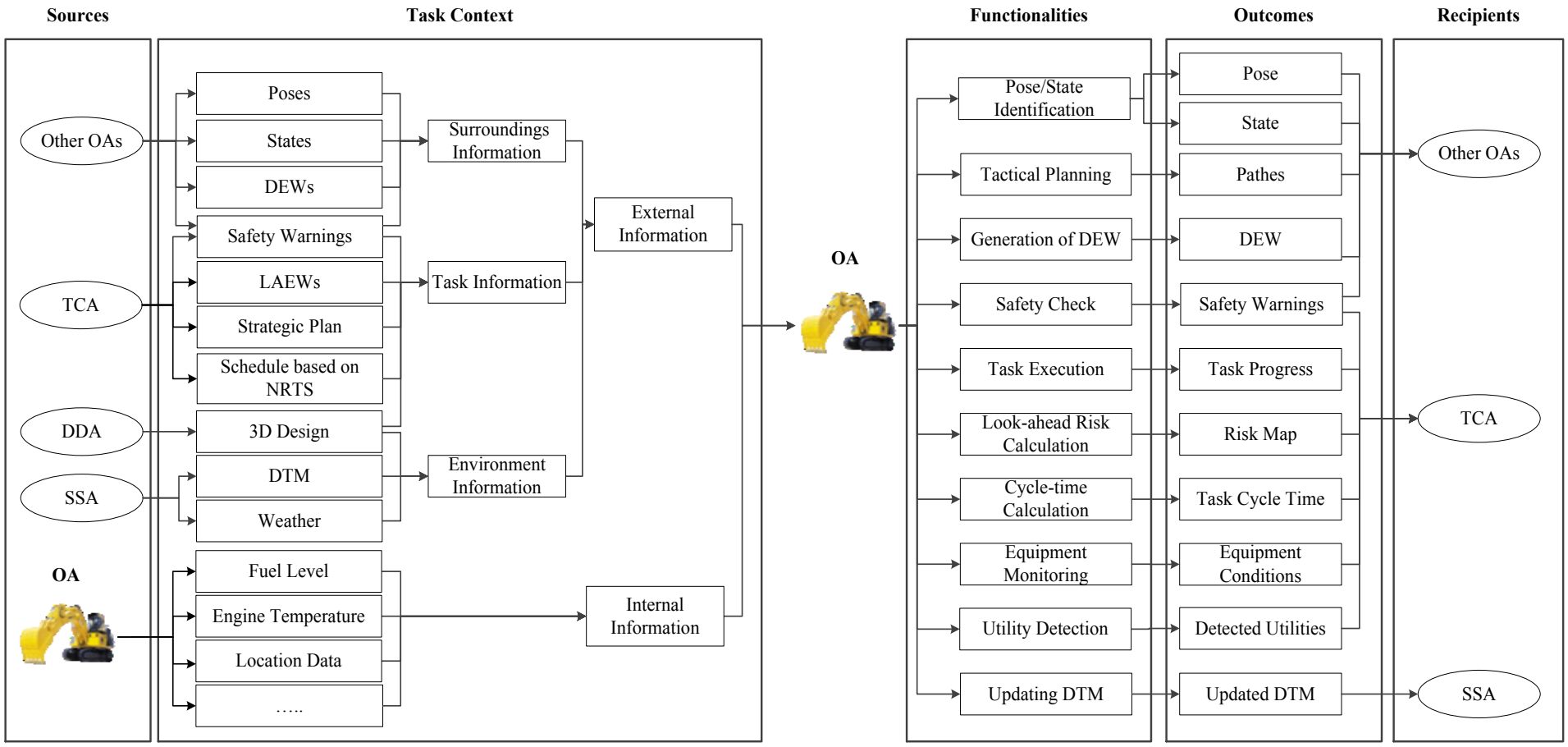


Figure 8-1: Architecture of OA

As for the pose and state identification, knowing that all pieces of equipment are equipped with *RTLS* and various other types of sensors, *OA*s can use accurate location and sensory data to determine not only their state but also the pose of the equipment. The pose of the equipment refers to the orientation and position of the equipment and represents a particular geometric relationship between various rigid components of the equipment at a certain position and changes as the equipment moves along any of its different *DOF*s. A robust optimization-based method that uses geometric and operational characteristics of the equipment is proposed to improve the quality of the pose estimation, as shown in Chapter 4.

State of the equipment, on the other hand, indicates the type of action a piece of equipment is engaged in, e.g., swinging, loading, dumping etc. A rule-based system is used to identify the states of different equipment with a high accuracy by leveraging a set of equipment proximity and motion rules, as shown in Chapter 5.

*OA*s have knowledge about the equipment specifications and the types of sensors it carries and the data needed for the pose and state identifications. Other components of *OA*s' knowledge pertain to methods required to extract pose and state information and to serve the functionalities specified in Figure 8-1. A high-level flowchart of the functionalities of an *OA* is shown in Figure 8-2. Given the unequal priorities of various functionalities of *OA*, and in order to embed these priorities in the structure, a modified subsumption architecture is chosen for agents. Subsumption architecture is based on breaking the activities of an agent in vertical modules where every module has limited responsibilities and the results of the higher modules always supersede those of the lower modules, if there is a conflict between various modules (Ferber 1999).

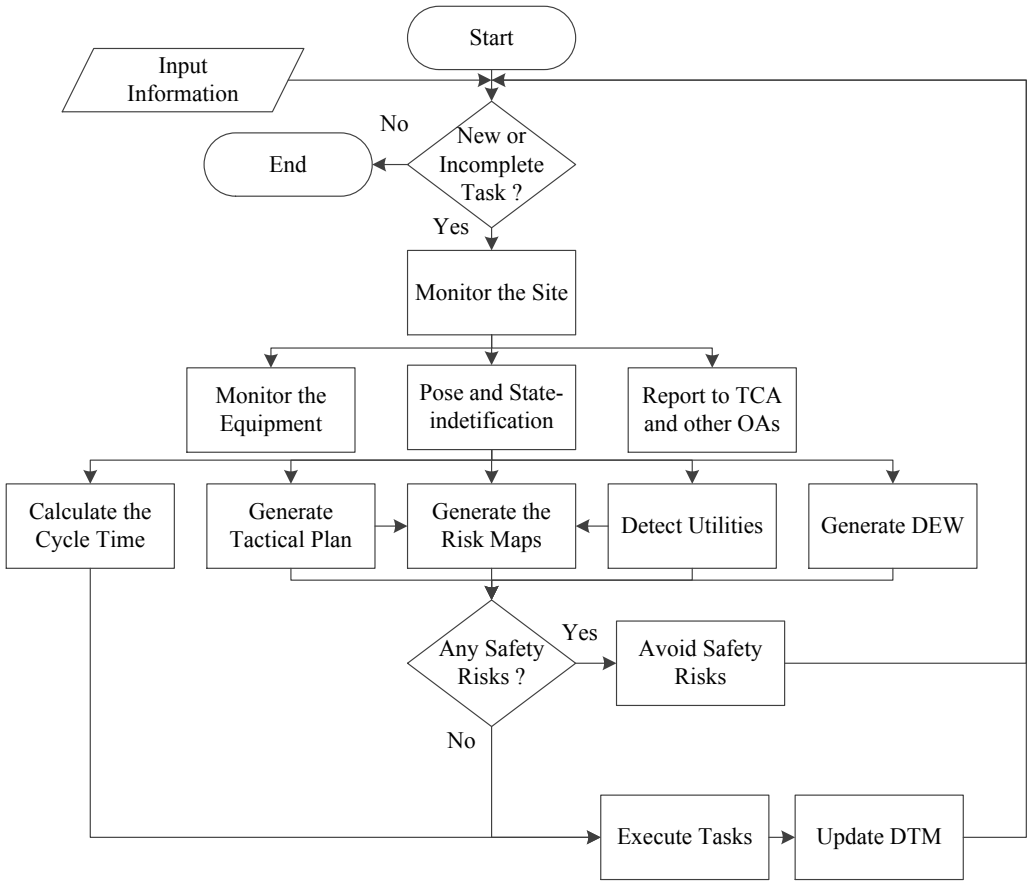


Figure 8-2: High-Level Flowchart of the OA Functionalities

In a nutshell, *OA*s constantly monitor the operations and perform the routine calculations for the equipment condition monitoring, pose and state-identifications, cycle time, generation of tactical plans, generation of risk maps, detecting utilities, and generating *DEW*s. However, this research focuses on the methods for the generation of *DEW*s and risk maps, as explained in Chapters 6 and 7. The development of the methods for the other calculations is part of the future work of this research. *OA*s also constantly report to *TCA* about their progresses. Meanwhile, if any safety risk is observed, immediate actions will be taken and the ongoing operation is halted. The type of the safety risk is communicated with the relevant *TCA* and the commands for the required action are provided to the *OA*s.

8.2.2 COORDINATION AGENTS

Coordination encompasses agents representing team coordinators who are responsible for making critical decisions, e.g., new work schedules or command for the suspension of the operation, using data from all other related agents, and further communicating their decisions with the appropriate *OA*s for the execution. Essentially, this component consists of one *GCA* and several *TCA*s. Each team is coordinated and supported by a *TCA*. However, depending on the characteristics of the project, the phase of the project and simultaneous operations, a hierarchy of several layers of teams and sub-teams can be formed.

8.2.2.1 TEAM COORDINATOR AGENTS

As shown in the architecture of a typical *TCA* in Figure 8-3, the role of a *TCA* is to assign tasks to the subordinate *OA*s or sub-*TCA*s and to collect information from them. The contextual awareness of a *TCA* is established by the information about the status of the surrounding teams, the operation to be accomplished, the environment, and the progress and status of the tasks for which it is responsible. The surroundings' information contains the *LAEW*s and safety warnings from other *TCA*s. This information helps the *TCA* to determine whether or not its own operation will spatio-temporally conflict with the other operations and decide if work rescheduling is required.

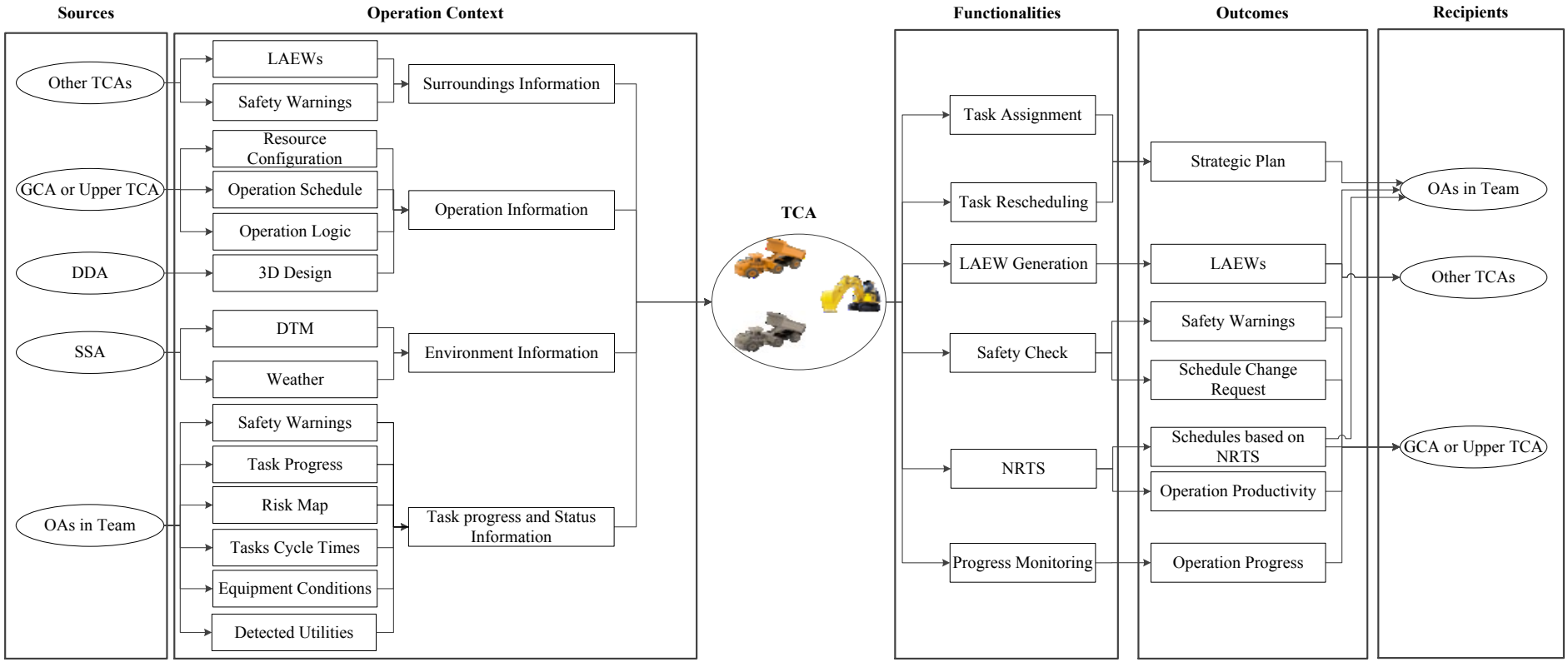


Figure 8-3: Architecture of TCA

In order for the *TCA* to be able to generate an initial work schedule for the subordinate *OA*s, it requires information about the operation to be accomplished. This information, termed as operation information, encompasses: (1) the resource configuration, operation schedule and operation logic from the *GCA* or superior *TCA*, and (2) the 3D design from the *DDA*. The *DTM* and weather information from the *SSA* are categorized as environment information. The operation logic refers to the essential construction method through which the work is done and includes the types of equipment required and the process that determines how different pieces of equipment are interacting, as will be explained in Section 8.2.5.1. The subordinate *OA*s of a *TCA* report about their encountered safety issues, tasks progress, risk maps, cycle-time, equipment conditions, and detected utilities, together termed as task progress and status information. The *TCA* will use this information to perform detailed progress monitoring and *NRTS* to ensure that the operation does not deviate from the project schedule. Accordingly, a *TCA* serves the following ends: (a) Generating the initial strategic plans for all subordinate *OA*s and assigning the corresponding tasks to them; (b) Performing task rescheduling when a delay is observed or a measure is taken to expedite the operation, e.g., a new equipment is added to the fleet or a safety issue dictated a change of schedule; (c) Generating the *LAEW*s for different *OA*s; (d) Ensuring the safety of operation with respect to the equipment in the same team and other teams; (e) Performing *NRTS* to ensure that the project will remain on schedule based on the most updated productivity pattern; and (f) Monitoring the progress of the work based on indices such as the earned value.

The knowledge of a *TCA* comprises methods to perform task assignments, task rescheduling, performing *NRTS* and progress measurements. Similar to *OA*s, a modified subsumption

architecture is chosen for *TCA*. Figure 8-4 shows the high-level flowchart of the *TCA* functionalities.

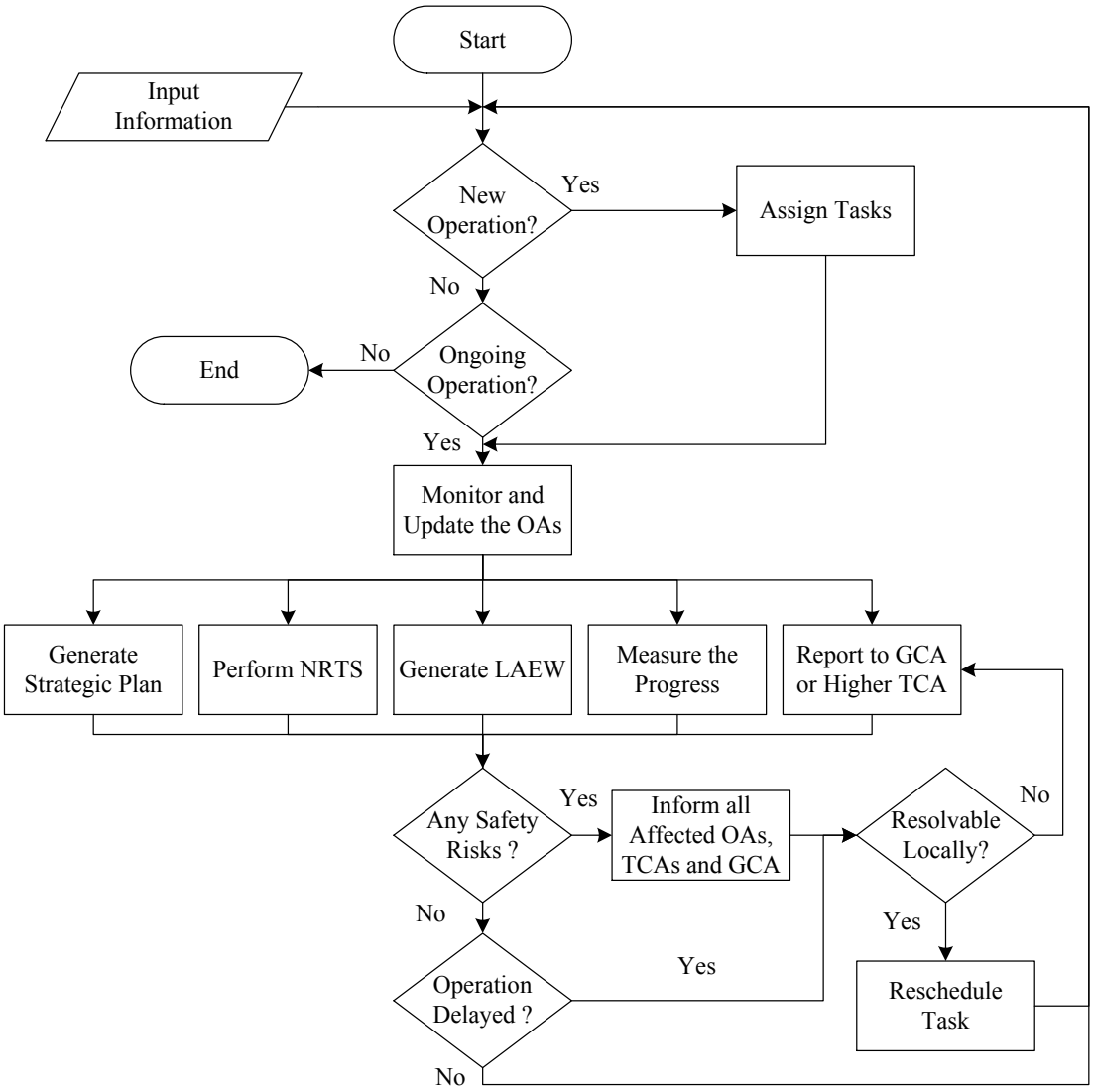


Figure 8-4: High-level Flowchart of the TCA Functionalities

At the top of the flowchart, *TCA* determines whether an operation is ongoing or a new operation is assigned, in which case the operation is broken into *OA*-executable tasks that are assigned to the relevant available *OA*s. Next, in view of the reports from subordinate *OA*s, the progress monitoring, *NRTS*, and *LAEW*, if any risk or delay is identified, either the tasks are rescheduled if the problem can be resolved locally, or the *GCA* (or higher *TCA*) is informed for direction. Local

resolvability means the problem can be solved by the information present to a single *TCA*, without the need to engage into negotiations with other *TCA*s. The negotiation between agents in a decentralized MAS structure is outside the scope of the present research. Of various functions performed by *TCA*, the detailed methods for performing *NRTS* and generating *LAEW* were presented in Chapters 5 and 7.

8.2.2.2 GENERAL COORDINATOR AGENTS

The *GCA* is responsible for monitoring and controlling the operations to ensure the smooth execution of the project. As shown in Figure 8-5, the *GCA* also generates the operations' schedule and the resource distribution based on a set of input data. The functionalities of the *GCA* are realized through the accumulation of information about the project and the progress of different operations. The project information is the combination of all essential documents/information based on which an earthwork project is executed. At a high level of abstraction, safety regulations, available resources, project schedule, construction methods, and available sub-contractors, all of which are coming from the *PDA*, are the main ingredients of the project information.

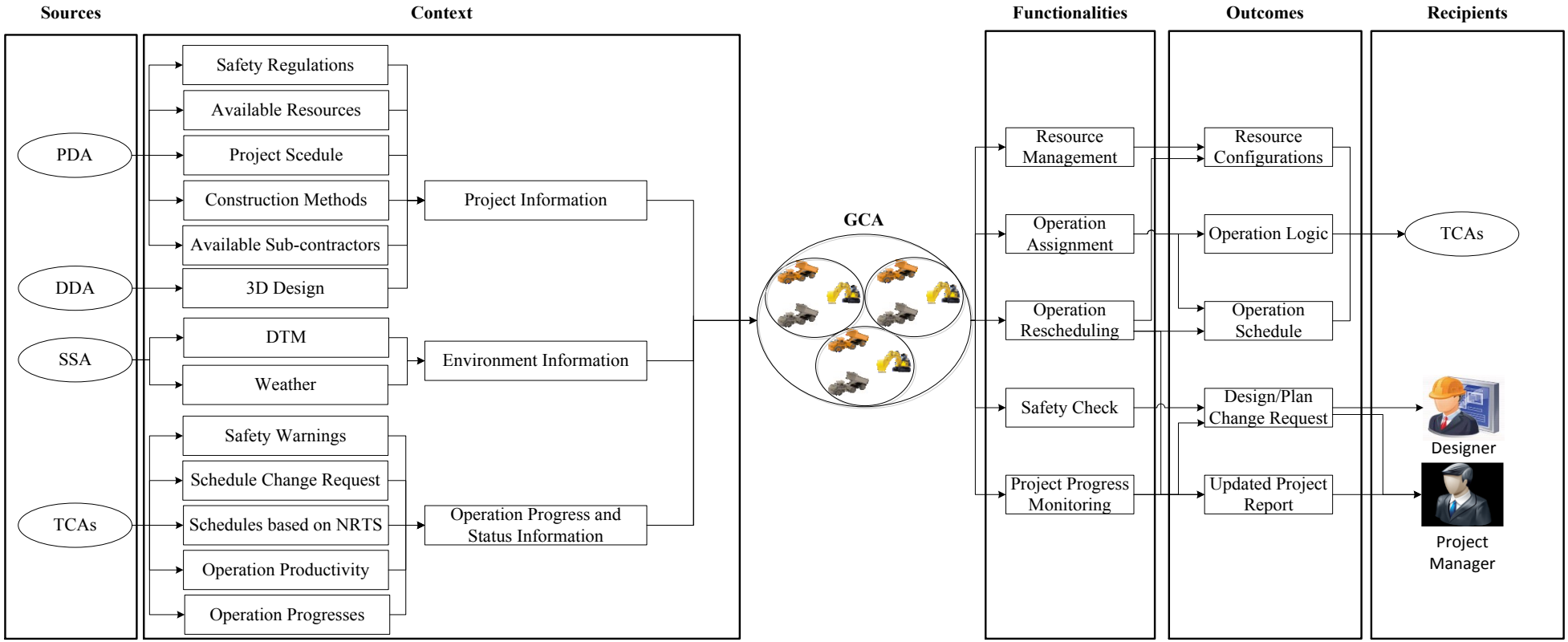


Figure 8-5: Architecture of GCA

Safety regulations are used to derive basic safety rules that need to be observed throughout the project. Available resources and available sub-contractors are used for the resource configurations and distribution. The project schedule is used for the generation of operation schedules that can be assigned to different *TCA*s. The Construction methods provide the *GCA* with the initial information needed to retrieve the right operation procedures. The *DTM* and the 3D design are required for the spatial aspects of operation assignments, and later for the task execution by the *OA*s. With the exception of the 3D design that comes from the *DDA*, all other ingredients of the project information are furnished by the *PDA*. The *DTM* and weather information are provided by the *SSA* as environmental information. In addition to the project information, the *GCA* requires feedback from the *TCA*s in form of operation progress and status information. This includes the *TCA*-provided information about: (a) safety warnings, (b) schedule change request, (c) forecasted schedule based on *NRTS*, (d) operation productivity, and (e) the progress made in different operations.

Example tasks of the *GCA* are: (a) Performing resource management and distributing available resources between various operations; (b) Disintegrating the project schedule into several operations and assigning each operation to a *TCA*; (c) Rescheduling the operations if required in face of delays or safety warnings; (d) Ensuring the safety of the entire project, in terms of avoiding disruption between the work of various *TCA*s; and (e) Constantly monitoring the progress of the project. In order to enable the *TCA*s to generate task plans, the *GCA* transfers the resource configuration, operation logic and operation schedule to them, and provides the design change requests to the designer and plan change requests and progress reports to the project manager.

In order to assign tasks to different *OA*s, the types and interactions of the equipment needed for every operation must be known. The *GCA* has various operation logics associated with different construction methods as part of its knowledge. Also, the *GCA* possesses knowledge about how to perform resource management, operation assignment, operation rescheduling, safety checks, and project progress monitoring. The high-level flowchart of the *GCA* Functionalities is shown in Figure 8-6.

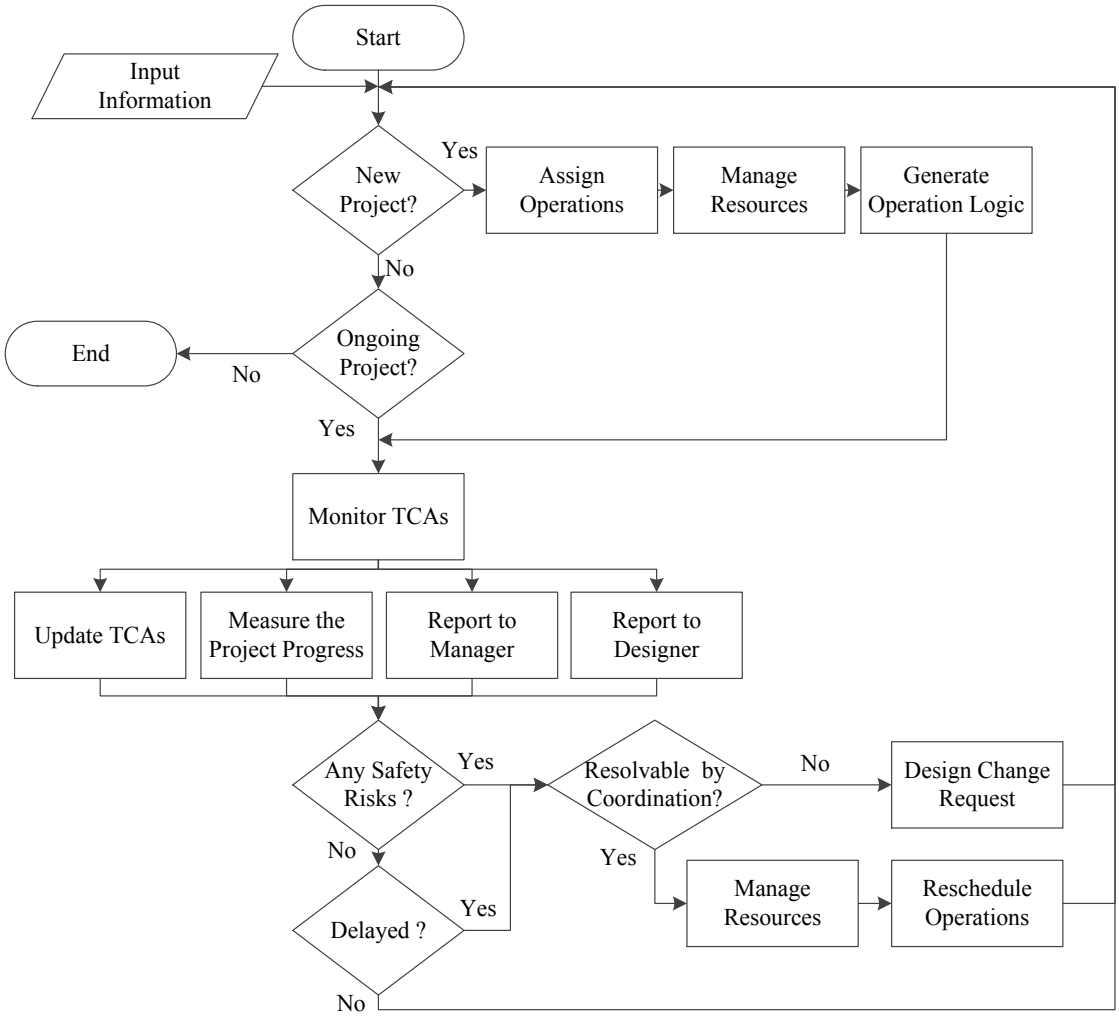


Figure 8-6: High-level Flowchart of the GCA Functionalities

Similar to the *TCA*, the first task of the *GCA* is to identify if there is an ongoing project or if there is a new project, in which case it proceeds with assigning the operations, managing the

resources, and generating the operations' logics. Later, through performing the project progress monitoring, if any safety risks are identified that cannot be locally resolved by a *TCA*, the *GCA* tries to resolve it through coordinating several *TCA*s, re-organizing resources and/or rescheduling operations. However, if the safety threats cannot be alleviated, due to reasons such as detecting unexpected underground utilities, a design change request is sent to the designer. If the project is found to be delayed, then a similar process is followed to investigate whether the problem can be resolved through resource re-configuration and operation rescheduling.

8.2.3 INFORMATION AGENTS

This component is in charge of handling the information required for *LGS* and encompasses the *SSA*, *DDA* and *PDA*. In the scope of this research, information agents are inherently reactive agents in the sense that they do not have any reasoning mechanism. They hold information and react to the requests made by coordinator/operator agents and provide the required information or update their contents based on the most recent changes on the site. Figure 8-7 shows the structure of the *SSA*, *DDA*, and *PDA*. The *SSA* provides the *DTM*, which is often obtained from Light Detection and Ranging (LIDAR) scans by surveyors. Additionally, the *SSA* uses a variety of local sensors coupled with the information from weather agencies to constitute a database of the weather conditions as expected at the planning time, at the current time and as forecasted. The main functionalities of the *SSA* are to provide information to the *TCA*s and *OA*s, when needed, and to update their data whenever a change befalls.

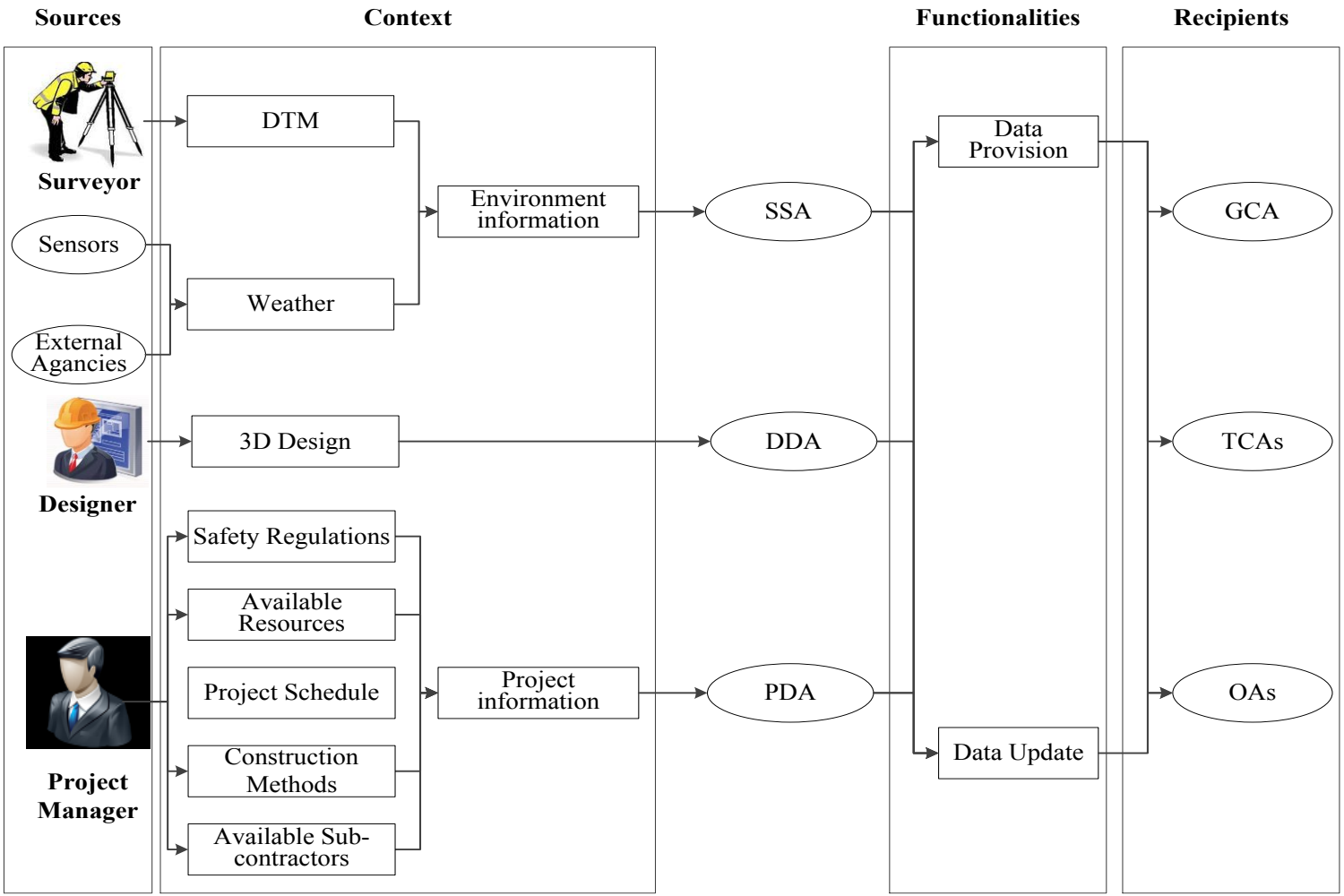


Figure 8-7: Architecture of SSA, DDA, and PDA

The *DDA*, on the other hand, encapsulates the designer-provided 3D models and updates them should any changes are made in the course of the project. The 3D design is used by all the *TCA*s and *OA*s as a reference for decision making and task execution.

Finally, the *PDA* hosts all the basic documents based on which an earthwork project is typically managed, including safety regulations, available resources, project schedule, construction methods, and available sub-contractors. Similar to the *SSA* and *DDA*, the *PDA* is responsible for providing the required information to the *TCA*s and *OA*s and keeping its own information updated.

Given that information agents are reactive in nature, they are devoid of sophisticated procedural knowledge. The process can be simply limited to providing the requested information in time and updating the information based on the recent changes to the site.

8.2.4 DATA STRUCTURE OF THE PROPOSED MAS

The Unified Modeling Language (UML) (Booch et al. 1999) provides the basis to visually outline the structure of an objected-oriented (OO) system and comprises a multitude of diagram types that use *UML* semantics and notations to represent various aspects of a system. With the *MAS* structure following the principles of the object-oriented design paradigm, *UML* can be used to represent a *MAS* structure. An object in an *OO* system, as well as an agent in a *MAS*, is considered as an instance of a class. Classes are represented in terms of a block that contains attributes and operations, where attributes represent the details of a class and the operations imply how to invoke a particular behavior (Pilone and Pitman 2005). Figure 8-8 illustrates the high-level class diagram of the proposed *MAS*. It indicates that *OA*s, *TCA*s and the *GCA* are three types of coordination/operation agents that interact with information agents to constitute the *MAS*. The presentation of *MAS* data structure in form of a class diagram can be discerned as the

starting step in developing ontology for the *LGS*-enabled earthwork project, wherein the formal representation of different types of data and their interdependences are presented.

Information Agents generate the *Project Context* which is an inseparable part of the *Coordination and Operation Agent* class. Note that the inheritance rule of a class diagram dictates that all child classes inherit the features of the parent class. A *Project Context* consists of several operations, each of which in turn comprises several tasks. Different agent types use different context levels, according to their role, and update them regularly as the project execution continues. Contexts have such attributes as ID, the scheduled start and completion times. At the inception of the project, the information agents together with *GCA* generate the initial *Project Context* that will be later used by the *GCA* and *TCA* to generate the *Operation Contexts*. Similarly, a *TCA* uses the *Operation Context* to generate several *Task Contexts* with the help of *OA*s. Information agents are kept updated by the coordination and operation agents along the execution of the projects.

As suggested by Figure 8-8, every *OA* belongs to a piece of equipment which is equipped with several types of sensors. Equipment has several types of attributes that represent its basic features and calculate its performance. Each type of sensors, too, has specific operations to provide the sensory information required by the *OA*s.

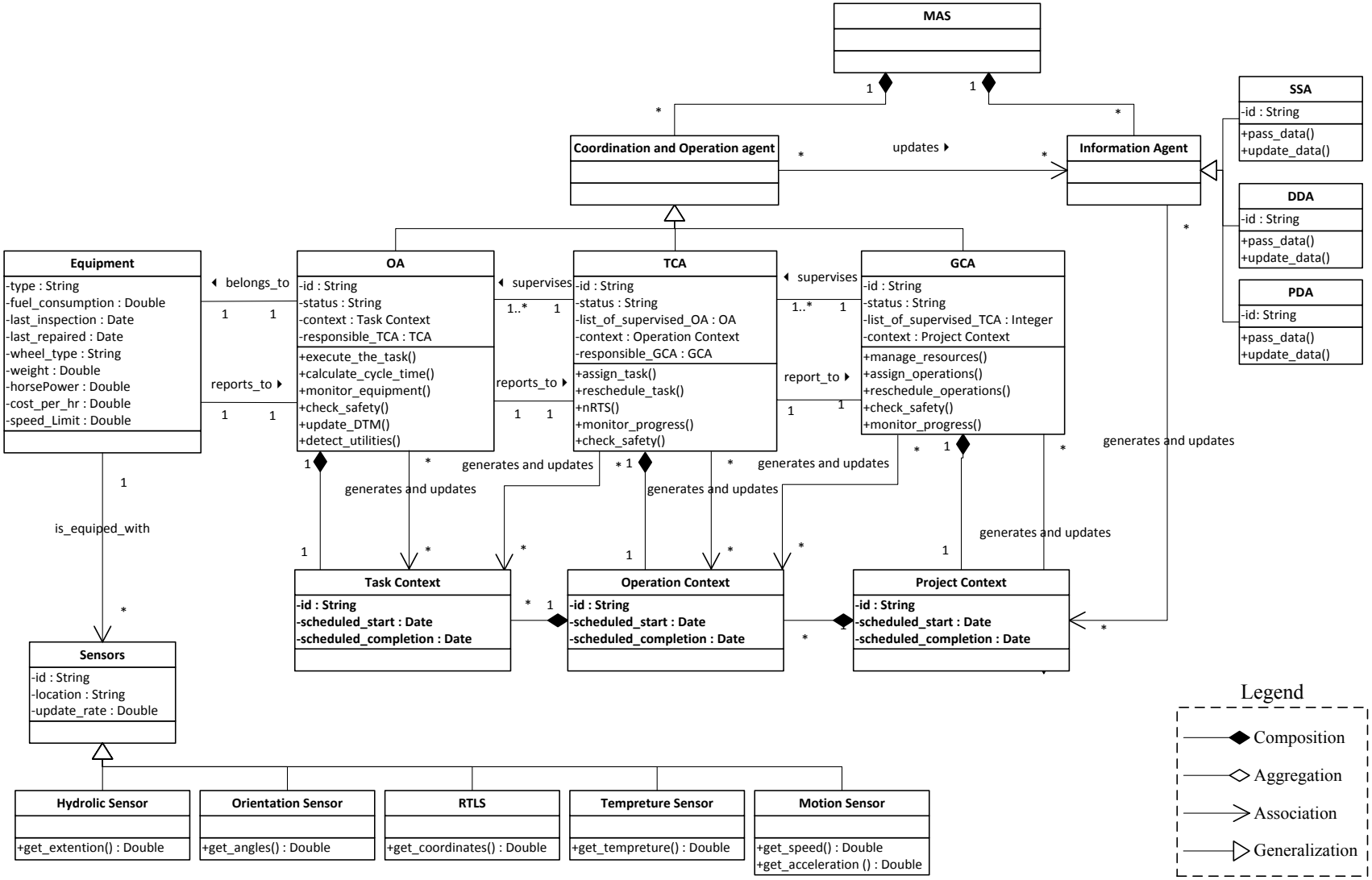


Figure 8-8: High-Level Class Diagram of the Proposed MAS

8.2.4.1 CLASS DIAGRAM OF THE TASK CONTEXT

The high-level class diagram of the proposed MAS, shown in Figure 8-8, can be further extended to illustrate the components of each context. Figure 8-9 depicts the extended class diagram of a task. In conformity with Figure 8-1, a task context includes *Surroundings Information*, *Task Information*, *Environment Information*, and *internal information*. The *surroundings information* consists of pose, state, and *DEW*s. The pose is determined by the location and orientation of the equipment. The state, on the other hand, is represented in terms of a type of action that happens in a workspace. It should be noted that the action types shown in Figure 8-9 are not all-encompassing by any means, and can be further extended based on the needs. The task information class, also, encompasses classes for *Strategic Plan*, *LAEW*, and *Safety Warnings* as well as a *3D Design*. *LAEW*s are represented in terms of their starting and end time, the risk level, and dimensions. The *Environment Information* encompasses *Weather Information* and the *DTM*, which is in turn represented by a set of Triangulated Irregular Network (TIN) consisting of points in the space. Finally, the Internal Information, which is determined by the type of equipment *OA*s represent, comprises such information as equipment fuel level, engine temperature, location data, etc.

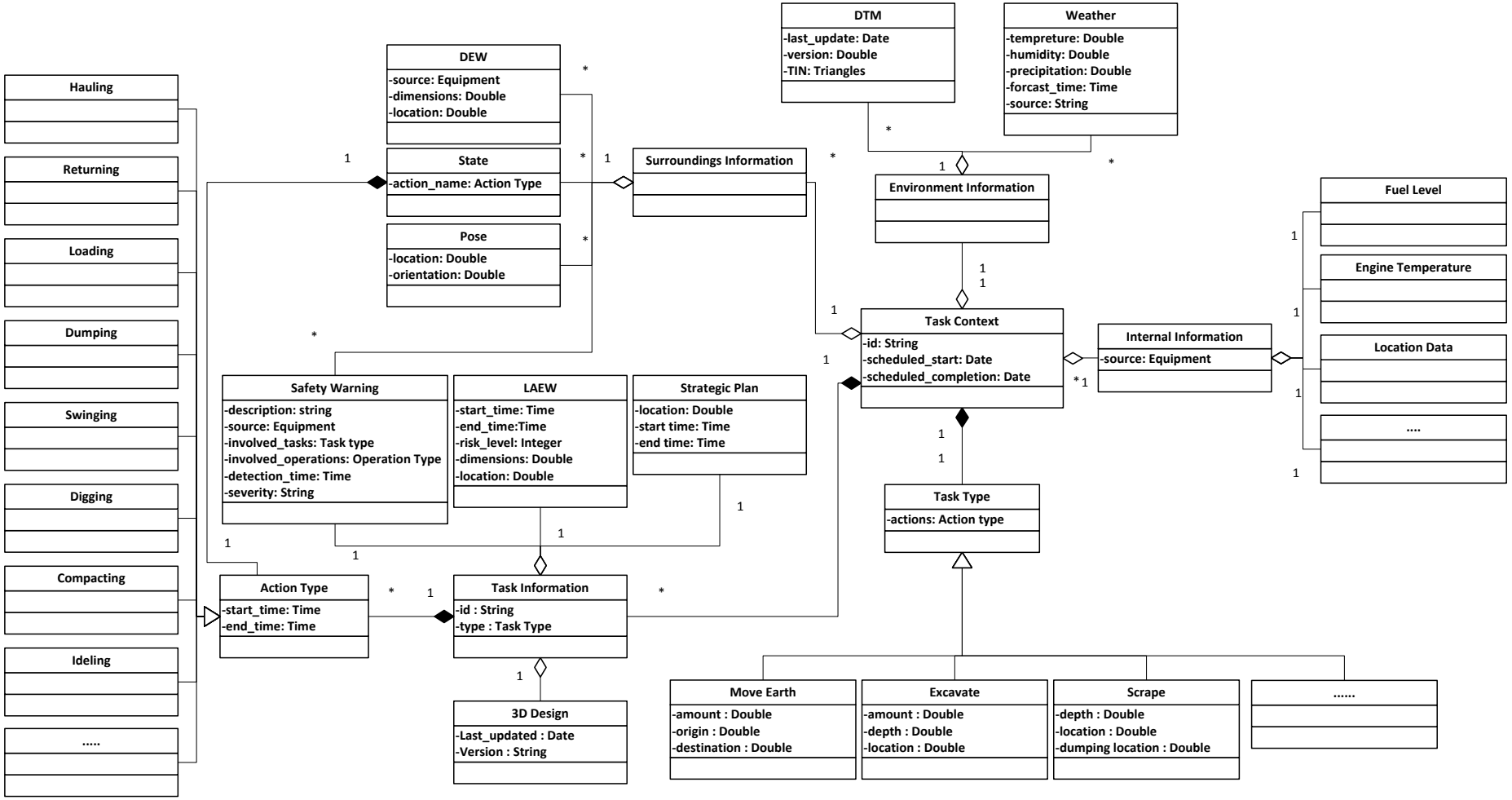


Figure 8-9: Extended Class Diagram of Task Context

8.2.4.2 CLASS DIAGRAM OF THE OPERATION CONTEXT

Figure 8-10 portrays the class diagram of the operation context, which consists of *Task Progress and Status Information*, *Operation Information* and *Environment Information*. The *Operation Information* contains resource configurations, operation schedule, operation logic, and 3D design.

Each operation has a particular type, e.g., grading, compaction, etc., which determines the type of equipment required, amount of work to be done, and the operation logic that needs to be followed. Operation schedule encapsulates information about the start time, end time and the number of tasks as well as the types of tasks that need to be accomplished.

Another component of the operation context is the task progress and status, which comprises a wide gamut of information types as shown in Figure 8-10 and explained in Section 8.2.2. Similar to the task context, environment information is also a part of the operation context and contains *DTM* and weather information.

8.2.4.3 CLASS DIAGRAM OF THE PROJECT CONTEXT

The highest level of the information layer is the project context which includes information classes concerning the *Operation Progress and Status*, and *Project Information* and *Environment Information* as shown in Figure 8-11. The *Project Information* comprises *Project Schedule*, *Available Resources*, *Construction Method*, *Available Sub-contractors*, *Safety Regulations* and *3D Design*. Additionally, the *Progress and Status* encompass various types of information classes as discussed in Section 8.2.2. Similar to the aforementioned contexts, the project context also includes environment information classes.

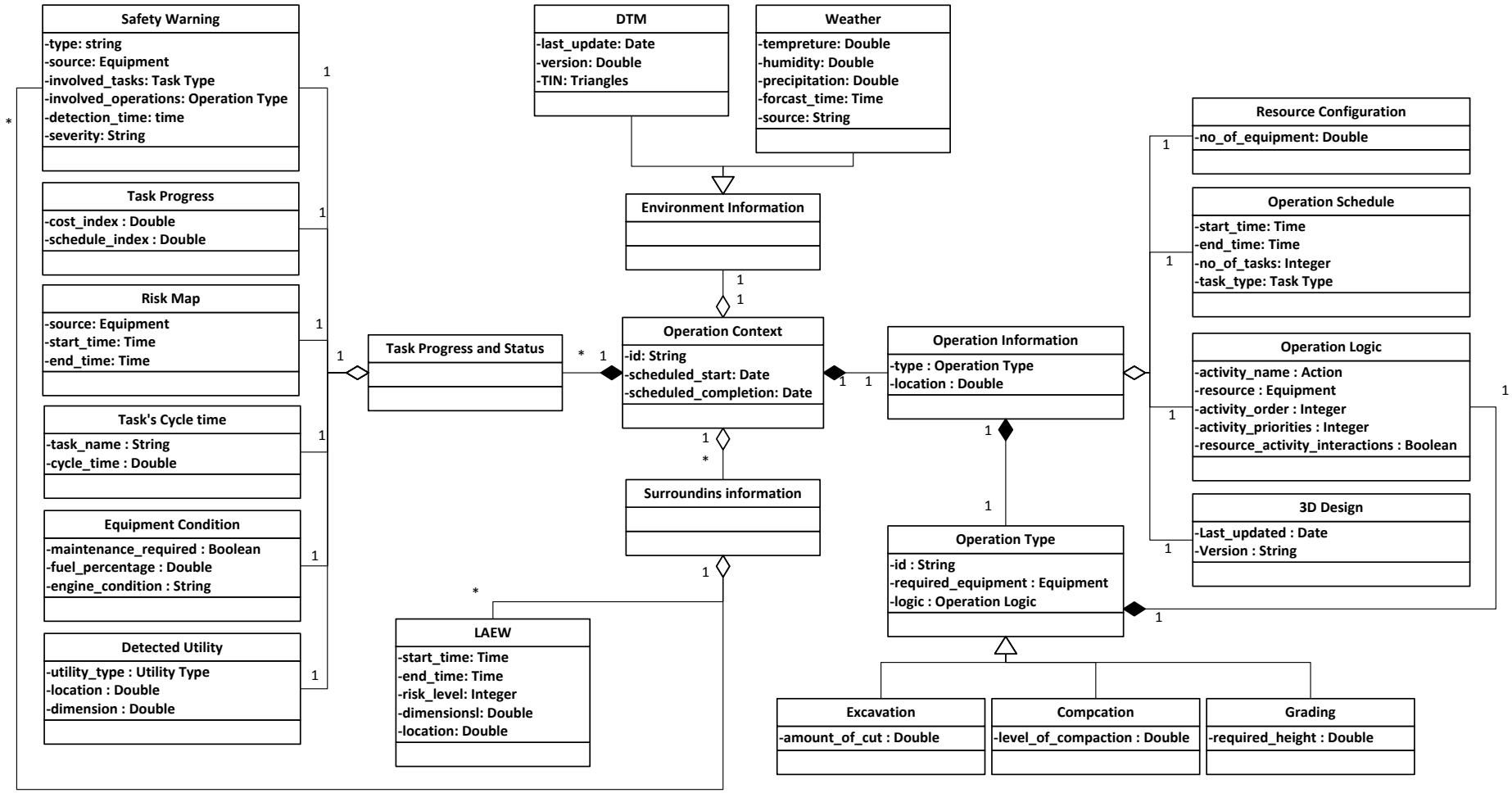


Figure 8-10: Extended Class Diagram of Operation Context

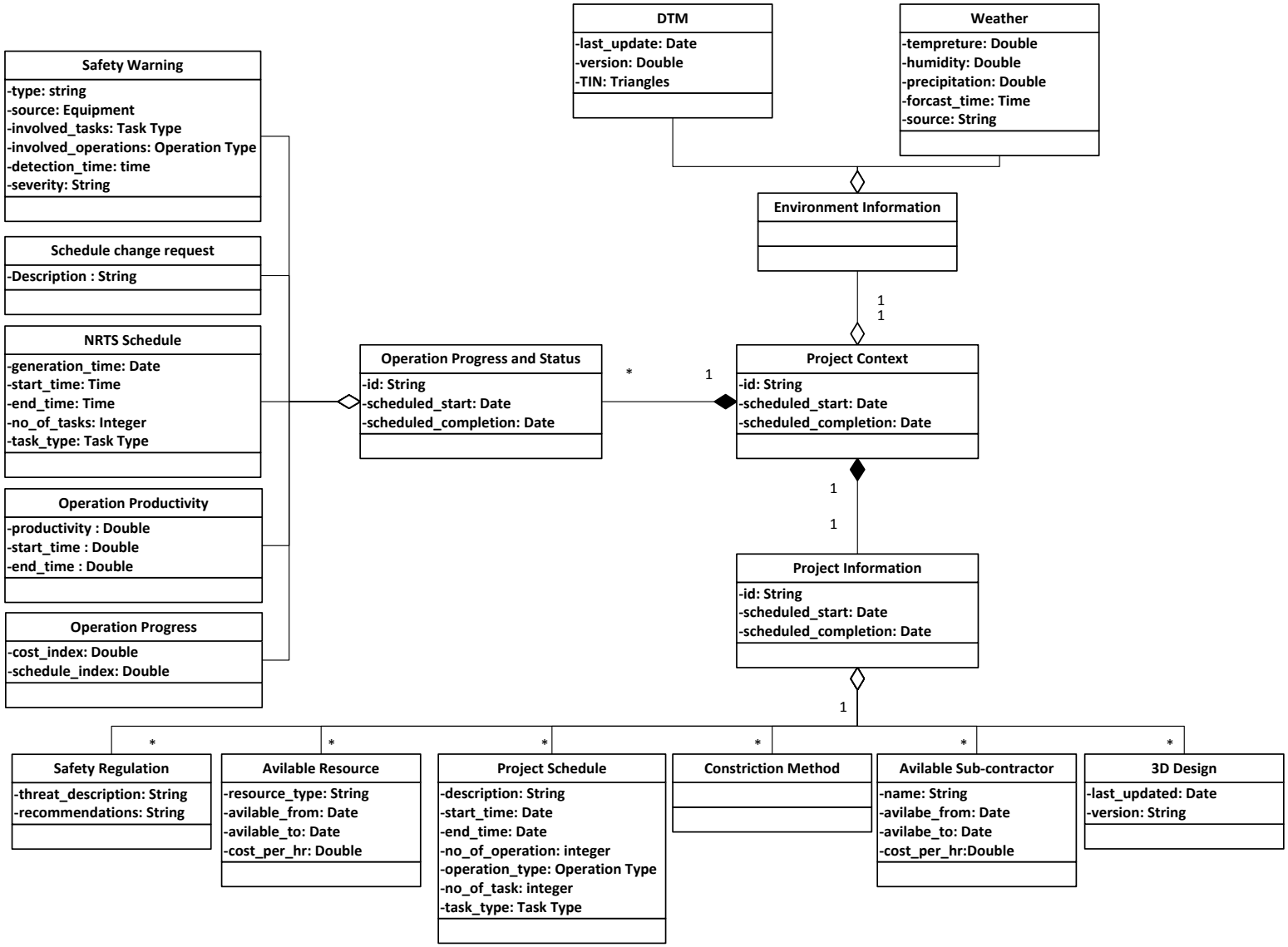


Figure 8-11: Extended Class Diagram of Project Context

8.2.5 TASK AND OPERATION MANAGEMENT IN MAS

In the MAS architecture, agents are required to communicate a large amount of information and to assign operations and tasks to the relevant *TCA*s and *OA*s, respectively. In order to better elaborate on the stream of data flow and the communications sequences, UML sequence diagram is drawn to represent the high level data communication between agents. A sequence diagram represents the messages transferred between several agents in order to perform certain actions.

Figure 8-12 illustrates the sequence diagram that represents the communication schemes between agents from the inception of the project to the end of one cycle of site monitoring, assuming no safety warnings are triggered. At the onset of the project, the *GCA* sends a request to the *SSA*, *DDA* and *PDA* for *DTM*, weather data, 3D design and project data, i.e., the components of the project and environment information. Upon the delivery of the requested documents, the *GCA* assigns operations to different *TCA*s based on the determination of the number of operations, their respective schedules, logics, and needed resources. Then, the *TCA*s re-evaluate weather data, *DTM* and the 3D design to assign tasks to *OA*s in terms of strategic plans.

Once the *OA*s are assigned, they combine the *TCA*-provided data with the data from the *SSA* and *DDA* to generate tactical plans and execute the tasks. Meanwhile, the *OA*s continuously calculate their poses and states, and communicate these data along with cycle times, equipment conditions, and task progress to the *TCA*s, and thus enable the *TCA*s to perform *NRTS*. Subsequently, the *TCA*s share operation productivity, *NRTS*-generated schedule, operation progress and schedule change requests, which might have been necessitated as a result of delays or safety issues.

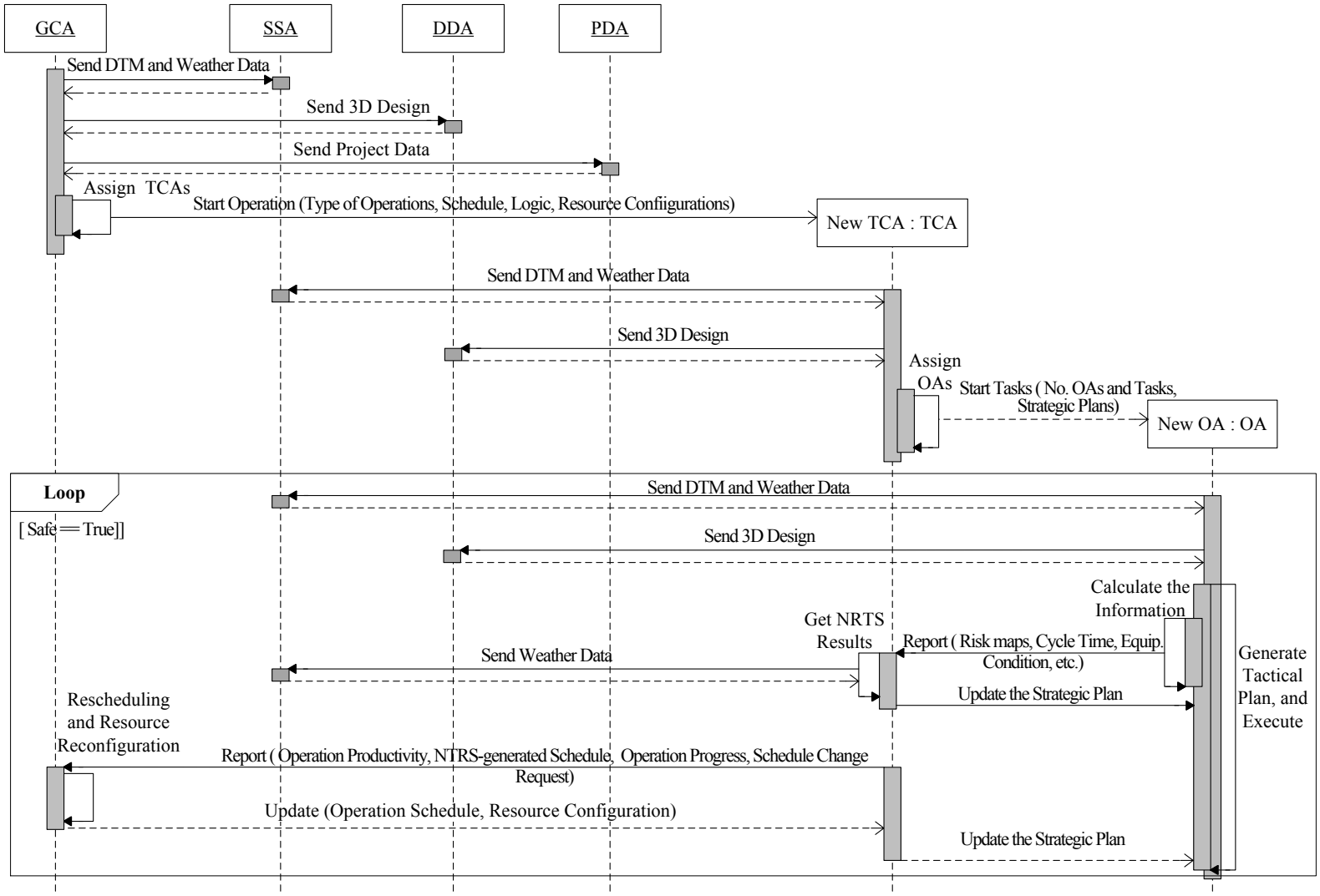


Figure 8-12: Sequence Diagram of Data Communication between Agents for Task and Operation Assignment

The loop box in Figure 8-12 indicates that the monitoring cycles keep iterating to the completion of the project (unless a safety threat is identified).

8.2.5.1 OPERATION LOGIC

One important input from *GCA* to *TCA* is the operation logic. As stated in Section 8.2.2, operation logic represents the logical sequences of tasks performed by different equipment to complete an operation. The operation logic can be provided in terms of an underlying simulation model that formalizes the resources and their interactions. To better put this concept in perspective, a simple excavation operation, where excavators are loading trucks, can be used as an example. Figure 8-13 shows the operation logic for the excavators and trucks in this example. According to Figure 8-13(a), the first step for the excavator is to move to the first designated Digging Station (DS). Once it is at the *DS* and a truck is ready for the loading, the excavator moves to the digging point and performs the digging. Upon completion of the digging, the excavator moves to the truck, and dumps the material onto the truck. This cycle is repeated until the excavation is completed in the *DS*. The excavator moves to the next *DS* if more *DS*s are assigned to the excavator based on the strategic planning. Figure 8-13(b) illustrates the operation logic of a truck. The truck finds the excavator with the shortest queue in the team and wait in the queue until the excavator is ready for the truck. Once the truck is loaded by the excavator, it proceeds to the dumping point. When combined with the pose and state information, the operation logic can be used by *OA*s to determine their current and future actions and to execute their tasks in the logical sequence.

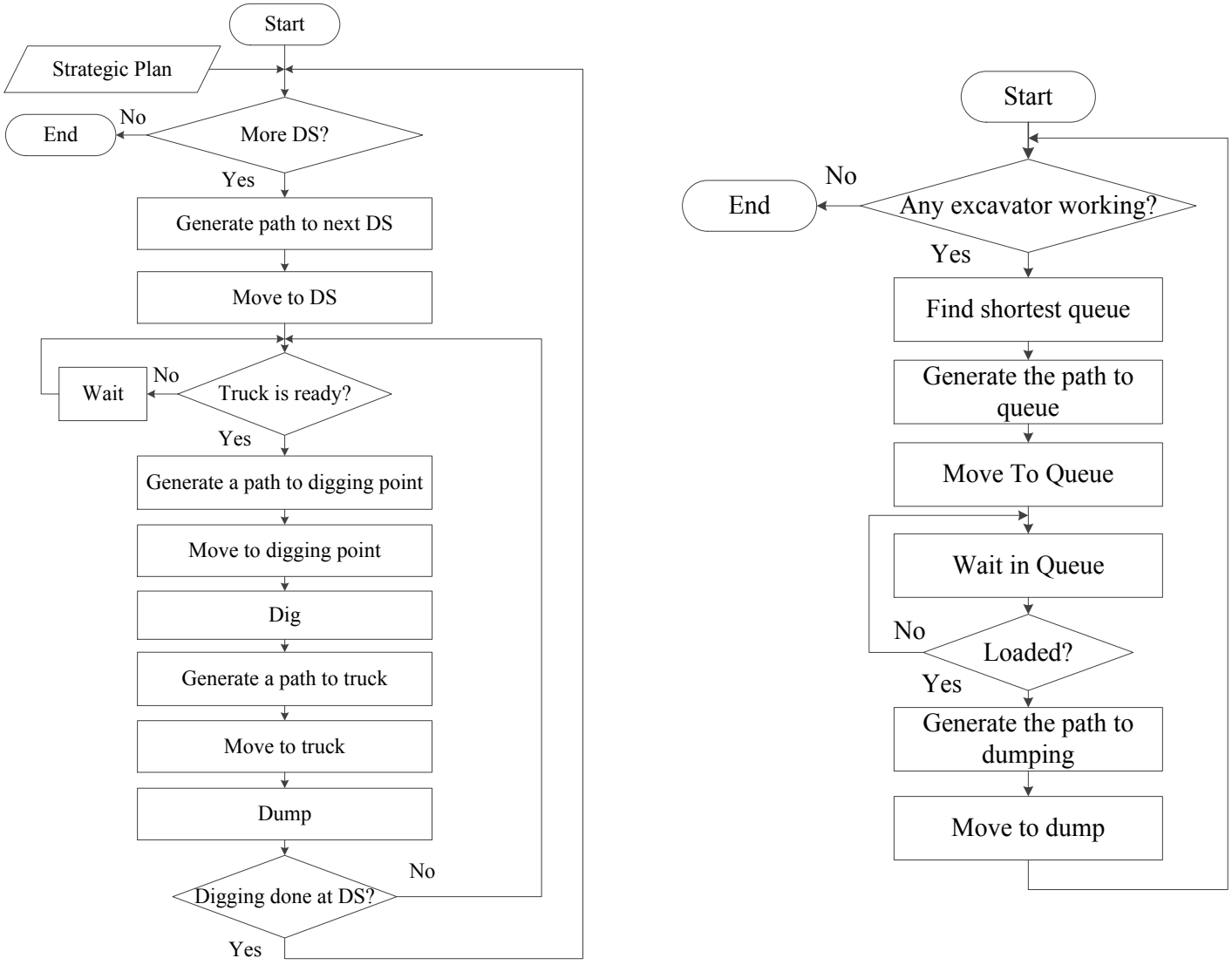


Figure 8-13: Algorithms Representing the Operation Logic of (a) Excavator, and (b) Truck

8.2.5.2 STRATEGIC AND TACTICAL PLANNING

While operation logic can help the *OA*s determine their current and future tasks, more detailed information is required in terms of strategic and tactical planning. Methods for strategic and tactical planning have been previously presented by Hammad et al. (2014), which are explained briefly in the following.

The strategic plan can be either based on the input from the project manager, for one-of-a-kind and complex operations, or based on the semi-automated method presented by Seo et al. (2011) for common and repetitive mass excavation operations. In the proposed semi-automated method, the sequence of *DS*s is determined considering the dimensions of the excavators and the number of required turns for the excavator.

Tactical planning, on the other hand, can be performed using advanced path planning algorithms at two levels, namely macro and micro. Macro level tactical planning is concerned with the movement of a piece of equipment (as a whole) from one point to another on the site, for instance the traversal movement of an excavator. On the other hand, micro level path planning focuses on the articulated movements of different parts of the equipment along different *DOF*s that enables the execution of such tasks as digging or swinging for an excavator. Several algorithms have been proposed for this purpose in the literature. For instance, Kim et al. (2012) proposed intelligent navigation strategies for an automated earthwork system considering safety and productivity at the macro level using Sense-bug algorithm for path planning. A* is another suitable 2D path planner that generates an optimal macro-level tactical plan. In this research, A* is proposed for the macro-level tactical planning because it is complete and optimal, which can dramatically help in reducing the cycle time of the equipment.

In spite of the effectiveness of A* for macro-level planning, it cannot be effectively used for micro-level tactical planning given that it is not designed to handle complex planning for equipment with multiple *DOF*s. Other algorithms such as the Rapidly-exploring Random Tree (RRT) and parametric scripting are suggested for the purpose of micro-level tactical planning in several researches (Stentz et al. 1999; AlBahnassi and Hammad 2011). In this research, parametric scripting is proposed to be used in the planning phase, given its effectiveness in capturing the behavior of expert operators and generating fast and reliable tactical plans. However, in case of any unexpected situation, such as possible collisions with other equipment, parametric scripting may fail since it is based on a set of predefined rules. As a result, *RRT* is proposed to be used for tactical re-planning.

8.2.6 SAFETY MANAGEMENT IN MAS

With respect to safety issues, depending on the criticality of the safety hazard, a variety of actions can be adopted, ranging from immediate autonomous suspension of tasks/operations until the threat is cleared to sending a request for the schedule change. The proposed *MAS* manages the safety of earthwork operations through the effective use of an spectrum of information (e.g., *DEW*s, *LAEW*s, equipment conditions). Two types of scenarios are conceivable, namely: (1) safety hazards that can be resolved using collision avoidance mechanisms by individual *OA*s (e.g., potential collision between equipment and workers), and (2) safety hazards that require more managerial decision-making by *TCA*s, *GCA*s, or designers (e.g., detected underground utilities and damaged equipment).

8.2.6.1 OPERATION-LEVEL COLLISION AVOIDANCE MECHANISM

While the paths generated by *OA*s through tactical planning at the task and operation management levels are planned to be collision-free, a collision avoidance mechanism is required to ensure that the paths remain collision-free in view of all the human factors and uncertainties involved in the execution phase. As stated in Section 3.2, the collision avoidance in the proposed *MAS* structure is supported through a two-layer mechanism which includes near real-time collision-free path re-planning using *LAEW*s and real-time collision avoidance using *DEW*s. These two layers are running independently in parallel with different update rates. Given the nature and functionality of *DEW*s, they are updated in real time with the same rate offered by pose estimation method (Δt). *LAEW*s, on the other hand, require intensive computations and communications between various agents, and thus they are updated with a rate less than *DEW*s. The *LAEW*s are generated over every δt and whenever a deviation from the predicted path of various pieces equipment is observed.

8.2.6.2 MANAGEMENT-LEVEL SAFETY SUPPORT MECHANISM

Safety threats that cannot be addressed by *OA*-level decision-making are tackled through a chain of communication between different layers of agents, as shown in Figure 8-14. Typically, an *OA* informs its *TCA* about the safety threats and requests direction. Then, depending on whether or not the *TCA* is capable of resolving the issue through intra-team coordination (e.g., assigning new tasks to the *OA*s), either the solution is provided to the subordinate *OA*s by means of a new strategic plans or the request is further escalated to the *GCA*. If the *GCA* is solicited for the intervention, in a similar manner to the *TCA*, depending on whether or not the problem can be resolved through resource re-configuration and operation-level re-scheduling, either the solution

is provided to the *TCA*s, who subsequently transfer the solution to the subordinate *OA*s, or the problem is sent to the designer for review and potential design change.

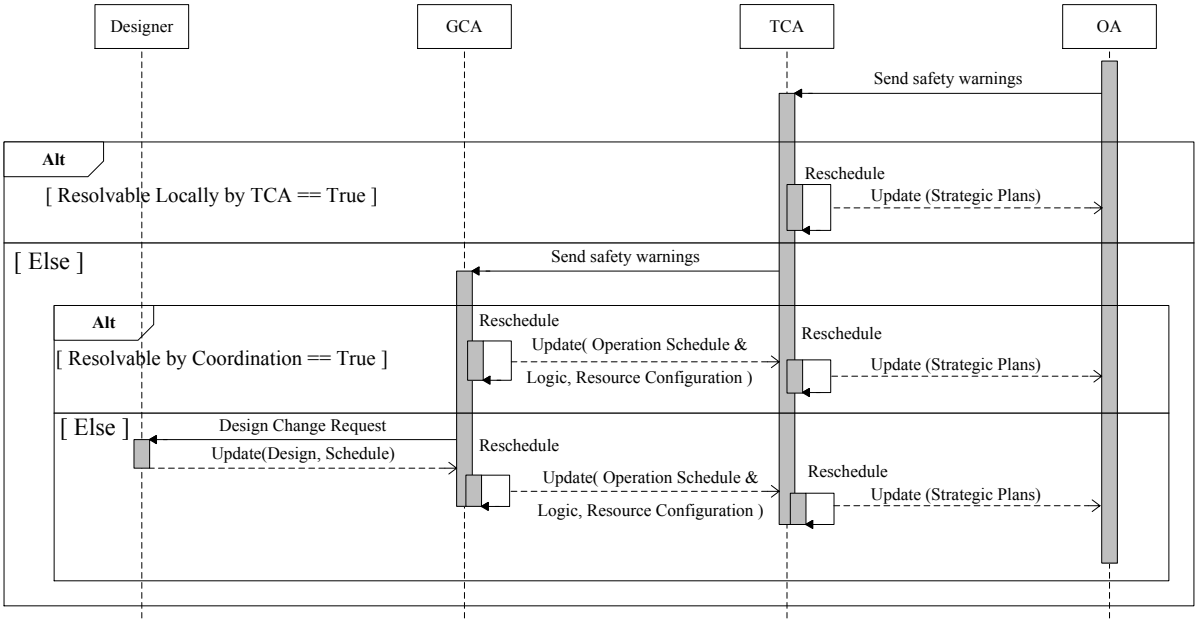


Figure 8-14: Sequence Diagram of Data Communication between Agents for Safety Warnings

For instance, if a collision between two pieces of equipment with similar priorities is identified through *DEW*s and both pieces of equipment stop, then the *TCA* would determine which equipment to proceed first. However, if the two teams are working in dangerously close proximities, the *TCA*s ask the *GCA* to require the team whose operation has the greater free slack to move to a new location. In another example, if a pipe is found at an unexpected location on the excavation site, the *TCA* would ask the *GCA* for direction. The *GCA* then would ask the designer to make a decision about the potential change of design.

Priority rules need to be used to resolve problems both at the *OA* level and managerial level, as explained above. As for the priority rules within a team, a set of factors such as operating costs, productivity rates, equipment size, type of material being hauled, etc., are considered. However, the priority rule between two pieces of equipment is not necessarily fixed throughout the

operation. Depending on the states of the equipment and which one of them is better able to avoid collisions, the priority rules amid the pieces of equipment can change in the course of the operation. For instance, Table 7-1 shows the state-dependent priority rules for a truck and an excavator when they are working together in a team. If the truck is in the stationary states (e.g., loading or waiting for load), it is the responsibility of the excavator to avoid the truck. On the other hand, if the truck is in the traversal states (e.g., hauling or maneuvering) and the excavator is in the stationary states (e.g., loading or swinging), the truck should avoid the excavator. Finally, if both the excavator and truck are in traversal states, the equipment with the higher operation costs and productivity rate (i.e., often the excavator) should have the right of way.

At the team level, the priority of a team is decided according to several criteria such as the criticality of the operation undertaken by the team and its impact on the timely completion of the project.

Table 7-1: State-dependent Priority Rules for a Team Consisting of a Truck and an Excavator

		Excavator	
		Stationary States	Traversal States
Truck	Stationary	Excavator should avoid the truck	Excavator should avoid the truck
	Traversal states	Truck should avoid the excavator	Equipment with the lower operation costs and productivity rate (usually truck) should avoid the other one

8.3 IMPLEMENTATION AND CASE STUDIES

In order to demonstrate the feasibility of the proposed method, an implementation is developed and tested in three simulated case studies. The first case study addresses the task and operation management in *MAS*, i.e., combining the strategic and tactical planning, and how *DEW*-based safety management acts in a typical operation. The second case study aims to demonstrate the applicability of the *LAEW*-based safety management and how it can be used in conjunction with tactical re-planning. Finally, the third case study addresses the management-level safety system and agent's communication in MAS.

8.3.1 IMPLEMENTATION

Two prototype systems have been implemented. The first prototype system is developed using C# *Application Programming Interface* (API) in Unity3D game engine (2015) to demonstrate the feasibility of agent-based decision-making and planning within a spatially-aware environment. This prototype system is developed in three steps:

(1*) 3D environment creation*: The case studies were inspired from the Turcot interchange project (Transports Quebec 2014), which is a large interchange reconstruction project in the city of Montreal. In this step, the DTM of the area of the case studies was converted to a *TIN* format using ArcGIS (2015). Then, it is exported to *FBX* file format and imported into Unity. It is then converted to the terrain type used by Unity (Height Map). An aerial photo is draped on top of the height Map.

(2) *Task and Operation Management:* The *OA*s for excavators and trucks are developed using the operation logics shown in Figure 8-13(a) and (b), respectively. In this implementation, the project manager defines the strategic plan for different pieces of equipment using the Graphical

User Interface (GUI) developed in Unity. Apart from the initial micro-level tactical planning, which is performed using parametric scripting (Rowe 1999; Kamat and Martinez 2005; Sarata et al. 2006), A*, parametric scripting, and RRT algorithms are developed inside Unity to perform macro-level and micro-level tactical planning, as mentioned in Section 8.2.5.2.

(3) *Operation-level Safety Management:* The methods shown in Figure 6-3 and Figure 7-10 for the *DEW*-based and *LAEW*-based safety management by *OA*s, respectively, are developed in Unity. As mentioned in Section 8.2.6.1, the *LAEW*-based and *DEW*-based collision avoidance calculations are performed over every δt and Δt, respectively, where δt is longer than Δt.

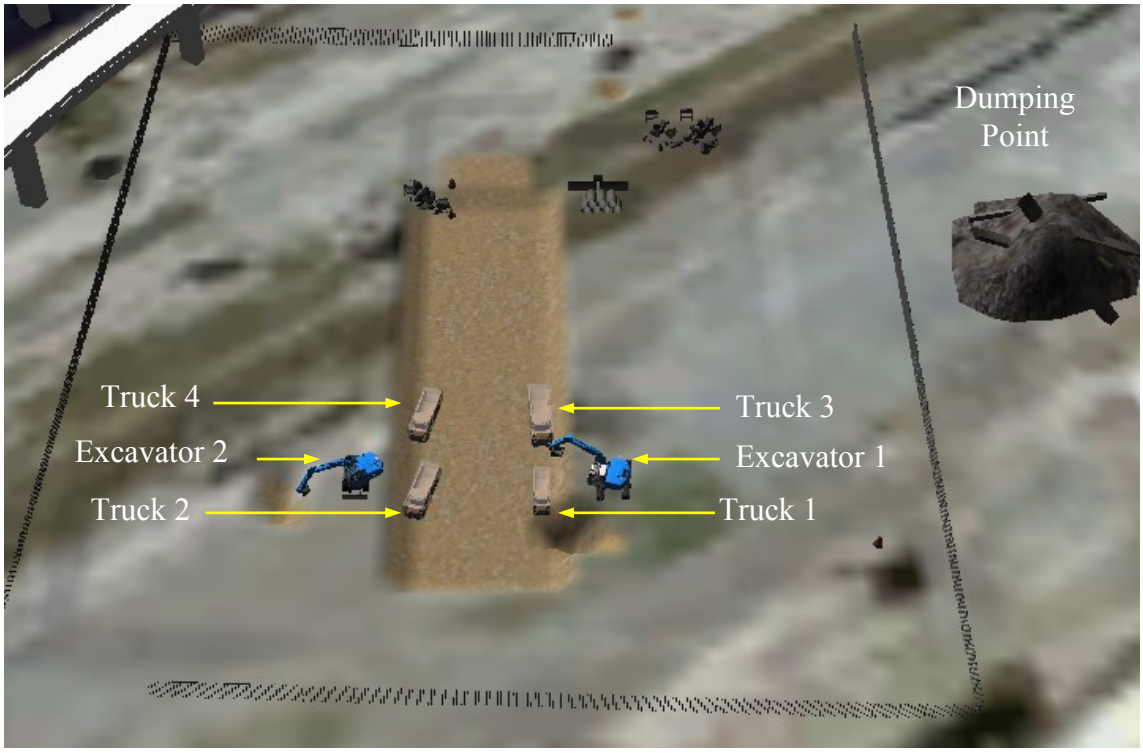
The second prototype systems is implemented in *Java Agent Development Framework* (JADE) (Bellifemine et al. 2007) to demonstrate the applicability of this specialized *MAS* development tool for representing the proposed *MAS* architecture, focusing on agent communication. In this prototype system the communication scheme between a network of agents is simulated using the *Foundation for Intelligent Physical Agents* (FIPA) specifications. *FIPA* allows the platform to standardize the communication between heterogeneous agents (Pitt and Bellifemine 1999). In the specific scenario simulated in this prototype, the managerial-level safety management aspects of proposed *MAS* are investigated.

Although the two prototype systems are loosely-coupled at this stage of the research, the full integration will be addressed in the future work of the author. In the fully integrated system, spatial calculations related to *MAS* within unity can be smoothly handled by *JADE* to apply the *MAS* structure for dispersed agents.

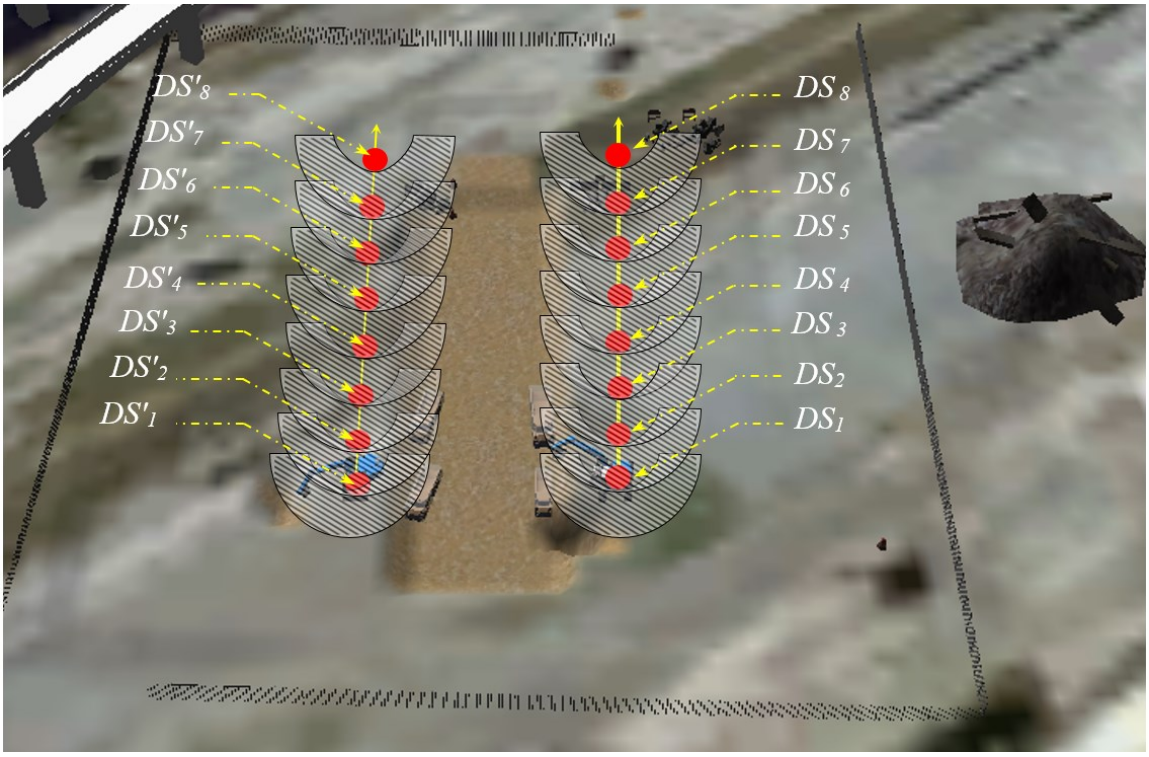
8.3.2 CASE STUDIES

8.3.2.1 FIRST CASE STUDY

As shown in Figure 8-15(a), a simulated earthmoving operation is designed in Unity where two excavators and four trucks are working together as a team. The case study assumes that the *GCA* has already transferred the operation information to the *TCA*, through the resource configuration, operation schedule, operation logic, and the 3D design of the model. The starting point of the case study is where a *TCA* uses these inputs to generate appropriate *OA*s for the two excavators and four trucks. The operation schedule is given in the form of the type of tasks to be completed by the *OA*s, and the start and end times of the operation. The operation logic is provided through the implementation of the algorithm shown in Figure 8-13. In this case study, the strategic plans for the excavator are given through manager-defined plans as shown in Figure 8-15(b), following the pattern presented by Seo et al. (2011). As shown in this figure, each excavator has a designated start and end points and a route connection the two points, which are defined through a graphical user interface. Points DS_1 to DS_8 and DS'_1 to DS'_8 represent the strategic plans for excavators one and two, respectively. Once the simulation runs, the *OA*s follow their plans based on their respective operation logics. While the operation is in progress, the *OA*s monitor and collect progress data and transfer them to the *TCA*.



(a)



(b)

Figure 8-15: (a) Layout of the First Simulated Case Study and (b) Schematic Representation of the Strategic Plans for Excavators

Figure 8-16 shows an instance of how *DEW*-based safety management is capable of avoiding immediate collisions between equipment.

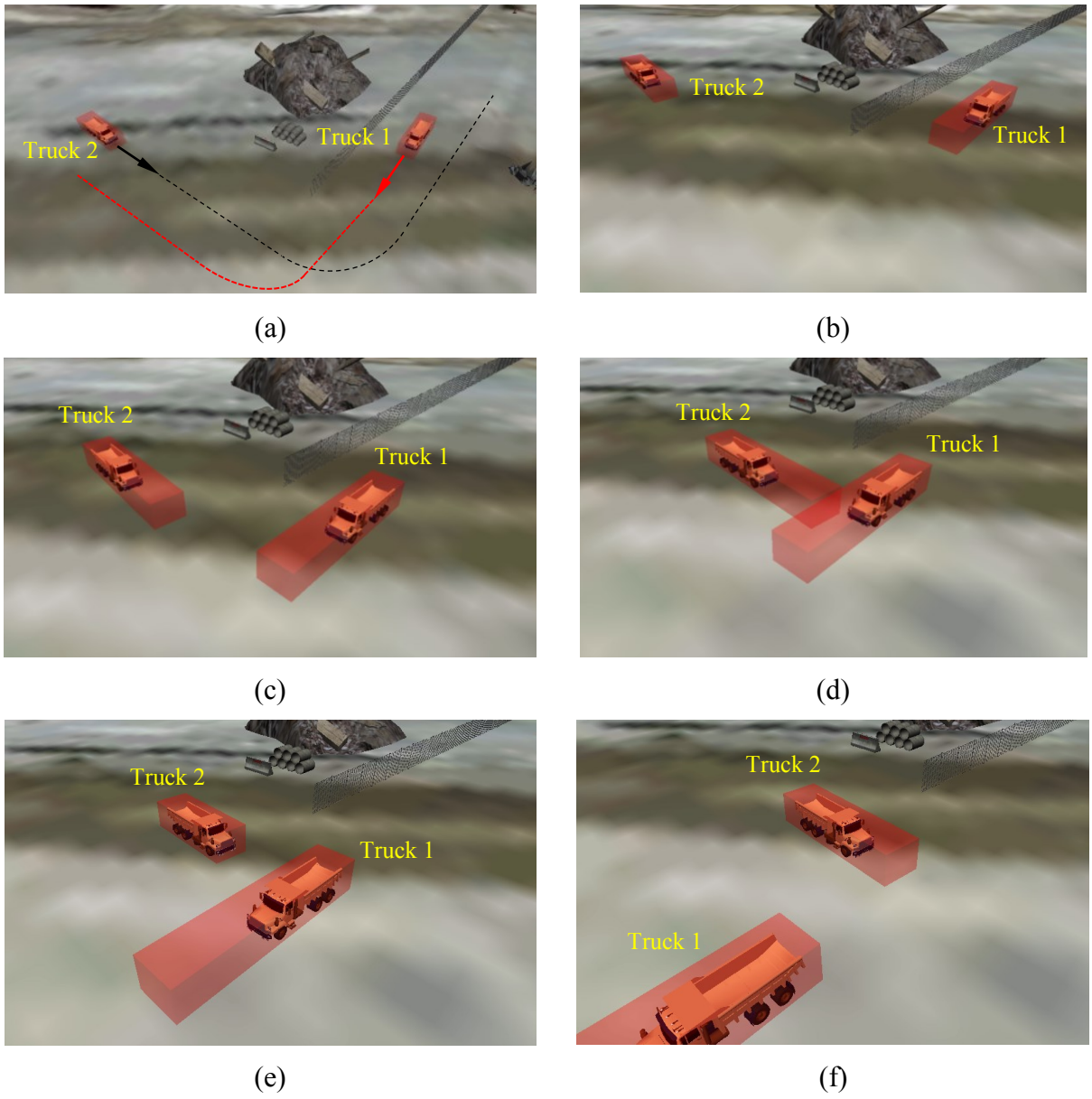


Figure 8-16: (a) The Layout of the Second Scenario, (b) and (c) Different Snapshots Showing the Dynamic Sizes of DEWs, (d) Collision Detection between DEWs, (e) Stoppage of Truck 2, and (f) Both Equipment Proceeding with Their Paths

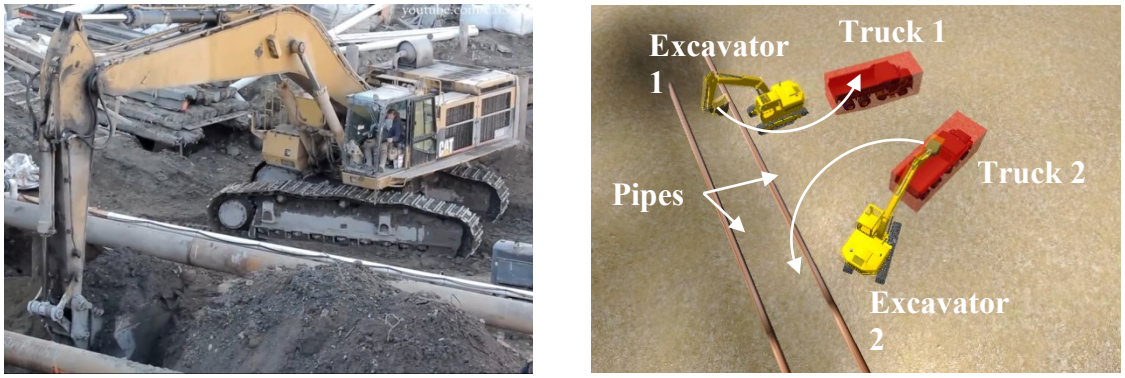
As shown in Figure 8-16(a), while Truck 1 is hauling the material to the dumping point, Truck 2 is returning from the dumping point to the corresponding *DS*. It is assumed that although their

planned paths have been collision-free, Truck 1 fell behind its planned path, which could lead to a potential collision. In this case, as explained in Section 7.4, *DEW*s can be used as the last line of defense to avoid the collision by requiring one of the equipment to stop. A higher priority is given to Truck 1 since it is loaded. Figure 8-16(b) and (c) show how the size of *DEW*s is changing based on the speed characteristics of the equipment. Figure 8-16(d) depicts the collision between the two *DEW*s. Given its lower priority, Truck B stops and waits until Truck A passes, as shown in Figure 8-16(e) and (f).

8.3.2.2 SECOND CASE STUDY

In the second case study, the feasibility of applying *LAEW*s for collision-free path planning is investigated. In this scenario, which is inspired from a real excavation site shown in Figure 8-17(a), two excavators are digging the ground between two pipelines. Excavator 1, which is assumed to have a higher priority than Excavator 2, is expected to swing to Truck 1, and Excavator 2 is expected to swing away from Truck 2. The excavators are working in close proximity and should avoid the pipes and the terrain as static obstacles. According to the proposed MAS, the *TCA* is in charge of the safe and efficient management of this operation.

As shown in Figure 8-17(b), the initial paths of the excavators are planned using parametric scripting, as explained in Section 8.2.5.2. Figure 8-17(c) shows the risk map of Excavator 1 with a threshold of 0.8 for the next 2.3 seconds. Since the initial path of Excavator 2 collides with the generated *LAEW*, Excavator 2 uses RRT path planner to generate a new path that will avoid the potential collision. The bucket trajectory of Excavator 2 and the final configurations of both excavators are shown in Figure 8-17(d). The results illustrate that the *MAS* is able to effectively use *LAEW*s to predict the potential future collisions and generate new paths to avoid such collisions.



(a) (b)

(c) (d)

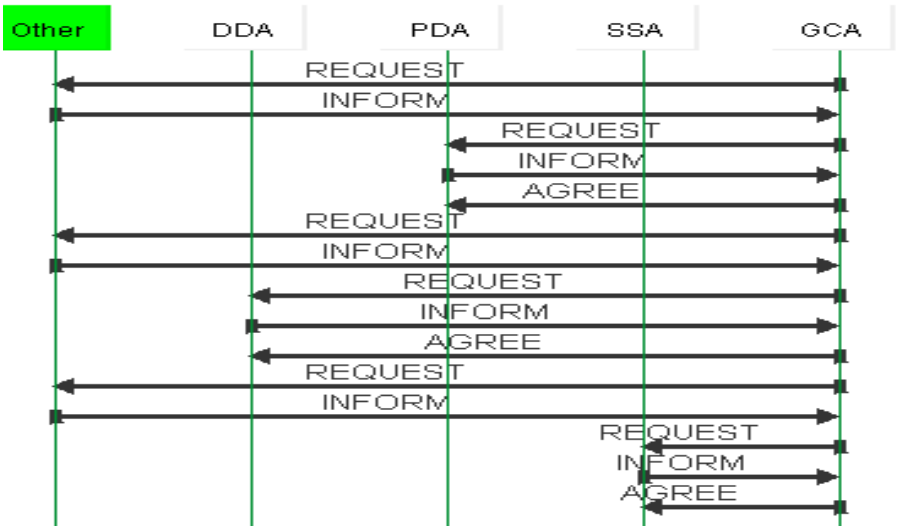
Figure 8-17: (a) a Real Case Scenario (b) Current Poses and Initial Paths of Excavators, (c) LAEW of Excavator 2 and (d) Final Path of Equipment 2

8.3.2.3 THIRD CASE STUDY

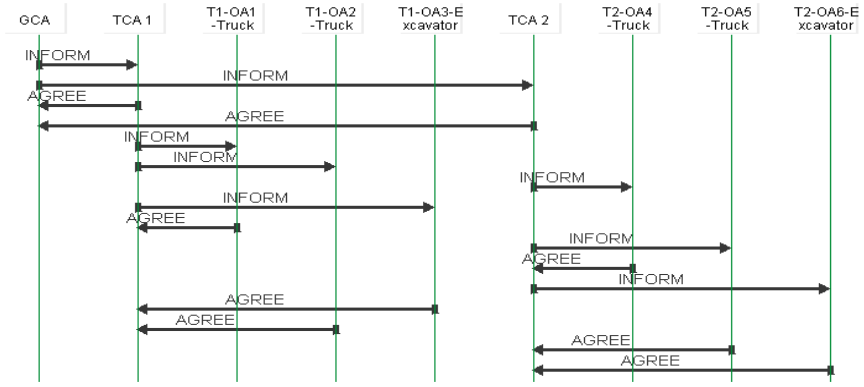
In order to demonstrate the agents' communications in the *MAS* structure, this case study was implemented. The scenario simulated in this case study is composed of two teams of equipment working in an earthwork site. In each team, one excavator is coupled with two trucks. This scenario starts from a point where the *GCA* uses the input from the project manager to form the required operation profiles. In addition to simulating the communication between agents for task assignment and monitoring, a safety-related issue is also simulated where the *OA* of the excavator in team one discovers an unexpected pipe, which was not accounted for in the plan, and escalates this issue to the *TCA* 1.

At the current level of the implementation the project manager needs to define the number of *OA*s and *TCA*s, but this can be further enhanced so as to use rule-based systems or optimization methods for the determination of the number of *OA*s and *TCA*s based on the daily work volume, available resources, and geographic dispersion of scheduled work. Figure 8-18 shows the sequences diagrams generated by *JADE* for different communications between agents. As shown in Figure 8-18(a), the *GCA* requests data from *SSA*, *PDA*, and *DDA* using messages in form of *Extensible Markup Language* (XML) files, in accordance with Figure 8-12. The agent *Other* is a default agent in the *JADE* platform that is responsible to inform different agents about how to find their target agents, in form of *Internet Protocol* (IP) addresses. Each *Inform* message from an agent is followed by an acknowledgement of the receipt of the data in form of an *Agree* message from the recipient agent. As shown in Figure 8-18(b), once all the required information is gathered from *Information Agents*, the *GCA* generates the operation *XML* files and sends them to the *TCA*s, who will in turn send the tasks files to the relevant *OA*s. Upon the start of the execution, each *TCA* starts monitoring its team and receives the *XML* messages from the *OA*s, as shown in Figure 8-18(c). The types of data communicated in this scenario are the data about surroundings, pose, state, liner speed, angular speed, and safety warnings. When the warning message from the *OA* of the excavator in the first team is received about the identified underground pipe, TCA_1 informs the *OA* in form of a *Refuse* message that this issue cannot be resolved by the team level coordination because the pipe was not accounted for in the design. The safety issue is then escalated to the *GCA* who in turn communicates with the designer agent to request for a design change. In this case study, only the escalation of the safety issue to the designer was demonstrated. However, in future research, more decision-making parameters can

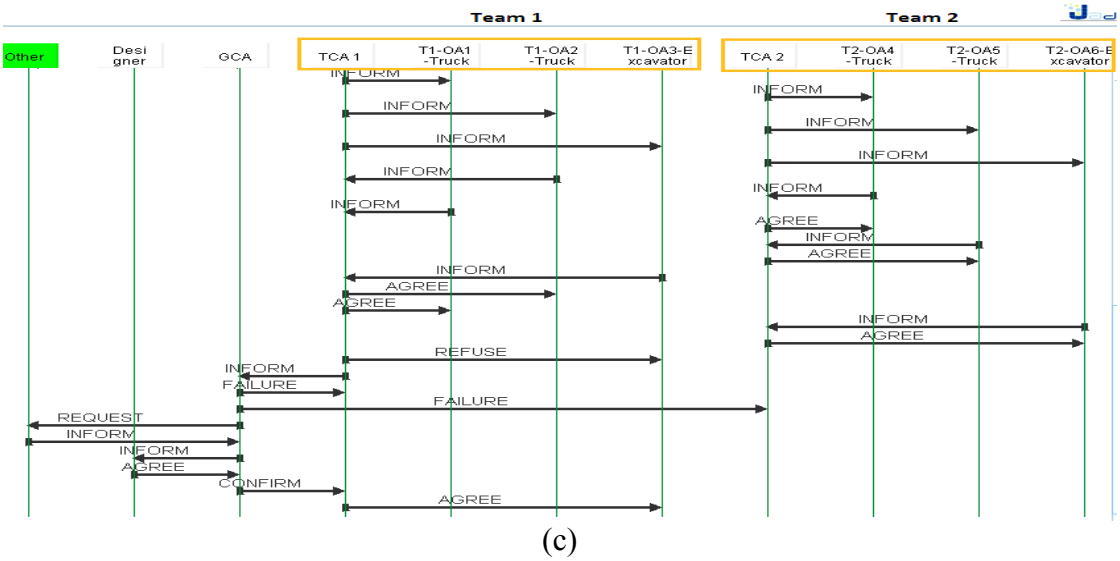
be added to enable the *GCA* to dispatch the involved teams to a new location while the managerial decisions are being made about the discovered pipe.



(a)



(b)



(c)

Figure 8-18: JADE Sequence Diagrams for Communication between (a) the GCA and Information Agents (b) the GCA, TCAs, and OAs, (c) Different Agent During Monitoring

8.4 SUMMARY AND CONCLUSIONS

The present chapter is the result of an effort to develop a *MAS* structure to provide the basis for a fleet-level automated equipment guidance system. In the proposed *MAS* structure, each piece of equipment is supported by an operator agent to oversee the task and provide guidance whenever needed. A multi-layer agent hierarchy assigns, monitors and coordinates the task execution, and a set of three types of agents feed the system with the relevant information. The functionalities, jurisdictions and the input-output scheme of every type of agents were discussed in detail. The sequence diagram was used to schematize the communication inter-dependencies and inter-connectivity of various types of agents, the type of information being communicated among them. Also, the task and operation management and safety management mechanisms were elaborated in detail.

The proposed approach helps expand the effective domain of *LGS* from machine-level guidance to fleet-level coordination. In the view of the presented case studies, the *MAS* structure is found to be effective in assigning different operations and tasks of a project to the specific agents that will be responsible for their realization. Using a combination of strategic and tactical planning methods, the *MAS* is able to effectively provide readily executable guidance/control for equipment operators considering a variety of safety issues. The proposed *MAS* structure is shown to have a potential for improving the safety of earthwork sites through enabling smooth coordination between a fleet of equipment. Safety issues and conflict-prone activities are addressed using the proposed two-layer safety mechanism that accounts for a wide range of human factors and uncertainties, e.g., operators' potential non-compliance with the suggested paths.

While the present research mainly focused on the safety aspect of *LGS*-based fleet coordination, the *MAS* structure is expected to contribute as well to improving the productivity of the operations. Similarly, the proposed method can be leveraged to improve the communication and decision-making at the managerial level through the provision of real-time sensory and simulation data.

The main limitation of the presented research is that the full-fledged implementation of the proposed framework and complete integration of different methods in near real-time requires a level of computational power which currently may not be economically justified to provide to earthmoving equipment. Having said that, considering the current rate with which the available computational power improves, fully integrated implementation is expected to become viable in the near future. The second limitation of the present research is the challenge of comparing this method with the available commercial fleet management systems (Noland 2010; Jamasmie 2010). Given the proprietary nature of these systems and the fact that there is very little information available on their technical aspects, a comprehensive comparison is not feasible.

The future work aims to investigate the impact of the proposed method on improving the productivity of operations and managerial decision making. Also, the author is dedicated to further improving the efficiency of the proposed algorithms and to developing a fully-integrated system that can be tested on construction sites.

CHAPTER 9: CONCLUSIONS, LIMITATIONS AND FUTURE WORK

9.1 SUMMARY

The present research is the result of an effort to improve the coordination between a fleet of earthwork equipment through enabling the inter-equipment communication and interaction using *LGS* and *MAS* approach. The particular focus of the this research was place on: (1) enabling the project-level coordination, monitoring and control through the integration of a *MAS* architecture and *LGSs*; (2) improving the pose estimation accuracy of more affordable Real-Time Location Systems so that a large number of older earthwork equipment can benefit from the advantages of *LGS*s without going through the costly retrofitting procedure; (3) developing an *LGS*-based Near Real-Time Simulation approach; and (4) improving the safety management system under the *MAS* structure that can leverage the capabilities of *LGSs* and Near Real-Time Simulation to ensure the safety of the earthwork operation.

9.2 CONTRIBUTIONS AND CONCLUSIONS

This research made the following contributions to the body of knowledge:

(1) A novel approach was introduced to improve the quality of data captured by less expensive Real-Time Location Systems so that the pose of the equipment can be estimated with an acceptable level of accuracy. A two-phase optimization-based method was proposed that uses a set of geometric and operational features of an excavator and the information about the position of data collectors attached to different parts of an excavator to correct the location

data captured by the Real-Time Location System. The proposed optimization-based method improved the location accuracy over raw Real-Time Location System data by more than 60%. Also, the proposed optimization-based method improved the accuracy of orientation estimation by 77%, in the first case study, and 18% in the second case study. With regard to this contribution the following conclusions can be drawn:

- The conventional simplified method of correction was substantially outperformed by the proposed optimization-based method;
- The proposed optimization-based method has a very strong potential to improve the pose estimation using redundancy of more affordable Real-Time Location Systems data collectors;
- The estimated pose can be used to identify the potential safety hazards and also to determine different states of the excavator, which can be later used to calculate the cycle time.

(2) A comprehensive approach was presented for Near Real-Time Simulation where the input from the pose estimation module is used to identify equipment state, track the cycle time of the operation, and update the initial simulation model to make a more accurate and reliable prediction about the project progress. Based on the case study that demonstrated the feasibility of the proposed approach, the following conclusions are drawn:

- The proposed overarching Near Real-Time Simulation approach provided a tracking-technology-independent method for processing, analyzing, filtering and visualizing the equipment states that can work with various types of Real-Time Location Systems technologies and under the availability of different levels of sensory data;

- The developed rule set were able to capture the finest details of truck and excavator motions through the concurrent consideration of the fleet as a whole, and for several scenarios where various range of sensory and location data are individually or collectively deployed. Furthermore, pattern analysis and cycle logic further improved the results of the state identification.

(3) A novel method was presented for the generation of real-time dynamic equipment workspaces considering the pose, state, geometry, and the speed characteristics of the equipment. This method is built on the previous two contributions, where robust methods for the calculation of pose and state of different pieces of equipment based on Real-Time Location Systems data were presented. In the light of the results of the developed prototype and case study, the following conclusions are drawn:

- The proposed method provided a balance between economic use of space and the ability to warn against potential collisions in an effective manner using the pose, state, geometry, and speed characteristics of the equipment;
- The flexibility of the method in using more than one speed vector in the calculation of Dynamic Equipment Workspaces enabled effective capturing of the operation of skilled operators where multiple degrees of freedom can be used simultaneously.

(4) A novel method for Look-Ahead Equipment Workspace for earthwork equipment was proposed that uses the predictive power of Near Real-Time Simulation to evaluate the site safety based on a number of parameters including Shortest Distance to Equipment, Time to Shortest Distance, and Visibility Index. This method enables different pieces of equipment to ensure that their initially planned paths are collision-free, or alternatively adjust their path planning to avoid potential collisions. The following conclusions are drawn:

- The proposed method provided a reliable basis for the generation of the risk maps of earthwork equipment, using the expected pose and state and considering the proximity-based and visibility-based risks;
- The risk maps can be combined to generate Look-Ahead Equipment Workspaces with different risk levels that can be used by different equipment and crews based on the varying levels of risk they can tolerate to adjust their initial paths.

(5) A *MAS* structure was developed to provide the basis for a fleet-level automated equipment guidance system. In the proposed *MAS* structure, each piece of equipment is supported by an operator agent to oversee the task and provide guidance whenever needed. A multi-layer agent hierarchy assigns, monitors and coordinates the task execution, and a set of three types of agents feeds the system with the relevant information. With regard to this contribution the following conclusions are made:

- The proposed framework helped expand the effective domain of *LGS*s from machine-level guidance to fleet-level coordination. In the view of the presented case studies, the *MAS* structure was found to be effective in assigning different operations and tasks of a project to the specific agents that will be responsible for their realization. Using a combination of strategic and tactical planning methods, the *MAS* was able to effectively provide readily executable guidance/control for equipment operators considering a variety of safety issues;
- The proposed *MAS* structure was shown to have a potential for improving the safety of earthwork sites through enabling smooth coordination between a fleet of equipment. Safety issues and conflict-prone activities were addressed using the proposed two-

layer safety mechanism that accounts for a wide range of human factors and uncertainties, e.g., operators' potential non-compliance with the suggested paths;

- While the present research mainly focused on the safety aspect of *LGS*-based fleet coordination, the *MAS* structure is expected to contribute as well to improving the productivity of the operations. Similarly, the proposed method can be leveraged to improve the communication and decision-making at the managerial level through the provision of real-time sensory and simulation data.

9.3 LIMITATIONS AND FUTURE WORK

While this research managed to address its set objectives successfully, the following limitations have been faced during various phases of the research:

(1) While the proposed pose estimation method is very promising in terms of meeting its objective, the current computation costs are high. This has consequences on the real-time aspect of monitoring practices that ensues from the pose estimation. Therefore, it is required to study other optimization algorithms that are capable of solving the presented objective functions faster. Additionally, the future efforts in this section of the research can be directed to:

- Performing sensitivity analysis to study the impact of the weight of the penalty (A) in Equation 4-9 and the tolerance of the Genetic Algorithm on the performance of proposed method;

- Applying the method to other types of equipment and to actual equipment on the construction site;

- Investigating other types of sensory data, e.g., inertial measurement unit, in terms of their efficiency and applicability for accurate pose estimation.

(2) The developed Near Real-Time Simulation approach, as presented in Chapter 5, does not perform distribution fitting to the newly captured data from the site and assumes that the statistical nature of data remains intact in the course of the project. However, this functionality can be added to the *Model Refiner* component so as to enable capturing the variation in the statistical characteristics of various elements of the cycle-time. The future efforts can aim to:

- Develop and implement the Information filter proposed in this approach and automate the scenario analyzer module;
- Add features for enhancing the project safety using updated simulation in short intervals;
- Improve the Model Refiner to enable distribution fitting to the data captured from the site;
- Consider the variability in the operational constraints over the project life-cycle and over the earthwork site;
- Conduct a more comprehensive case study in an actual earthmoving project;
- Extend the functionality of the proposed approach to incorporate the possibility of updating the logic of the model and the sequence of activities, when required.

(3) With regard to the proposed method for the generation of Dynamic Equipment Workspace, some false warnings resulted from capturing the movement along various degrees of freedom only in 2D. Therefore, the future efforts can be dedicated to avoiding this problem by considering the details of the movement in the third dimension in the generation of Dynamic

Equipment Workspaces. Another interesting aspect, to investigate in the future, is the psychological human factors that pertain to the use of Dynamic Equipment Workspace for collision warnings. The study of these factors would help identify the acceptable ratio of false alarms to triggered warnings, which can in turn be used to develop the usability criteria for the evaluation and development of practical systems.

(4) As for the Look-Ahead Equipment Workspaces workspace, the main limitation of the research was that the full integration of the method with pose estimation and Near Real-Time Simulation modules was not performed. At the current stage of the research the data transfer was done manually. This integration can be considered for the future work. Additionally, the following aspects of the proposed method can be studied in the future:

- Applying advanced path planning algorithms to enable the generation of the equipment motion paths in cases where the parametric motion planning method may not work because of the presence of many obstacles;
- Using a probabilistic risk assessment method to consider the uncertainties of the predicted motion paths of the equipment;
- Conducting a comprehensive survey to identify the optimal weights for the three risk factors based on the views of safety managers and equipment operators.

(5) The main limitation of the presented research is that the full-fledged implementation of the proposed framework and complete integration of different methods in near real-time requires a level of computational power which currently may not be economically justified to provide to earthmoving equipment. Having said that, considering the current rate with which the available computational power improves, which is roughly 100% growth rate in computation

speed every 18 month to three years according to Moore's law (Cavin et al. 2012), fully integrated implementation is expected to become viable in the near future. The second limitation of the present research is the challenge of comparing this method with the available commercial fleet management systems (Noland 2010; Jamasmie 2010). Given the proprietary nature of these systems and the fact that there is very little information available on their technical aspects, a comprehensive comparison is not feasible.

(6) While the operational and managerial benefits of the proposed method are well discussed in this thesis, it is necessary to further analyze the economics of applying the multi-agent based LGSs to further demonstrate the advantages of the proposed method over existing off-the-shelf systems.

REFERENCES

AbouRizk, S. (2010). "Role of simulation in construction engineering and management." *J.Constr.Eng.Manage.,* 136(10), 1140-1152.

Acharya, B., Fagerman, J., and Wright, C. (2000). "Accuracy assessment of DTM data: a cost effective approach for a large scale digital mapping project." 33(B2), 105-111.

Akhavian, R., and Behzadan, A. H. (2013). "Knowledge-based simulation modeling of construction fleet operations using multimodal-process data mining." *J.Constr.Eng.Manage.,* 139(11), 04013021.

Akhavian, R., and Behzadan, H.,Amir. (2012). "An integrated data collection and analysis framework for remote monitoring and planning of construction operations." *Advanced Engineering Informatics,* 26(4), 749-761.

Akinci, B., Fischer, M., and Kunz, J. (2002). "Automated generation of work spaces required by construction activities." *J.Constr.Eng.Manage.,* 128(4), 306-315.

AlBahnassi, H., and Hammad, A. (2011). "Near real-time motion planning and simulation of cranes in construction: Framework and system architecture." *J.Comput.Civ.Eng.,* 26(1), 54-63.

Alshibani, A., and Moselhi, O. (2007). "Tracking and forecasting performance of earthmoving operations using GPS data." *In CME 25 Conference on Construction Management and Economics,* 1377-1388.

Alzraiee, H., Moselhi, O., and Zayed, T. (2012). "Dynamic planning of earthmoving projects using system dynamics." *Gerontechnology,* 11(2), 316.

Andayesh, M., and Sadeghpour, F. (2014). "The time dimension in site layout planning." *Autom.Constr.,* 44, 129-139.

ArcGIS. (2015). Retrieved on May 4, 2015 from: <http://www.esri.com/software/arcgis/arcgis-for-desktop>

AutoDesk. (2014). "AutoDesk SoftImage." Retrieved on May 4, 2015 from: <http://www.autodesk.com/products/softimage/overview>

AWCBC. (2013). Association of Workers' Compensation Boards of Canada, National Work Injury/Disease Statistics Program (*NWISP*). Retrieved on May 15, 2015 from: < http://awcbc.org/?wpdmdl=7064 >

Azimi, R., Lee, S., AbouRizk, S. M., and Alvanchi, A. (2011). "A framework for an automated and integrated project monitoring and control system for steel fabrication projects." *Autom.Constr.,* 20(1), 88-97.

Baerwald, J. E., Huber, M. J., and Keefer, L. E. (1965). *Transportation and traffic engineering handbook.* Prentice Hall, Englewood Cliffs, New Jersey.

Barfield, W., and Dingus, T. A. (1998). *Human factors in intelligent transportation systems.* Lawrence Erlbaum Associates, Mahwah, New Jersey.

Beavers, E.,J., Moore, R.,J., Rinehart, R., and Schriver, R.,W. (2006). "Crane-related fatalities in the construction industry." *J.Constr.Eng.Manage.,* 132(9), 901-910.

Bellifemine, F. L., Caire, G., and Greenwood, D. (2007). *Developing multi-agent systems with JADE.* John Wiley & Sons.

BLS. (2012). "Injuries, Illnesses, and Fatalities." *Bureau of Labour Statistics*, Retrieved on May 4, 2015 from: <http://stats.bls.gov/iif/oshcfoi1.htm#2012>

Bonchis, A., Hillier, N., Ryde, J., Duff, E., and Pradalier, C. (2011). "Experiments in autonomous earth moving." *In World Congress,* 18(1), 11588-11593.

Booch, G., Rumbaugh, J., and Jacobson, I. (1999). *The unified modeling language user guide.* Addison-Wesley, Reading, PA.

Borshchev, A., and Filippov, A. (2004). "From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools." *In Proceedings of the 22nd international conference of the system dynamics society*.

Boulanger, G.,D. (1961). "Program evaluation and review technique." *Advanced Management,* 26(7), 8-12.

Bradley, A.,D., and Seward, W.,D. (1998). "The development, control and operation of an autonomous robotic excavator." *Journal of Intelligent and Robotic Systems,* 21(1), 73-97.

M. Bryant. (2006). "Feature: 3D Machine Control: Where does the surveyor fit in?" Retrieved on May 4, 2015 from: <http://archives.profsurv.com/magazine/article.aspx?i=1530 .>

Burns, R. L. (2002). "Dynamic safety envelope for autonomous-vehicle collision avoidance system." *U.S. Patent No. 6,393,362*.

Carbonari, A., Giretti, A., and Naticchia, B. (2011). "A proactive system for real-time safety management in construction sites." *Autom.Constr.,* 20(6), 686-698.

Caterpillar. (2015). "MineStar." Retrieved on May 4, 2015 from: <http://www.cat.com/en_US/support/operations/technology/cat-minestar.html >

Cavin, R. K., Lugli, P., and Zhirnov, V. V. (2012). "Science and engineering beyond Moore's law." *Proceedings of the IEEE, 100(Special Centennial Issue)*, 1720-1749.

CGA. (2012). "DIRT Annual Report for 2012." Common Ground Alliance. Retrieved on May 4, 2015 from: <http://commongroundalliance.com/sites/default/files/DIRT_Report_for_2012_Final_09.10.13_reduced.pdf>

Chae, S. (2009). "Development of warning system for preventing collision accident on construction site." *Proceedings of the 26th Int. Symposium on Automation and Robotics in Construction*.

Chae, S., and Yoshida, T. (2010). "Application of RFID technology to prevention of collision accident with heavy equipment." *Autom.Constr.,* 19(3), 368-374.

Chang, D. (1987). "RESQUE: A resource based simulation system for construction process planning." *Doctoral Dissertation*, University of Michigan, Ann Arbor, Michigan.

Chavada, R., Dawood, N. N., and Kassem, M. (2012). "Construction workspace management: the development and application of a novel nD planning approach and tool." *Journal of Information Technology in Construction (ITcon),* 17, 213-236.

Cheng, T., and Teizer, J. (2013). "Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications." *Autom.Constr.,* 34, 3-15.

Cheng, T. (2013). "Automated safety analysis of construction site activities using spatio-temporal data." *Doctoral Dissertation*, Georgia Institute of Technology.

Chi, S., and Caldas, C. H. (2011). "Image-based safety assessment: automated spatial safety risk identification of earthmoving and surface mining activities." *J.Constr.Eng.Manage.,* 138(3), 341-351.

Chi, S., and Caldas, C. H. (2011). "Image-based safety assessment: automated spatial safety risk identification of earthmoving and surface mining activities." *J.Constr.Eng.Manage.,* 138(3), 341-351.

Choe, S., Leite, F., Seedah, D., and Caldas, C. (2014). "Evaluation of sensing technology for the prevention of backover accidents in construction work zones." *Journal of Information Technology in Construction,* 19, 1-19.

CRC Mining, (2015). "Shovel load assist program." Retrieved on May 4, 2015 from: <http://www.crcmining.com.au/breakthrough-solutions/shovel-load-assist-project/>

Davies, V. J., and Tomasin, K. (1996). *Construction safety handbook.* Thomas Telford.

Dawood, N., and Mallasi, Z. (2006). "Construction workspace planning: assignment and analysis utilizing 4D visualization technologies." *Computer-Aided Civil and Infrastructure Engineering,* 21(7), 498-513.

de Athayde Prata, B., Ferreira Nobre, E., and Cordeiro Barroso, G. (2008). "A stochastic colored petri net model to allocate equipments for earth moving operations." *Journal of Information Technology in Construction (ITcon),* 13, 476-490.

De Berg, M., Van Kreveld, M., Overmars, M., and Schwarzkopf, O. C. (2000). *Computational geometry.* Springer.

De Giovanni, L., and Pezzella, F. (2010). "An improved genetic algorithm for the distributed and flexible job-shop scheduling problem." *Eur.J.Oper.Res.,* 200(2), 395-408.

Dunston, S.,P., and Monty, J. (2009). "Practices for seamless transmission of design data from design phase to construction equipment operation - a synthesis study." Publication FHWA/IN/JTRP-2008/08. *Joint Transportation Research Program*, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.

El-Rayes, K., and Khalafallah, A. (2005). "Trade-off between safety and cost in planning construction site layouts." *J.Constr.Eng.Manage.,* 131(11), 1186-1195.

Fambro, D. B., Fitzpatrick, K., and Koppa, R. J. (1997). *Determination of stopping sight distances.* NCHPR Report 400*,* Transportation Research Board.

Fan, X., and Yen, J. (2007). "R-CAST: Integrating team intelligence for human-centered teamwork." *In Proc. of the national conference on artificial intelligence,* 1535-1541.

Feng, C., Liu, L., and Burns, S. (1997). "Using genetic algorithms to solve construction time-cost trade-off problems." *J.Comput.Civ.Eng.,* 11(3), 184-189.

Ferber, J. (1999). *Multi-agent systems: an introduction to distributed artificial intelligence.* Addison-Wesley, Harlow, UK.

Fishman, S.,G. (2001). *Discrete-event simulation: modeling, programming, and analysis.* Springer.

Forshaw, J., and Smith, G. (2009). *Dynamics and relativity.* John Wiley & Sons.

Gan-Mor, S., Clark L., R., and Upchurch, B. (2007). "Implement lateral position accuracy under RTK-GPS tractor guidance." 59(1), 31-38.

Gazis, D., Herman, R., and Maradudin, A. (1960). "The Problem of the Amber Signal Light in Traffic Flow." *Oper.Res.,* 8(1), 112-132.

Georgiadis, P., Vlachos, D., and Iakovou, E. (2005). "A system dynamics modeling framework for the strategic supply chain management of food chains." 70(3), 351-364.

Ghavami, M., Michael B., L., and Kohno, R. (2004). *Ultra Wideband Signals and Systems in Communication Engineering.* Wiley.

Goldberg, D. E., and Holland, J. H. (1988). "Genetic algorithms and machine learning." *Mach.Learning,* 3(2), 95-99.

Golparvar-Fard, M., Heydarian, A., and Niebles, J. C. (2013). "Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers." *Advanced Engineering Informatics,* 27(4), 652-663.

Google 3D Warehouse. (2015). "Google Sketchup." Retrieved on May 4, 2015 from: < https://3dwarehouse.sketchup.com/?redirect>

Guenther, N., and Salow, H. (2012). "Collision avoidance and operator guidance innovating mine vehicle safety." *Queensland Mining Industry Health & Safety Conference,* Townsville.

Guy, P. (2011). *Personal Communication*. Belitech Inc., Email correspondence on September 2011.

Hajjar, D., and AbouRizk, S. (1999). "Simphony: An environment for building special purpose construction simulation tools." *Processings of 1999 Winter Simulation Conference,* 998-1006.

Halder, B., and Vitale, A. J. (2010). "Autonomous Machine Control System." *U.S. Patent Application 12/902, 612.*

Halpin, D. (1977). "CYCLONE-method for modeling job site processes." *Journal of the Construction Division,* 103, 489-499.

Halpin, D., and Riggs, S.,L. (1992). *Planning and analysis of construction operations.* Wiley, New York.

Halpin, D. W., and Martinez, L. H. (1999). "Real world applications of construction process simulation." *Proceeding of the 1999 Winter Simulation Conference: Simulation---a bridge to the future,* 956.

Hammad, A., El Ammari, K., Langari, S., Vahdatikhaki, F., Soltani, M., AlBahnassi, H., and Paes, B. (2014). "Simulating macro and micro path planning of excavation operations using game engine." *Proceedings of the 2014 Winter Simulation Conference*.

Hammad, A., and Zhang, C. (2011). "Towards real-time simulation of construction activities considering spatio-temporal resolution requirements for improving safety and productivity." *Winter Simulation Conference,* Phoenix.

Hammad, A., Zhang, C., Al-Hussein, M., and Cardinal, G. (2007). "Equipment workspace analysis in infrastructure projects." *Canadian Journal of Civil Engineering,* 34(10), 1247-1256.

Hartigan, J., and Wong, A.,Manchek. (1979). "Algorithm AS 136: A k-means clustering algorithm." *Applied Statistics,* 100-108.

Hinze, W.,J., and Teizer, J. (2011). "Visibility-related fatalities related to construction equipment." *Saf.Sci.,* 709-718.

Hinze, J., Huang, X., and Terry, L. (2005). "The nature of struck-by accidents." *J.Constr.Eng.Manage.,* 131(2), 262-268.

Howarth, T., and Watson, P. (2009). *Construction Safety Management.* Wiley-Blackwell.

Huhns, M., and Stephens, L. (1999). "Multiagent systems and societies of agents." *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence,* MIT Press.

Hukkeri, R., B. (2012). "Machine control system implementing intention mapping." *U.S. Patent Application 12/951, 848*.

Jackson, T. (2008). "Technology: GPS by the numbers." *Equipment World*, Retrieved on May 4, 2015 from:
< www.equipmentworld.com/technology-gps-by-the-numbers/>

Jahren, C. (2013). "Implementation: 3D engineered models for highways." *Report submitted to Federal Highway Administration*, Iowa State University.

Jamasmie, C. (2010). "The mine of the future might be a thing of the past." Retrieved on May 5, 2015from:
<http://MINING.com>

Kalk, A. (1978). "INSIGHT: Interactive simulation of construction operations using graphical techniques." *Doctoral dissertation,* Department of Civil Engineering, Stanford University.

Kamat, V. R., and Martinez, J. C. (2005). "Dynamic 3D visualization of articulated construction equipment." *J.Comput.Civ.Eng.,* 19(4), 356-368.

Kaming, P. F., Holt, G. D., Kometa, S. T., and Olomolaiye, P. O. (1998). "Severity diagnosis of productivity problems—a reliability analysis." *Int.J.Project Manage.,* 16(2), 107-113.

Kaufmann, K., and Anderegg, R. (2008). "GPS-based compaction technology." *Proceedings of the 1th International Conference on Machine Control & Guidance,* 24-26.

Kelley, J.,J., and Walker, R.,M. (1959). "Critical-path planning and scheduling." *In Papers presented at Eastern Joint IRE-AIEE-ACM Computer Conference,* 160-173.

Kim, C., Haas, C. T., Liapi, K. A., and Caldas, C. H. (2006). "Human-assisted obstacle avoidance system using 3D workspace modeling for construction equipment operation." *J.Comput.Civ.Eng.,* 20(3), 177-186.

Kim, S., Seo, J., and Russell, J. S. (2012). "Intelligent navigation strategies for an automated earthwork system." *Autom.Constr.,* 21, 132-147.

Kim, S. K., and Russell, J. S. (2003a). "Framework for an intelligent earthwork system: Part I. System architecture." *Autom.Constr.,* 12(1), 1-13.

Kim, S. K., and Russell, J. S. (2003b). "Framework for an intelligent earthwork system: Part II. Task identification/scheduling and resource allocation methodology." *Autom.Constr.,* 12(1), 15-27.

Kiongoli, S. (2010). "Testing the accuracy of machine guidance in road construction." *Undergraduate dissertation*, University of Southern Queensland.

Kolodziej, W.,Krzysztof, and Hjelm, J. (2006). *Local positioning systems: LBS applications and services.* CRC Press.

Kumler, M. (1994). "An intensive comparison of Triangulated Irregular Networks (TINs) and Digital Elevation Models (DEMs)." *The International Journal for Geographic Information and Geovisualization,* 31(2), 1-99.

Lambropoulos, S., Manolopoulos, N., and Pantouvakis P., J. (1996). "SEMANTIC: Smart earthmoving analysis and estimation of cost." *Construction Management & Economics,* 14(2), 79-92.

Lee, J., and Bernold, L. E. (2008). "Ubiquitous agent-based communication in construction." *J.Comput.Civ.Eng.,* 22(1), 31-39.

Leica. (2015). "Leica Jigsaw." Retrieved on May 5, 2015 from: <http://mining.leica-geosystems.com/assets/Secure/Brochure/Leica-Jigsaw-J3autonomous.pdf >

Leica Geosystems. (2015). Retrieved on May 5, 2015 from: <http://www.leica-geosystems.us/>

Li, H., and Love, P. (1997). "Using improved genetic algorithms to facilitate time-cost optimization." *J.Constr.Eng.Manage.,* 123(3), 233-237.

Lluch, J., and Halpin, W.,D. (1982). "Construction operation and microcomputers." *Journal of the Construction Division,* 108(1), 129-145.

Lu, M., Dai, F., and Chen, W. (2007). "Real-time decision support for planning concrete plant operations enabled by integrating vehicle tracking technology, simulation, and optimization algorithms." *Canadian Journal of Civil Engineering,* 912-922.

Luo, X., O'Brien, W. J., Leite, F., and Goulet, J. A. (2014). "Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks." *Advanced Engineering Informatics,* 28(4), 287-296.

Maalek, R., and Sadeghpour, F. (2013). "Accuracy assessment of Ultra-Wide Band technology in tracking static resources in indoor construction scenarios." *Autom.Constr.,* 30, 170-183.

MacCollum, D. V. (1995). *Construction safety planning.* John Wiley & Sons.

Makkonen, T., Nevala, K., and Heikkilä, R. (2006). "A 3D model based control of an excavator." *Autom.Constr.,* 15(5), 571-577.

Mallasi, Z., and Dawood, N. (2004). "Workspace competition: assignment, and quantification utilizing 4D visualization tools." *Proceeding of Conference on Construction Application of Virtual Reality,* 13-22.

Mallasi, Z., and Dawood, N. (2001). "Assessing space criticality in sequencing and identifying execution patterns for construction activities using VR visualisations." *In ARCOM doctoral research workshop: Simulation and modelling in construction,* 22-27.

Marks, E. (2014). "Active safety leading indicators for human-equipment interaction on construction sites." *Doctoral dissertation*, Georgia Institute of Technology.

Martinez, J. (1996). "STROBOSCOPE State and Resource Based Simulation of Construction Processes." *Doctoral dissertation,* University of Michigan.

Martinez, J. C. (2001). "EZStrobe: general-purpose simulation system based on activity cycle diagrams." *In Proceedings of the 33nd conference on Winter simulation,* IEEE Computer Society, 1556-1564.

Martinez, J. C., and Ioannou, P. G. (1999). "General-purpose systems for effective construction simulation." *J.Constr.Eng.Manage.,* 125(4), 265-276.

Martinez, J. (2010). "Methodology for conducting discrete-event simulation studies in construction engineering and management." *J.Constr.Eng.Manage.*136(1), 3-16

Marzouk, M., and Ali, H. (2013). "Modeling safety considerations and space limitations in piling operations using agent based simulation." *Expert Syst.Appl.,* 40(15), 4848-4857.

Mathworks. (2015). "MATLAB." Retrieved on May 5, 2015 from: <http://www.mathworks.com/help/curvefit/sum-of-sine.html>

McCann, M. (2006). "Heavy equipment and truck-related deaths on excavation work sites." *J.Saf.Res.,* 37(5), 511-517.

Mintah, B., Price, R. J., and King, K. D. (2011). "Semi-autonomous excavation control system" *U.S. Patent No. 7,934,329.* Washington, DC: U.S. Patent and Trademark Office.

Montaser, A., and Moselhi, O. (2012). "RFID for tracking earthmoving operations." *Construction Research Congress,* 1011-1020.

Moon, H., Kim, H., Kim, C., and Kang, L. (2013). "Development of a schedule-workspace interference management system simultaneously considering the overlap level of parallel schedules and workspaces." *Autom.Constr.,* 39, 93-105.

Motamedi, A., and Hammad, A. (2009). "Lifecycle management of facilities components using radio frequency identification and building information model." *Journal of Information Technology in Construction (ITcon),* 14, 238-262.

Navon, R. (2007). "Research in automated measurement of project performance indicators." *Autom.Constr.,* 16(2), 176-188.

Navon, R., Goldschmidt, E., and Shpatnisky, Y. (2004). "A concept proving prototype of automated earthmoving control." *Autom.Constr.,* 13(2), 225-239.

Navon, R., and Sacks, R. (2007). "Assessing research issues in automated project performance control." *Autom.Constr.,* 16(4), 474-484.

Noland, R. (2010). "Site connectivity - positioning & machine control's next technology step, machine control online." Retrieved on May 5, 2015 from: <http://machinecontrolonline.com/content/view/6390/136/>

Ogunlana, O.,S., Li, H., and Sukhera, A.,F. (2003). "System dynamics approach to exploring performance enhancement in a construction organization." *J.Constr.Eng.Manage.,* 129(5), 528-536.

OHS. (2011)."Occupational Injuries and Diseases in Alberta: Construction and Construction Trade Services Industries." *Government of Alberta and Work Safe Alberta*, Retrieved on May 15, 2015 from: < http://work.alberta.ca/documents/OID-construction-industries.pdf >

Oloufa, A., Ikeda, M., and Oda, H. (2003). "GPS-based wireless collision detection of construction equipment." *NIST Special Publication,* 461-466.

Ören, I.,T., Numrich, S., Uhrmacher, M.,A., Wilson, F.,L., and Gelenbe, E. (2000). "Agent-directed simulation: challenges to meet defense and civilian requirements." *In Proceedings of the 32nd conference on Winter simulation,* 1757-1762.

OSHA. (2011). "Construction Focus Four: Caught-In or -Between Hazards." OSHA Directorate of Training and Education.

Peralta, R. C., Forghani, A., and Fayad, H. (2014). "Multiobjective genetic algorithm conjunctive use optimization for production, cost, and energy with dynamic return flow." *Journal of Hydrology,* 511, 776-785.

Perkinson, L.,C., Bayraktar, E.,M., and Ahmad, I. (2010). "The use of computing technology in highway construction as a total jobsite management tool." *Autom.Constr.,* 19(7), 884-897.

Petschko, H. (2008). "Universal developer platform for machine control applications." *1 st International Conference on Machine Control & Guidance,* 257-262.

Peurifoy, L.,R., Schexnayder, J.,C., Shapira, A., and Schmitt, L.,R. (2011). *Construction planning, equipment, and methods.* McGraw-Hill, New York.

Peyret, F., Jurasz, J., Carrel, A., Zekri, E., and Gorham, B. (2000). "The computer integrated road construction project." *Autom.Constr.,* 9(5-6), 447-461.

Pilone, D., and Pitman, N. (2005). *UML 2.0 in a Nutshell.* " O'Reilly Media, Inc.

Pitt, J., and Bellifemine, F. (1999). "A protocol-based semantics for FIPA97 ACL and its implementation in JADE." *AI* IA*.

Pradhananga, N. (2014). "Construction site safety analysis for human-equipment interaction using spatio-temporal data." *Doctoral dissertation*, Georgia Institute of Technology.

Pradhananga, N., and Teizer, J. (2013). "Automatic spatio-temporal analysis of construction site equipment operations using GPS data." *Autom.Constr.,* 29, 107-122.

Rebolj, D., Babič, N. Č, Magdič, A., and Podbreznik, P. (2008). "Automated construction activity monitoring system." *Advanced Engineering Informatics,* 22(4), 493-503.

Reclus, F., and Drouard, K. (2009). "Geofencing for fleet & freight management." *In Intelligent Transport Systems Telecommunications,(ITST), 2009 9th International Conference,* 353-356.

Ren, Z., and Anumba, C. J. (2002). "Learning in multi-agent systems: a case study of construction claims negotiation." *Advanced Engineering Informatics,* 16(4), 265-275.

Rezazadeh Azar, E., Feng, C., Kamat, V. R. (2015). "feasibility of in-plane articulation monitoring of excavator arm using planar marker tracking, *Journal of Information Technology in Construction (ITcon)*, 20, 213-229.

Rezazadeh Azar, E., and McCabe, B. (2011). "Automated visual recognition of dump trucks in construction videos." *J.Comput.Civ.Eng.,* 26(6), 769-781.

Rojas, M.,E., and Mukherjee, A. (2006). "Multi-agent framework for general-purpose situational simulations in the construction management domain." *J.Comput.Civ.Eng.,* 20(3), 165-176.

Rojas, J. A., and Giachetti, R. E. (2009). "An agent-based simulation model to analyze team performance on jobs with a stochastic structure." *SIMUL'09. First International Conference on,* Advances in System Simulation, 148-154.

Rowe, P., and Stentz, A. (1997). "Parameterized scripts for motion planning." *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems,* IEEE, 1119-1124.

Rowe, P. S. (1999). "Adaptive motion planning for autonomous mass excavation." *Doctoral dissertation*, Carnegie Mellon University.

Ruff, T. (2006). "Evaluation of a radar-based proximity warning system for off-highway dump trucks." *Accident Analysis & Prevention,* 38(1), 92-98.

Russell, J. S., and Kim, S. K. (2003). "Chapter 6." *Construction Automation,* CRC Press, .

Russell, S., and Norvig, P. (2003). *Artificial Intelligence, A Modern Approach,* Prentice Hall, Englewood Cliffs, NJ.

SAFEmine. (2014). "Advanced traffic safety solution for surface mining." Retrieved on May 5, 2015 from:
< http://www.safe-mine.com/index.php/products#cas>

Saito, H., Sugiura, H., and Yuta, S. (1995). "Development of autonomous dump trucks system (HIVACS) in heavy construction sites." *Proceedings of 1995 IEEE International Conference on Robotics and Automation,* Nagoya, 2524-2529.

Sarata, S., Koyachi, N., Tsubouchi, T., Osumi, H., Kurisu, M., and Sugawara, K. (2006). "Development of autonomous system for loading operation by wheel loader." *In Proceedings of the 23rd International Symposium on Automation and Robotics in Construction, ISARC,* 466-471.

Sawhney, A., Bashford, H., Walsh, K., and Mulky R., A. (2003). "Agent-based modeling and simulation in construction." *Proceedings of the 2003 Winter Simulation,* 1541-1547.

Siddiqui, H. (2014). "UWB RTLS for construction equipment localization: experimental performance analysis and fusion with video data." *Maters dissertation,* Concordia University

Schaub, H., and Junkins, J. L. (2003). *Analytical Mechanics of Space Systems.* Aiaa.

Seo, J., Lee, S., Kim, J., and Kim, S. (2011). "Task planner design for an automated excavation system." *Autom.Constr.,* 20(7), 954-966.

Setayeshgar, S., Hammad, A., Zhang, C., and Vahdatikhaki, F. (2013). "Real time safety analysis of construction projects using BIM and RTLS." *CSCE 2013 Annual Conference,* Montreal.

Shi, J.,J. (1999). "Computer simulation in AEC and its future development." *In Berkeley-Stanford CE&M Workshop,* Stanford.

Shi, J., and AbouRizk S., S. (1998). "An automated modeling system for simulating earthmoving operations." *Computer-Aided Civil and Infrastructure Engineering,* 13(2), 121-130.

Singh, R. (2010). "Construction automation: A six year plan in development." Retrieved on May 5, 2015 from: <http://www.oregon.gov/ODOT/HWY/GEOMETRONICS/docs/dozer/design_to_dozer_engineering_automation_mod.pps>

Singh, S. (1995). "Synthesis of tactical plans for robotic excavation." *Doctoral dissertation*, Carnegie Mellon University.

Singh, S., and Cannon, H. (1998). "Multi-resolution planning for earthmoving." *In Proceedings of 1998 IEEE International Conference on Robotics and Automation,* 121-126.

Smith, D.,S., Osborne, R.,J., and Forde, C.,M. (1996). "The use of a discrete-event simulation model with Erlang probability distributions in the estimation of earthmoving production." *Civ.Eng.Syst.,* 13(1), 25-44.

Soltani, A. R., Tawfik, H., Goulermas, J. Y., and Fernando, T. (2002). "Path planning in construction sites: performance evaluation of the Dijkstra, A∗, and GA search algorithms." *Advanced Engineering Informatics,* 16(4), 291-303.

Song, L., and Eldin, N.,Neil. (2012). "Adaptive real-time tracking and simulation of heavy construction operations for look-ahead scheduling." *Autom.Constr.,* 27, 32-39.

Stanfords Business Mapping. (2015). Retrieved on May 5, 2015 from: <http://www.stanfords.co.uk/Business-Mapping/BM/Products/LiDAR_BM-LIDAR.htm>

Stentz, A., Bares, J., Singh, S., and Rowe, P. (1999). "A robotic excavator for autonomous truck loading." *Autonomous Robots,* 7(2), 175-186.

Sterman, J. (2000). *Business dynamics.* Irwin-McGraw-Hill.

Sturm, A., and Vos, W. (2008). "New technologies for telematics and machine control." *1st International Conference on Machine Control & Guidance,* 189-198.

Sweetser, A. (1999). "A comparison of system dynamics (SD) and discrete event simulation (DES)." *17th International Conference of the System Dynamics Society,* 20-23.

Szczesny, K., König, M., Laußat, L., and Helmus, M. (2013). "Integration of uncertain real-time logistics data for reactive scheduling using fuzzy set theory." *Proceedings of the 30th*

International Symposium on Automation and Robatics in Construction and Mining (ISARC), Montreal.

Talmaki, S., and Kamat, R.,V. (2014). "Real-time hybrid virtuality for prevention of excavation related utility strikes." *J.Comput.Civ.Eng.,* 28(3).

Tambe, M. (1997). "Towards Flexible Teamwork." *Journal of Artificial Intelligence Research,* 7, 83-124.

Taoka T., G. (1989). "brake reaction times of unalerted drivers." *ITE J.,* 59(3), 19-21.

Tatum B., C., and Funke T., A. (1988). "Partially automated grading: construction process innovation." *J.Constr.Eng.Manage.,* 114(1), 19-35.

Teizer, J., Allread S., B., Fullerton E., C., and Hinze, J. (2010b). "Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system." *Autom.Constr.,* 19(5), 630-640.

Teizer, J., Allread, B. S., and Mantripragada, U. (2010a). "Automating the blind spot measurement of construction equipment." *Autom.Constr.,* 19(4), 491-501.

Teizer, J., Venugopal, M., and Walia, A. (2008). "Ultrawideband for automated real-time three-dimensional location sensing for workforce, equipment, and material positioning and tracking." *Transportation Research Record: Journal of the Transportation Research Board,* 2081(1), 56-64.

Topcon. (2015). "Machine Control." Retrieved on May 5, 2015 from: <http://www.topconpositioning.com/products/machine-control>

Transports Quebec. (2014). "Projet de reconstruction du complexe Turcot à Montréal, Montréal-Ouest et Westmount." Retrieved on May 5, 2015 from: <http://www.bape.gouv.qc.ca/sections/mandats/Complexe_Turcot/documents/liste_cotes.htm>

Trimble. (2015). "Machine Control." Retrieved on May 5, 2015 from: <http://construction.trimble.com/products/machine-control>

Turing, A. M. (1950). "Computing machinery and intelligence." *Mind,* 236, 433-460.

Ubisense. (2015). Retrieved on May 5, 2015 from: <http://ubisense.net/en>

Uddin, W. (2002). "Evaluation of airborne LiDAR digital terrain mapping for highway corridor planning and design." 15, 10-15.

Unity 3D. (2015). "Unity - Game Engine." Retrieved on May 5, 2015 from: <http://unity3d.com/>

Vahdatikhaki, F., and Hammad, A. (2014). "Framework for near real-time simulation of earthmoving projects using location tracking technologies." *Autom.Constr.,* 42, 50–67.

Van Sickle, J. (2008). "e-Education institute." Retrieved on May 5, 2015 from: <https://courseware.e-education.psu.edu/downloads/geog862/Lesson7/index.htm>

Van Tol, A.,A., and AbouRizk, M.,S. (2006). "Simulation modeling decision support through belief networks." *Simulation Modelling Practice and Theory,* 14(5), 614-640.

Veneziano, D., Hallmark, S., and Souleyrette, R. (2002). "Accuracy evaluation of lidar-derived terrain data for highway location." *Center for Transportation Research and Education*, Iowa State University.

Viljamaa, E., and Peltomaa, I. (2014). "Intensified construction process control using information integration." *Autom.Constr.,* 39, 126-133.

Vonderohe, P.,A. (2007). "Implementation of GPS controlled highway construction equipment." Construction Materials and Support Center; University of Wisconsin.

Vonderohe, P.,A., and Hintz, C. (2010). "3D design terrain models for construction plans and GPS control of highway construction equipment." Construction Materials and Support Center; University of Wisconsin.

Vonderohe, A. (2009). "Training on automated machine guidance." National Center for Freight & Infrastructure Research & Education.

Vonderohe, A., Brockman, K., Whited, G., and Zogg, J. (2009). "Development of a specification for gps-machine-guided construction of highway subgrade." *88th Annual Meeting of the Transportation Research Board,* Washington, DC.

Wambeke, B., Liu, M., and Hsiang M., S. (2010). "Causes of variation in construction task starting times and duration." *In Processing of Construction Research Congress, 980-989*.

Wang, J., and Razavi, S. N. (2015). "Low False Alarm Rate Model for Unsafe-Proximity Detection in Construction." *J.Comput.Civ.Eng*. (Available online)

Weiss, G. (1999). *Multiagent systems: a modern approach to distributed artificial intelligence.* The MIT press.

Worrall, S., and Nebot, E. (2008). "A probabilistic method for detecting impending vehicle interactions." *Robotics and Automation. ICRA 2008. IEEE International Conference on,* IEEE, Pasadena, CA, 1787-1791.

Wortman, R. H., and Matthias, J. S. (1985). "Evaluation of driver behavior at signalized intersections." *Transportation Research Record TRR,* 1027, 20-30.

Wu, H., Tao, J., Li, X., Chi, X., Li, H., Hua, X., Yang, R., Wang, S., and Chen, N. (2013). "A location based service approach for collision warning systems in concrete dam construction." *Saf.Sci.,* 51(1), 338-346.

Wu, W., Yang, H., Chew, D. A., Yang, S., Gibb, A. G., and Li, Q. (2010). "Towards an autonomous real-time tracking system of near-miss accidents on construction sites." *Autom.Constr.,* 19(2), 134-141.

Yabuki, N., (2011). "3D data for machine control & machine guidance in construction." *Asia Construction Information Technology Round Table Meeting*, Seoul, Korea.

Yang, H., Chew, D. A., Wu, W., Zhou, Z., and Li, Q. (2012). "Design and implementation of an identification system in construction site safety for proactive accident prevention." *Accident Analysis & Prevention,* 48, 193-203.

Yen, J., Fan, X., Sun, S., Hanratty, T., and Dumer, J. (2006). "Agents with shared mental models for enhancing team decision makings." *Decis.Support Syst.,* 41(3), 634-653.

Zayed, T.,M., and Halpin, D. (2001). "Simulation of concrete batch plant production." *J.Constr.Eng.Manage.,* 127(2), 132-141.

Zhang, C. (2010). "Improving crane safety by agent-based dynamic motion planning using UWB real-time location system." Doctoral Dissertation, Concordia University, Montreal, Canada.

Zhang, C., and Hammad, A. (2012). "Multi-agent approach for real-time collision avoidance and path re-planning on construction sites." *J.Comput.Civ.Eng.,* 26(6), 782-794.

Zhang, C., Hammad, A., and Rodriguez, S. (2012). "Crane pose estimation using UWB real-time location system." *J.Comput.Civ.Eng.,* 26(5), 625-637.

Zolynski, G., Schmidt, D., and Berns, K. (2014). "Safety for an autonomous bucket excavator during typical landscaping tasks." *New Trends in Medical and Service Robots,* Springer, 357-368.

APPENDICES

APPENDIX A: ADMINISTRATIVE INFRASTRUCTURE FOR AMC/G

In order to take full advantage of *LGS*s, it is important for project owners to develop specifications for the use of this technology, and to require the contractors to follow those specifications. The specifications must cover: accuracy limits, quality control and quality assurance (QC/QA) procedures, risks allocation for errors, and payment mechanisms. Furthermore, there is a need to develop an implementation strategy for adopting this technology in construction projects. Vonderohe (2007, 2009) and Dunston and Montey (2009) have reviewed the implemented measures and adopted practices of various Departments of Transportation (DOTs) in the US for accommodating *AMC/G*-friendly administrative infrastructure. These reports cover the types of *AMC/G* equipment allowed by *DOT*s and address the subsequent liability and financial issues. It could be construed from these reports that the terms *AMC/G* and *AMG* are used interchangeably. However, the reports provide evidences suggesting that the *DOT*s are moving towards *AMC*, particularly with their numerous references to grading operations. Table A-1 delineates the criteria which determine the variance in the practice of different *DOT*s.

Table A-1: Administrative Criteria for AMC/G Implementation

Technology	Can AMC/G be implemented? Can other new technologies be utilized?
Documentations and Responsibilities	What documents does the transportation agency provide to the contractor? Who is responsible for the development of 3D design models? What documents does the contractor need to provide to the transportation agency? Does the contractor need to provide training on AMC/G equipment?
Liabilities	Who is liable for the errors in the 3D models? Who is bearing the costs for the errors and subsequent corrections?
Control	How the quality of the project is controlled? Who is providing the control points?

Table A-2 and Figure A-1 summarize the state of the practice and the administrative adjustments made towards streamlining the application of *AMC/G* in construction projects based on the abovementioned reports (Dunston and Montey 2009; Vonderohe 2007). Table A-2 shows different major measures towards formalization of AMC/G and the DOTs that have adopted those measures.

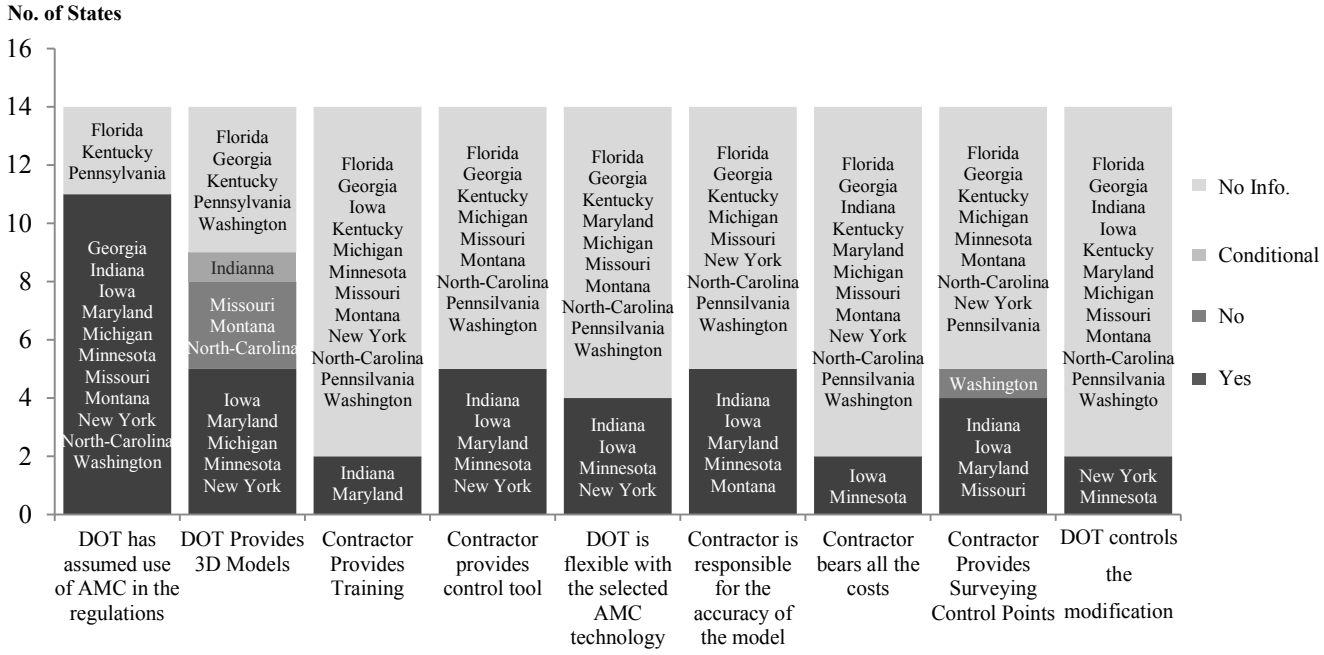


Figure A-1: State of the Practice in Implementing AMC/G at U.S. Dots (Adapted from Vonderode 2007, Vonderode 2009)

Table A-2 categorizes the *DOT*s' state of the practice into used technologies, obligations of the contractor, documents and data format provided by *DOT*s, liability issues regarding error correction and data adjustment, incurred costs and *QA/QC*. In Canada, according to several telephone interviews with most of the provincial ministries of transportation, including New Brunswick, Ontario, Saskatchewan, Quebec and British Columbia, although *AMC/G* is starting to be appreciated as a promising technology with palpable gains, there is a need for major actions towards the regularization of *AMC/G* in transportation projects. The details for the interviewed parties can be found in Table A-3 and Figure A-2.

Table A-2: State of The Practice in Implementing AMC/G at U.S. DOTs (Adapted from Vonderode 2007, Vonderode 2009)

State	Used Technologies	Contractor's Obligations	Data Provision by DOT	Liability w.r.t. Data Accuracy and Adjustment	Costs	QC/QA
Florida	Designers can submit files using available formats	Not Available	Not Available	Not Available	Not Available	Not Available
Georgia	GPS AMC Supplementary laser guidance when required	Not Available	DOT provides electronic digital files in their standard format	Not Available	Not Available	Not Available
Indiana	GPS and Robotic Total Station AMC/G All new technologies upon review and approval	Development of required 3D models 8 hours of training on GPS/RTS and AMC/G system Provision of GPS rover, professional surveying services and grading plan	DOT provides 3D model for AMC/G if available	DOT is not responsible for the accuracy of provided 3D models	Not Available	Contractor provides adequate control point,
Iowa	Mandatory use of AMC/G for designated areas Unspecified types of AMC	AMC/G equipment acquisition Transformation of provided data to compatible formats Provision of GPS rover Daily inspection of AMC/G equipment Uncompromised accuracy Submission of a AMC/G work plan	Engineer provides localized coordinate system DOT provides CAD files, AMC/G DTM and alignment data	Contractor is responsible for the accuracy of provided data Contractor modifies provided electronic data Contractor assumes the risk of error when information is used out of intended context	Contractor bears all the costs The bid for AMC/G grading is measured and paid for at the lump sum contact price	Engineer sets initial control points Contractor provides control points and stakes at critical points, sets hubs at the required points and preserves all the reference points and monuments
Kentucky	Not Available	Requires certain EDFs with the Final plans	Not Available	Not Available	Not Available	Not Available
Maryland	GPS and RTS AMC/G	Provision of GPS rover 8 hours of training on GPS/RTS and the AMC/G system Provision of a surveyor to perform verification Submission of DTM to DOT Demonstrate the accuracy of the AMC/G system to DOT	DOT provides contract documents and DTM data	Contractor corrects all errors and adjusts the data to the satisfaction of DOT	Not Available	Contractor provides adequate control point, stationing and stakes and sets initial control points Contractor provides control points and stakes at critical points
Michigan	GPS AMC	Not Available	DOT provides design files for AMC	Not Available	Not Available	Not Available
Minnesota	GPS and RTS AMC/G All new technologies upon review and approval	Provision of RTS to DOT for control Notification about utilized AMC/G equipment within 15 days after award of the contract	DOT will provide 2D and 3D or DTM files upon contract approval	No guarantee for the accuracy of the provided data by DOT Change of the model is performed by DOT	DOT does not pay for the correction of errors No direct payment is made for AMC	Not Available
Missouri	GPS AMC	Development of required 3D models	Electronic digital files of plans and profiles needed for 3D models is provided by DOT	Not Available	Not Available	DOT has contractor staking procedure Check elevations are based on project plans
Montana	GPS AMC Supplementary laser guidance when required	Development of required 3D models Quality control on AMC/G use	DOT provides paper plans	No checking of the models by DOT	Not Available	Filed checks by radial survey DOT has quality specification on final product
New York	GPS AMC All new technologies upon review and approval	Mandatory use of the DOT's CAAD software Provision of Detailed contract control plan Provision of GPS rover	DOT provides 3D model for AMC	DOT has specifications on use of the model DOT has to approve any modification of the model	Not Available	Checks on originally-measured points Contractor and DOT use same documents for control
North Carolina	GPS AMC	Contractor develops 3D models from paper plans	DOT provides paper plans	Not Available	Not Available	DOT requires full staking
Pennsylvania	DOT has formulated some addendum to their publications	Not Available	Not Available	Not Available	Not Available	Not Available
Washington	GPS AMC	Not Available	Not Available	Not Available	Not Available	DOT provides surveying control Construction staking is done by contractor

Table A-3: List of Contacted Experts at Various Transportation Agencies in Canada

Organization	Contact person
Info Excavation	Marc Jr. Colas, Technical Manager 514-331-0636
Hydro Quebec Research Institute	Robotic Regis Houde Head of Robotics 450 652 8107
MTQ	Odile Béland Technologies de l'information 418 643-4431
Belitec (Trimble)	Guy Perron 819-373-3880
Ontario Ministry of Transportation	Queen's Park/Minister's Office 77 Wellesley Street West Ferguson Block, 3rd Floor Toronto, ON M7A 1Z8 416-327-9200
Ontario Ministry of Transportation	Jeff Baker, Investment Planning Section, at (905) 704-2628, or at Jeff.Baker@mto.gov.on.ca
Ontario Ministry of Transportation	Tony Tuinstra, P.Eng., PQS Construction Engineer Contract Management Office Ministry of Transportation 2nd Floor, 301 St. Paul St. St. Catharines, ON L2R 7R4 P: 905-704-2197 F: 905-704-2040 E: Tony.Tuinstra@ontario.ca
Alberta Ministry of Transportation	Moh Lali Ex Dir Tech Standards 780-415-1083
British Colombia Ministry of Transportation	Keith Callander Director of provincial field services 256-828-4551 256-828-4151 Keith.callander@gov.bc.ca Ian Pilkington Director of rehabilitation & maintenance 250-387-7627

ALBERTA TRANSPORTATION
DESIGN BULLETIN #60

AMENDMENT TO ENGINEERING CONSULTANT GUIDELINES RE: SURVEY REQUIREMENTS TO SUPPORT GPS MACHINE CONTROL CONSTRUCTION

Introduction

This bulletin supplements information contained in Volumes 1 and 2 of the Engineering Consultant Guidelines for Highway and Bridge Projects in regards to providing survey and design information to the contractor on grade construction projects.

In recent years there has been a move by the contracting industry to use GPS machine control equipment whereas the operation of the equipment is partially controlled through GPS technology and computer software using electronic survey and design data. This technology has been widely reported as improving the efficiency of the contractors operations by minimizing the amount of re-grading required to meet specified grades and dimensions.

In addition to the requirements outlined in section 1.7.2 Survey Information (Volume 2 Construction Contract Administration) the following information, and when requested by the Contractor, is to be provided In order to facilitate GPS Machine Control Construction (GPS MCC).

It is recognized that not all types of grade construction (e.g. grade widening, intersection treatments) is conducive to using GPS MCC and there will be instances where the electronic survey and design data is not available in a suitable format for all portions of the project. The intent behind this bulletin is to have consultants enter into the preliminary survey and design phases with the goal to have good quality electronic design data which could be provided to the Contractor in support of GPS MCC.

Preliminary Survey

Original ground coordinates shall consist of cross section pick up at a maximum interval of 20 m stations with breakline information and additional pick-up collected at any ground irregularities (i.e. hills or valleys).

Construction Control Points

Construction control points as described in Section 6 Preliminary Engineering (Volume 1 Design and Tender) shall be selected to surround the project and encompass the construction limits (i.e. both sides of highway along the entire length) including borrow sources. Control point data is to be provided to the Contractor in P,N,E,Z,D file format with the intent that the same control points will be used during construction by both the Consultant and Contractor.

If additional control points or other design related information is required by the Contractor, it will be the Contractor's responsibility to obtain, however this does not preclude the Consultant in offering assistance depending upon the level of effort required and whether there is a benefit to the Department in so providing.

Figure A-2: Document Representing the AMC/G State in the Province of Alberta, Canada

ALBERTA TRANSPORTATION
DESIGN BULLETIN #60

Electronic Design Files

Electronic design files shall meet the following requirements.

- Electronic survey data to be provided as top of subgrade.
- Designs should be presented as a complete and continuous surface.
- The preferred file format is .xml however other formats (e.g. .tin) are acceptable if compatible with GPS MCC.
- For linear work (i.e. roadways or canals) the data should be presented in road files or template format.
- For non linear work (ponds, etc.) the design is best presented as a Digital Terrain Model (DTM) with a line work file defining all critical elements.

Survey Information on Contract Plans

Contract plans should include the following survey information.

- A coordinate list for horizontal alignment with northings and eastings for all PI locations
- Complete curve data
- A coordinate list for control points with design assumptions used (i.e. ground versus grid, scale factors, etc.)

Special Provisions

For grading projects the Consultant shall include special provisions detailing what portions of the work has survey and electronic design data that is suitable for GPS MCC according to the requirements outlined in this bulletin. Questions regarding this bulletin may be directed to Jim Gavin, Roadway Construction Standards Specialist @ 780-415-1008.

Recommended by:

Chuck McMillan
Director,
Surface Engineering and Aggregates
Technical Standards Branch

Reviewed by:

Ranjit Tharmalingam
Director, Professional Services

Approved by:

John Engleder
Executive Director
Program Management Branch

Figure A-2: Document Representing the AMC/G State in the Province of Alberta, Canada (cont.)

As shown in Table A-2 and Figure A-1 some *DOT*s are more advanced in terms of transition to 3D design models, and thus they provide a more *AMC/G* friendly workflow environment. For instance, *DOT*s of Indiana, Iowa, Minnesota and New York have managed to secure a level of flexibility which allows contractors to go beyond the recognized technologies and venture new methods. Iowa *DOT* has already mandated the use of *GPS*-based *AMC/G* for designated areas for which a *AMC/G DTM* exists.

As for the development of the 3D models, some *DOT*s, e.g., Iowa and New York, have taken the responsibility to develop *DTM*s while in some other states, e.g., Montana, the contractor is required to develop the *DTM*s. Nevertheless, almost no state has gone far enough to be confident of the accuracy of their developed models, and therefore they all transfer the responsibility for the accuracy of the provided data and corrections to the contractors. Similarly, all states have established *QC/QA* regulations based on 2D data used for the contract and refrain from leveraging the 3D models as the basis for *QC/QA*.

Although transportation agencies have started to appreciate the need for re-engineering their workflow for regularization of *AMC/G* in projects, the adopted measures are far from sufficient. The application of 3D models needs to be endorsed by, and formalized into, a structured framework whereby their validity is universally acknowledged across the entire project life-cycle, and rework and redundancies are eradicated.

The current workflow, for design-bid-build project delivery method, is schematically drawn using *IDEF*0, Integration Definition, process modeling platform as shown in Figure A-3(a). The process is viewed from the perspective of a strategy planner who intends to examine the efficiency of the current process and look into how to improve the process in terms of reducing redundancies and re-work as well as streamlining the data exchange and re-formatting. The

majority of U.S. transportation agencies we studied follow this workflow. A transportation agency approaches a designer for the design (A1 and A2 in Figure A-3(a)), and, regardless of the availability of 3D models, will later use 2D design models as the basis for bidding (A3). Once a contractor is appointed to the project, most frequently, the contractor is required to develop and adjust the *DTM*s and 3D design models usable by the *AMC/G* technology of their choice (A4, A5 and A6). Subsequently, the contractor will create an *AMC/G* work plan, directly executable by machines, (A7). Using the work and control plans, the engineers and contractor will control the operation. To conclude the project, the final work needs to be checked and approved by the engineers of the transportation agencies (A9).

The current workflow involves a great deal of redundancy and re-work that could be removed. For instance while in some cases the designers develop a 3D model, only a 2D model is given to the contractor who has to re-develop the same 3D model. This problem could be resolved through re-engineering of the administrative process and transiting from 2D design models to 3D design models as the primary source of the design. In this fashion, the design company will be obliged to deliver to the client the required 3D models.

Other issues surrounding the use of *AMC/G* in construction projects in the current workflow are: (1) Transformation of 2D models to *AMC/G* operable 3D models is reported to be cumbersome, error-prone and time-consuming (Yabuki 2011;Vonderohe 2007); (2) Clear *QC/QA* for *AMC/G* enhanced project execution is missing, and the adopted transplantation of conventional *QC/QA* regulations involves deficiencies in terms of inadequate *QC/QA* plans for AMC/G assisted project executions; (3) By not clarifying the liability issues for design errors and subsequent data adjustments, or in some cases by transferring all the risks to the contractors, transportation agencies fail to provide incentive for contractors to implement *AMC/G*; (4) The lack of

interoperability between different digital formats; (5) Security and ownership rights pertaining to the use of digital files are not properly addressed; (6) *AMC/G* tasks are susceptible to large errors due to the power vested in the technology; and (7) The data used for the execution of the project and further *QC/QA* are not unified.

Transition from 2D models to 3D models as the basis for contracting will contribute greatly towards the unification of documents used for execution and control. It is also recommendable that transportation agencies and engineers, investigating more meticulously the advantages of AMC/G, re-define the staking requirements and control plans for *QC/QA* purpose and harness the inherent power of new positioning technologies, e.g., *GPS*, towards *QC/QA*. Additionally, transportation agencies need to experience the use of *AMC/G* in projects more committedly so that they can reach to an articulation of responsibilities and liabilities that are more assuring to contractors. To resolve the ownership rights issues, it is proposed to utilize digital signature and data encryption scheme. This could be further enhanced through the implementation of version and change management using a unified data communication platform. It is noteworthy that the legitimization of the electronic exchange of data would allow better tracing and recording of the data exchange, and thus increases the security and ownership rights of digital files.

Incorporating the above consideration for improving the process, Figure A-3(b) illustrates our proposed workflow for the application of *AMC/G*. It suggests that 3D models and *DTM*s, unlike the conventional method, could be used as the base documents for bidding (B1-B3). In this fashion, the contractor needs only to improve, i.e., increase the level of detail, and convert the data format commensurate with their *AMC/G* technology of choice (B4).

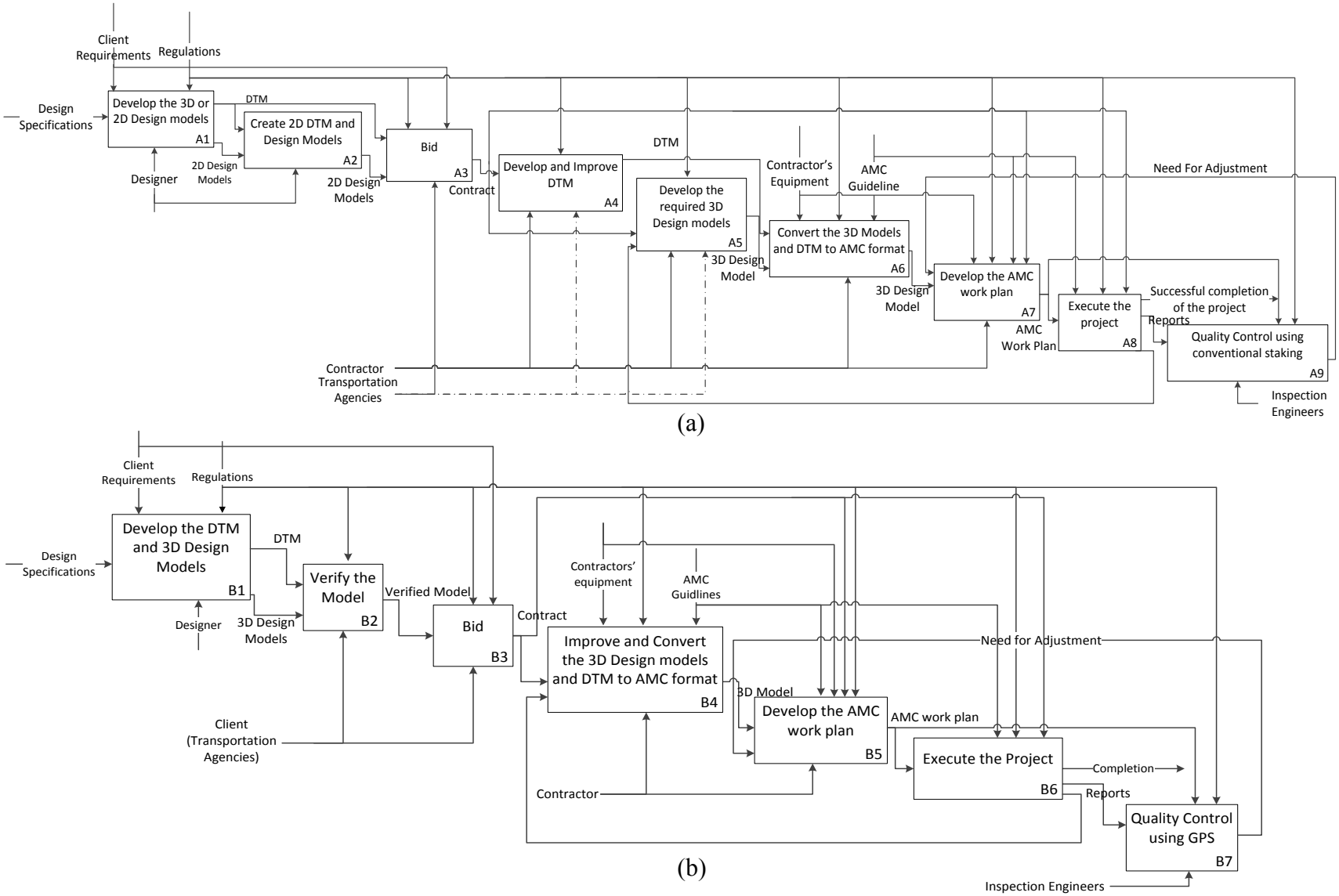


Figure A-3: Current (a) and Improved (b) Workflows for Application of AMC/G in Transportation Agencies Using Design-Bid-Build Delivery Method

Further amelioration to the process is the deployment of GPS and other state-of-the-art positioning technologies for the purpose of staking and quality control (B7). This will help reduce the amount of effort and time required for QC/QA in the course and at the end of the project. Another measure that is incorporated in the improved workflow process is the inclusion of a verification phase prior to handing over the documents to the contractor (B2). This practice would allow the engagement of transportation agencies more actively and committedly in the process, and thus offsetting the liability distribution amid major stakeholders. As shown in Figure A-3(b), the proposed improvement will shorten the process by two blocks, which represents a considerable level of smoothening.

While Figure A-3 schematizes the process for design-bid-build delivery method, we can further explore the improved process for the design-build delivery method. In this mode, the design and construction of the project will be assigned to the same entity. This will greatly help reducing the problem of interoperability as companies could confidently use their own standard for the entire process. The process model for this mode of project delivery is shown in Figure A-4.

A more futuristic vision of administrative process is to exploit a *MAS* structure. In this setting, every involved stakeholder is represented by an agent in the *MAS*. The agent system, using the power of *MAS* in coordinating communication, collaboration and negotiation of diverse players, would help smooth the data and information flow. Agents would monitor the data exchange in terms of compliance with the receivers' required data conditions, capacity, format, and delivery timeframe. In cases of conflicts, the agents could embark on negotiations and identify the most economic concession strategy to resolve the problem.

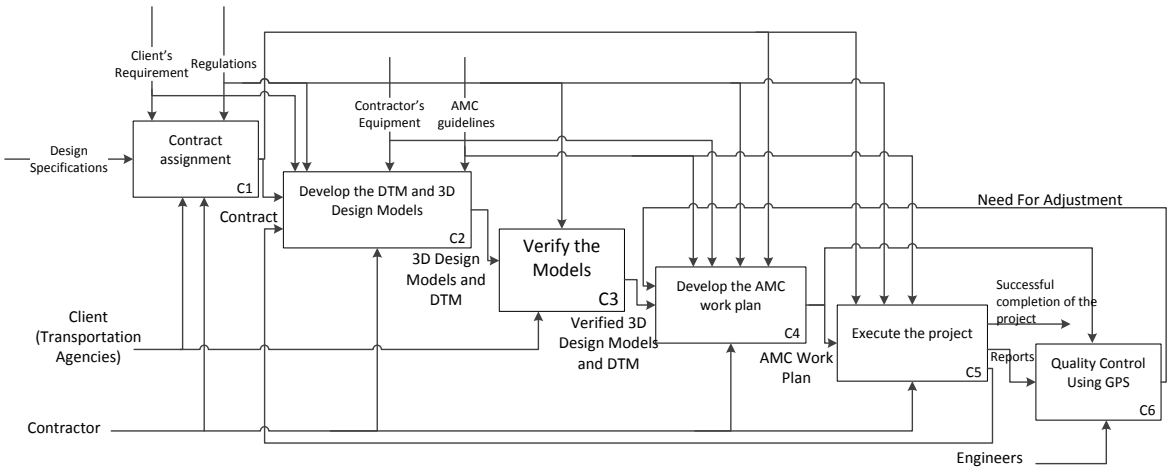


Figure A-4: Improved Workflow for Application of AMC/G in Transportation Agencies Using Design-Build Delivery Method

It is worth noting that 3D models will offer increased accuracy of any cost estimations to be performed throughout the process, be it by the contractor or transportation agencies. Conventionally, average-end-area method is used to calculate the earthwork volume and cost, which is based on the averaged area between two successive cross-sections multiplied by the distance between the cross sections in the interval ascertained by the required level of accuracy. However, the development of 3D models would allow transit to the surface-to-surface (S2S) method which offers greater accuracy (Vonderohe and Hintz 2010) at no extra cost and in shorter time (Yabuki 2011). *S2S* method is based on the matched triangulation of the two surfaces, as-is and as-designed, and the calculation of the volume of the triangular prisms formed by connecting the equivalent corners of the triangles from the two surfaces.

APPENDIX B: THE OUTDOOR TESTS

B.1 INTRODUCTION

Two case studies were performed to evaluate the performance of pose estimation module on actual construction equipment. The first test was carried out on a small construction site in the Loyola Campus of Concordia University, where a roller was tracked using *UWB*. The second case study was performed on a congested construction site in downtown Vancouver.

B.2 FIRST DYNAMIC TEST FOR TRACKING A ROLLER

Design of Test

In this test, which lasted 29.5 minutes, a roller on a small construction site (22.98 m × 14.03 m) on Loyola campus of Concordia University was tracked using four *UWB* tags, as shown in Figure B-1. Since there were 14 tags present on the monitoring cell, the update rate of *UWB* system was set to 8.3 Hz.

Performance Analysis

The analysis of the data was done in 2D, i.e., X-Y plane. For a more efficient analysis of the large amount of the collected data, the test was divided into six periods each of which lasted for five minutes.

In the first step of the analysis, the Missing Data Rate (MDR) of the tags was studied. It is observed that tags S2 and S4 had the worst and best performance with 73.91% and 23.93% of missing data rate, respectively. In general, tags S1 and S2 that were attached to the lower part of the roller had high *MDR* due to a low visibility. Figure B-2 shows the result of visualizing the raw location data of all four tags over the first period of the test. Through a visual inspection, it can be realized that there is little consistency between the locations of the tags. Also, in keeping

with the analysis of *MDR*, it can be observed that the data of tags S1 and S2 have more inconstancy compared to that of tags S3 and S4.

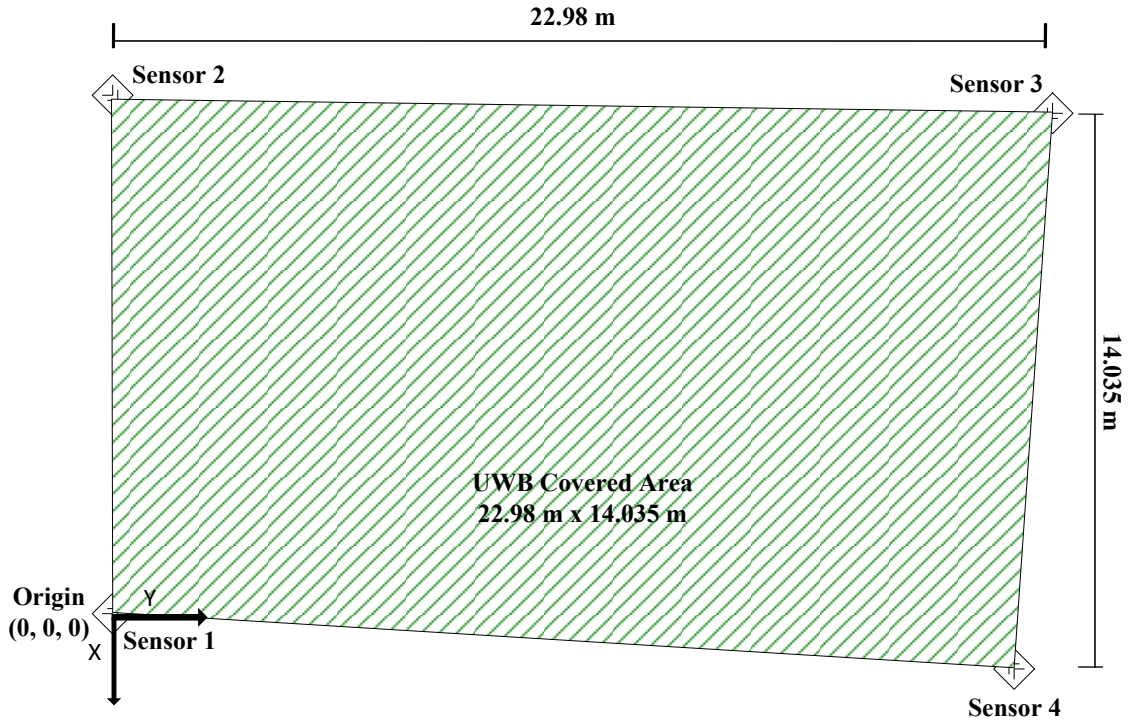
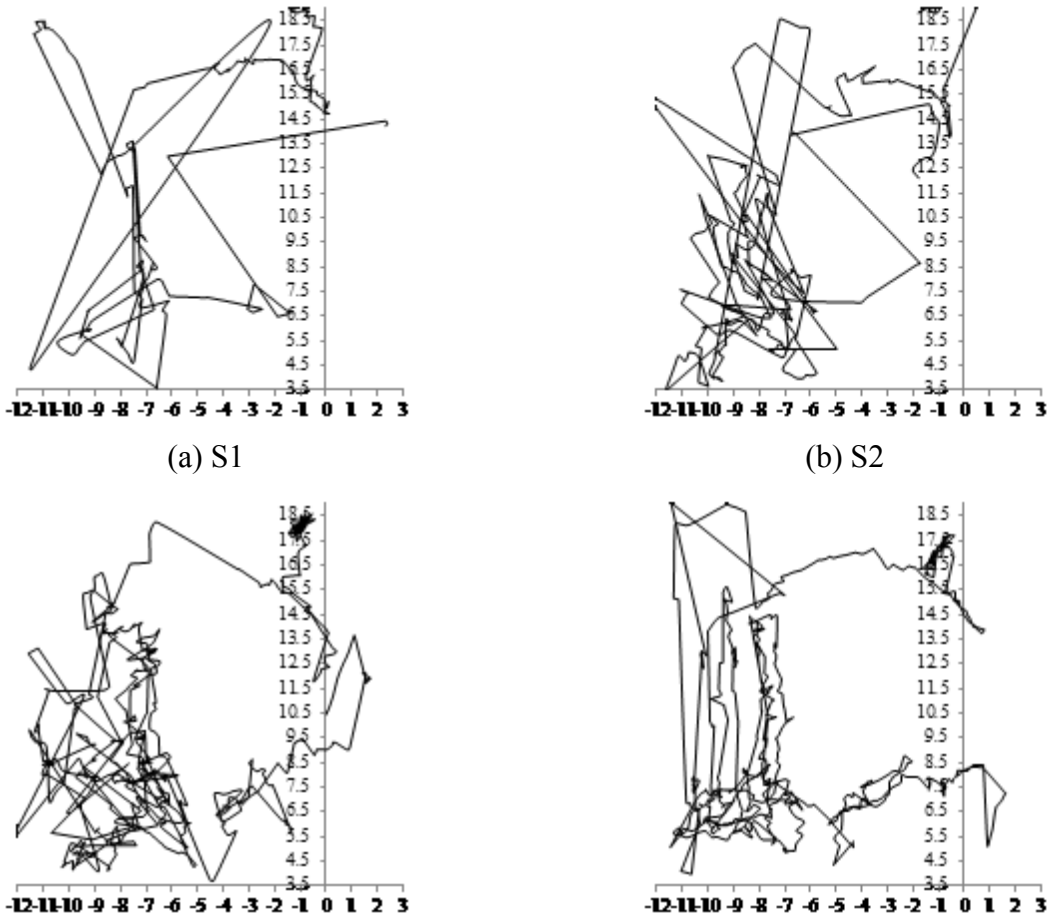


Figure B-1: (a) Monitoring Area Site, (b) Position of Tags, and (c) Roller during Compaction Process (Siddiqui 2014)



(a) S1 (b) S2

(c) S3 (d) S4

Figure B-2: Raw Data of All Tags for Period 1 (The X and Y axes are the coordinates of the tag in meters (Siddiqui 2014)

Data Enhancement

To enhance the quality of the pose estimation, the simplified correction method, which is explained in Section 4.2.2, is applied to the location data. The data were averaged over 3 seconds. There are six *GC*s that can be applied to the data, as shown in Figure B-3. Additionally, the maximum speed of the roller, which is obtained through considering the average and standard deviation of speed of Tag 4, was used to define the *OC*.

It is observed that when all four tags are considered, the quality of pose estimation after the correction method is not satisfactory. Therefore, the data of the tags with high *MDR*, i.e., S1 and S2, were removed from the analysis. Since the remaining tags, i.e., S3 and S4, had low *MDR*, the averaging period was set to 1 second. Figure B-4 shows the result of corrected data when the *GC* between S3 and S4 and *OC* is considered. Through the comparison of the Figures B-2 (c) and (d) with Figures B-4 (a) and (b) it can be observed that when the smoothness of the location data is considered, the applied simplified correction method improved the quality of data.

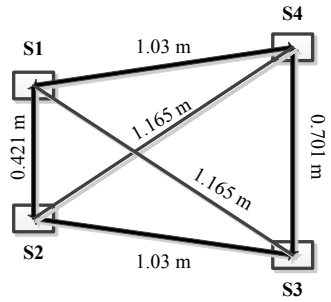
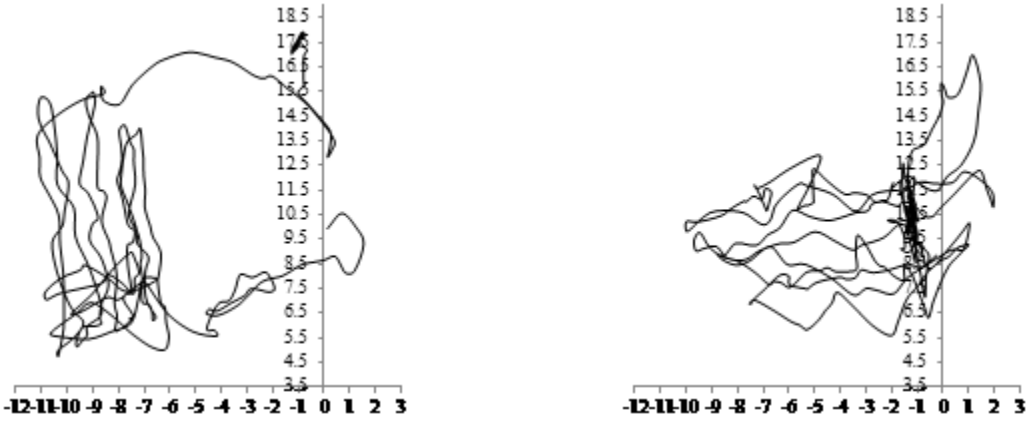


Figure B-3: Geometric Constraints of Roller (Siddiqui 2014)



(a) Period 1 (b) Period 2

Figure B-4: Results of Simplified Correction Method (The X and Y axes are the coordinates of the tag in meters) (Siddiqui 2014)

B.3 SECOND DYNAMIC TEST TO TRACK AN EXCAVATOR IN A CONGESTED SITE

Design of Test

To investigate the applicability of wireless *UWB* in an actual construction site a test was conducted in a site located in Downtown Vancouver. At the time of the test, two excavators were performing an earthwork operation in a relatively congested space (i.e., 36.5 m × 24 m). Figure B-5 shows the site layout. The main objective of the test was to track one of the excavators using *UWB* in a wireless setting.

Figure B-5: Site View on May 22, 2014 before Visit (Siddiqui 2014)

The bigger excavator in the center of Figure B-5 was the subject of the test. The test spanned over four days, i.e. from Monday, June 23, 2014 to Thursday, June 26, 2014. Figure B-6 depicts the progress of the construction work during the test period. Contrary to our expectations, a delay in the work resulted in a prolonged presence (i.e., 2 days) of a large mobile crane on the site. Figures B-6(a) and (b) demonstrate the level of congestion on the first 2 days of the test.

(a) Day 1

(b) Day 2

(c) Day 3

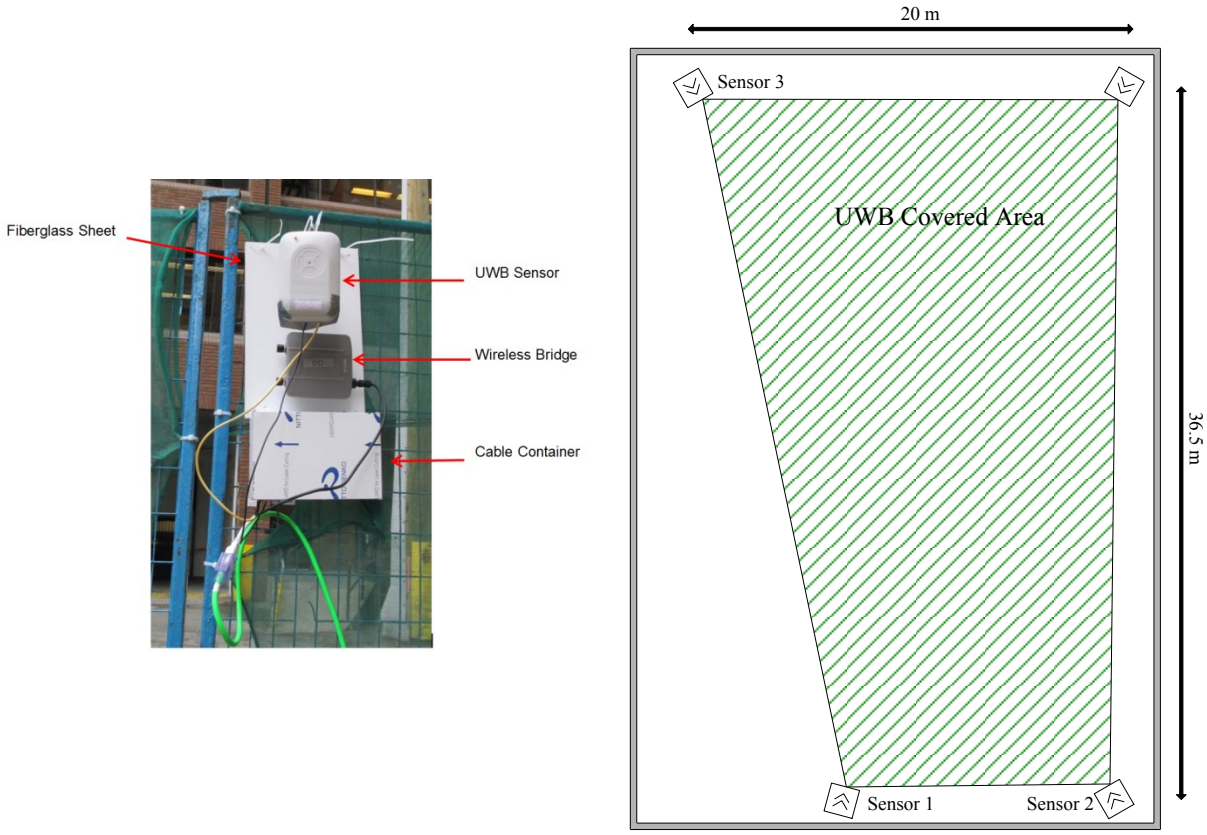
(d) Day 4

Figure B-6: Site Conditions for Each Day (Siddiqui 2014)

The presence of the crane was a setback for the designed test because the dominant operation during the first two days was the removal of a chimney structure rather than the expected

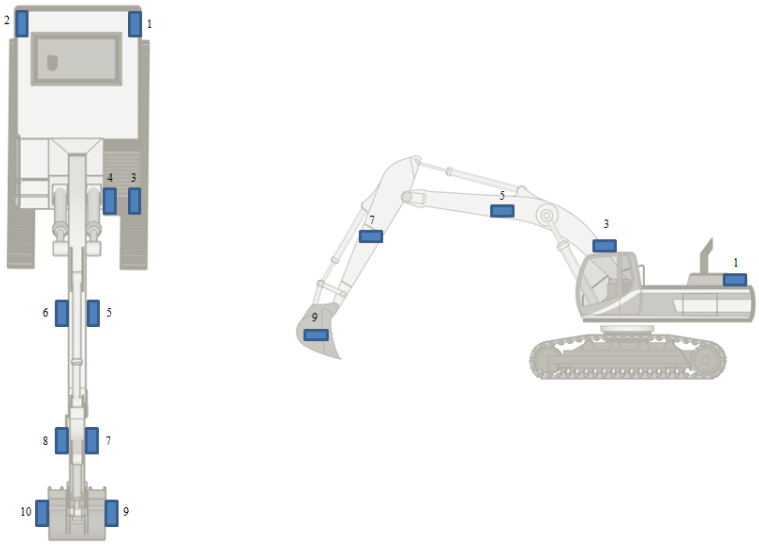
earthwork operation. Additionally, the heavy-metallic body of the crane introduced some noise for the wireless communication and, thus, disturbed the performance of *UWB* system.

Given the engineering constraints that hampered the use of tripods on the site, *UWB* sensors and their associated wireless bridges were placed on panels that could be easily installed on the fences or walls, as shown in Figure B-7(a). Once installed, the four *UWB* sensors covered an area of approximately 36.5 m × 20 m, as shown in Figure B-7(b). The working station, from where the *UWB* system was being monitored and controlled, was setup on the second floor of the existing building on the site. With regard to the power supply for *UWB* sensors, two separate generators were used to power the sensors that did not have access to the buildings' power outlets. A surveyor team who was present on the site helped measuring the sensors' position using total station. The Easting and Northing Coordinate System (ENCS) provided by the surveyors were transformed to coordinates in a local coordinate system. Nevertheless, the surveyors were not available during the calibration steps. As a result, the distances between sensors and the calibration tag were measured using a measurement tape. The presence of working equipment on the site together with the suboptimal measurement for the calibration have resulted in some difficulties in the smooth calibration of the *UWB* sensors. As shown in Figure B-8, ten *UWB* tags were strategically attached to the excavator using magnets. An IP camera was used to record the video of operation for further validations.



(a) (b)

Figure B-7: (a) UWB Sensor Panel and (b) UWB Covered Area (Siddiqui 2014)



(a) Upper View *(b) Side View*

Figure B-8: Excavator Tags' Positions (Siddiqui 2014)

Performance Analysis

For the purpose of data analysis, two portions of the overall collected data were analyzed in 2D plane. The chosen periods were both on the 4th day when the crane had left the site. While during the first chosen period the excavator was stationary, during the second period the excavator was performing some dirt removal from the site.

As a first observation, it was noted that 5 out of the 10 tags have a *MDR* of more than 90%, which has resulted in the overall Actual Update Rate (AUR) of lee than 1 Hz. As for the remaining tags, the *AUR* and *MDR* were found to be acceptable. The root of this inconsistency was found in the particular orientation of the excavator during the first period. In this period, as shown in Figure B-9, the excavator was facing two of the sensors from cabin side. Thus, the 5 tags attached to this side of the excavator had a better visibility and therefore had a higher *AUR* and less *MDR*.

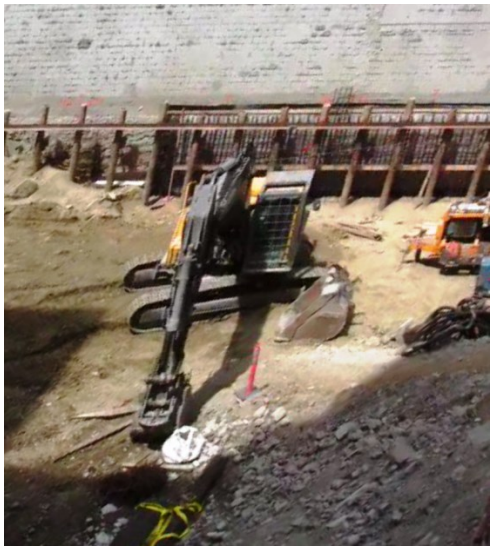
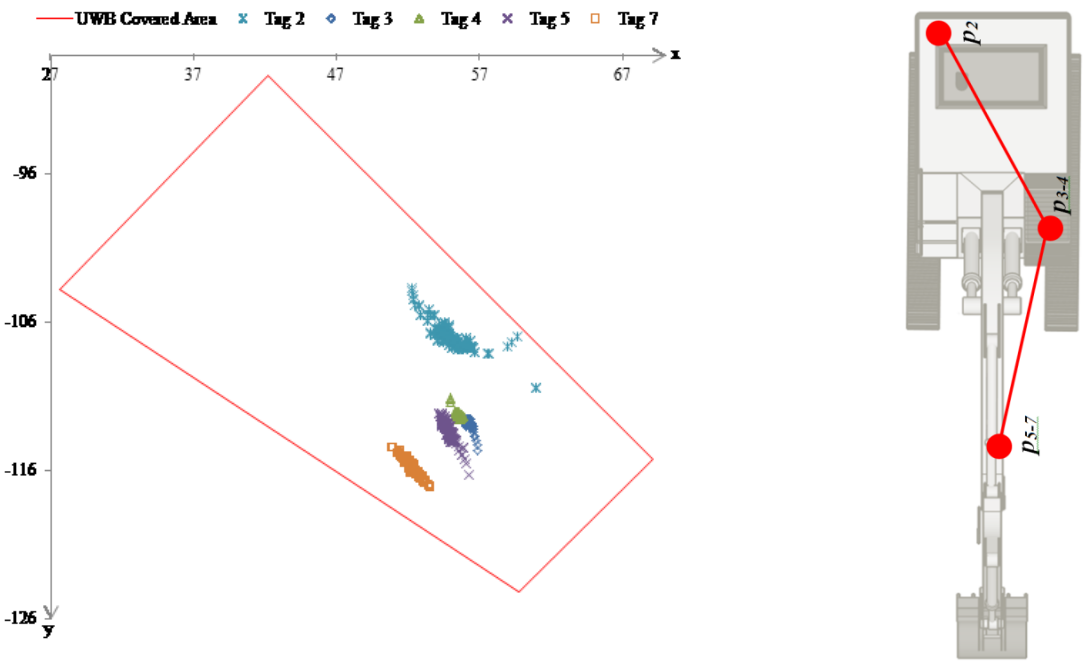


Figure B-9: Excavator Position at 12:53 PM on Day 4 (Siddiqui 2014)

From this point of analysis onward, only the five tags with acceptable *AUR* were accounted. The fact that the excavator has been stationary during this period was exploited to analyze the rate of

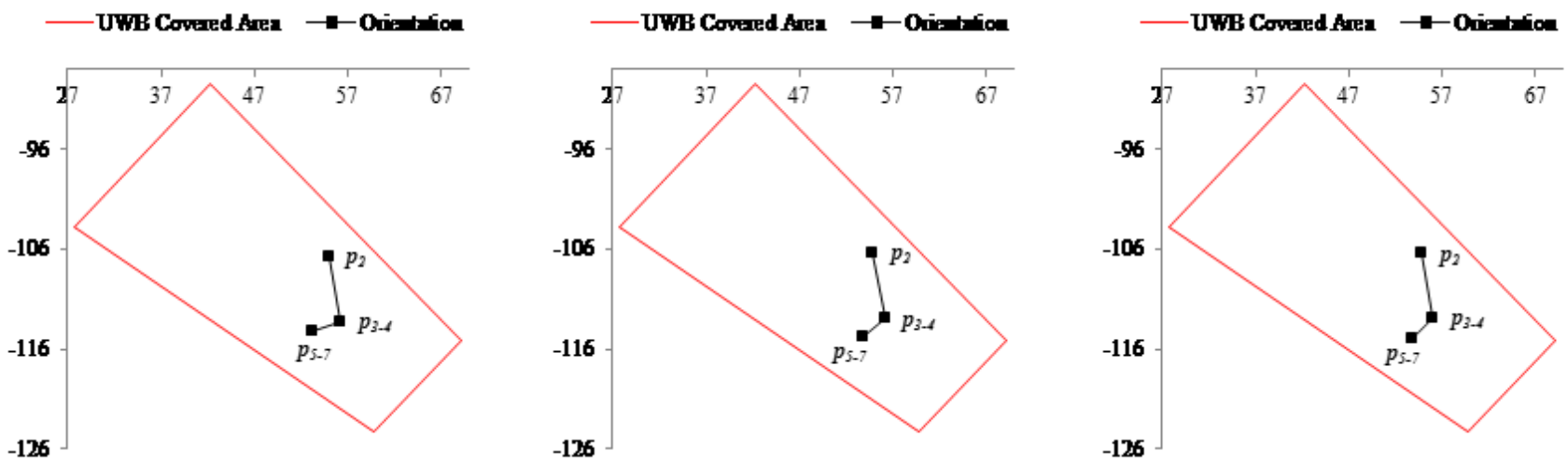
fluctuations in the registered location of each tag. It was discovered that Tag 2 had a high standard deviation (i.e., 1.37 m in x direction and 0.92 in the y-direction), representing high fluctuation of location error. The standard deviations for the remaining tags were smaller, e.g., 0.5 m in both directions for Tag 7. This observation can be seen in Figure B-10(a), where the registered locations of the tags were plotted.

In the next step of the analysis, the orientation of the excavator was scrutinized. To unify the timing of the data, the locations of all the tags were averaged over 1 second. Next, the tags that were attached in proximity of one another were averaged (i.e., Tags 3 with Tag 4 and Tag 5 with Tag 7). As a result, three data points were analyzed: (1) Tag 2 (p2); (2) Tags 3 & 4 (p3-4); and (3) Tags 5 & 7 (p5-7), as shown in Figure B-10(b). Given that the equipment has been stationary during this period, the angle between the two lines shown in Figure B-10(b) is expected to be constant throughout the period. Figure B-11 shows the plot of the three data points during six seconds of the analyzed period. Based on the visual comparison, no major discrepancies between the expected orientation and the measured orientations are observed.

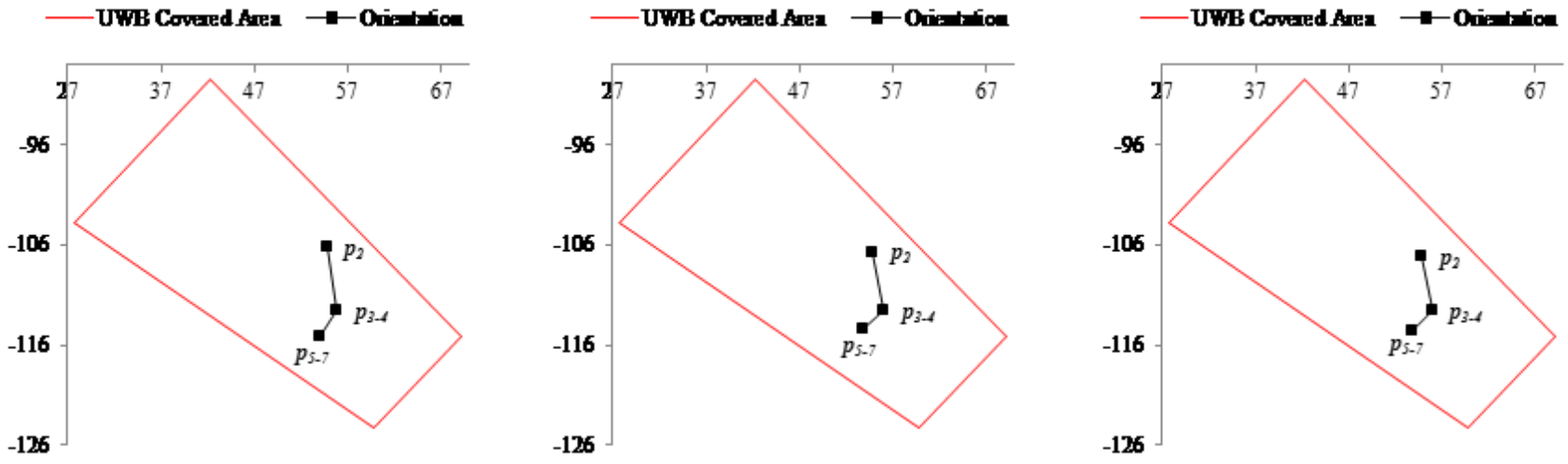


(a) (b)

Figure B-10: (a) Raw Data Analysis of Five Tags for Full Scale Outdoor Test, and (b) Schematic View of Orientation of Excavator (Siddiqui 2014)



(a) Second 1 (b) Second 2 (c) Second 3



(d) Second 178 (e) Second 179 (f) Second 180

Figure B-11: Scatter Plots for Orientation of Excavator – Period 1 (Siddiqui 2014)

APPENDIX C: LIST OF PUBLICATIONS

Journal Papers

- Vahdatikhaki, F., Lnagari, S. M., Hammad, A., Taher A. (2015). "Framework for Fleet-level Automated Equipment Guidance in Earthwork Projects Using Multi-agent System." *Expert Systems with Applications* (Under Review).

- Mawlana, M., Vahdatikhaki, F., Doriani, A., Hammad, A. (2015). "Integrating 4D Modeling and Discrete Event Simulation for Phasing Evaluation of Elevated Urban Highway Reconstruction Projects." *Journal of Automation in Construction* (Under Review).

- Vahdatikhaki F., Hammad A. (2015). "Risk-based Look-ahead Workspace Generation for Earthwork Equipment Using Near Real-time Simulation." *Journal of Automation in Construction* (accepted with modifications).

- Vahdatikhaki F., Hammad A. (2015). "Dynamic Equipment Workspace Generation for Improving the Earthwork Safety Using Real-time Location Systems." *Journal of Advanced Engineering Informatics* (Available online).

- Vahdatikhaki F., Hammad A., Siddiqui H., (2015). "Optimization-based Excavator Pose Estimation Using Real-time Location Systems." *Journal of Automation in Construction* (Available online).

- Vahdatikhaki F., Hammad A. (2014). "Framework for Near Real-Time Simulation of Earthmoving Projects Using Location Tracking Technologies." *Journal of Automation in Construction*, 42, 50-67.

- Hammad A., Vahdatikhaki F., Cheng Z. (2013). "A Novel Approach to Project-Level Automated Machine Control/Guidance Systems in Construction Projects." *Journal of Information Technology in Construction*, 18, 161-181.

Conference Papers

- Vahdatikhaki, F., Hammad, A., Langari, S. M. (2015). "Multi-agent System for Improved Safety and Productivity of Earthwork Equipment Using Real-time Location Systems." *5th International and 11th Construction Specialty Conference*, Vancouver, Canada (Accepted).

- Vahdatikhaki, F., Hammad, A. (2015). "Dynamic Equipment Workspace Generation for Improving Earthwork Safety Using Equipment Pose and State Data." *32nd International Symposium on Automation and Robotics in Construction and Mining*, Oulu, Finland (Accepted).

- Vahdatikhaki, F., Hammad, A. (2015). "Visibility and Proximity-based Risk Map of Earthwork Site Using Real-time Simulation." *32nd International Symposium on Automation and Robotics in Construction and Mining*, Oulu, Finland (Accepted).

- Vahdatikhaki, F., Hammad, A., Taher, A. (2015). "Multi-agent Approach for Automated Guidance and Control of Earthwork Equipment." *2nd International Conference on Civil Engineering Informatics,* Tokyo, Japan. (selected as an honored paper)

- Vahdatikhaki, F., Hammad, A., Setayeshgar, S. (2013). "Location-aware Real-time Simulation Framework for Earthmoving Projects Using Automated Machine Guidance." *Winter Simulation Conference*, Washington D.C.

- Vahdatikhaki, F., Hammad, A., Setayeshgar, S. (2013). "Real-time Simulation of Earthmoving Projects Using Automated Machine Guidance." *the 30th International Symposium on Automation and Robotics in Construction and Mining*, Montreal, Canada.

- Hammad, A., Vahdatikhaki, F., Cheng, Z., Mawlana, M., Doriani, A. (2012). "Towards the Smart Construction Site: Improving Productivity and Safety of Construction Projects Using Multi-Agent Systems, Real-Time Simulation and Automated Machine Control." *Winter Simulation Conference*, Berlin, Germany.

- Siddiqui, H., Vahdatikhaki, F., Hammad, A. (2014). "Performance Analysis and Data Enhancement of Wireless UWB Real-time Location System for Tracking Construction Equipment." *The 21st International Workshop on Intelligent Computing in Engineering*. Cardiff, UK: EG-ICE.

- Setayeshgar, S., Hammad, A., Vahdatikhaki, F. (2013). "Real-time Safety Risk Analysis of Construction Projects Using BIM and RTLS." *the 30th International Symposium on Automation and Robotics in Construction and Mining*, Montreal, Canada.

- Hammad, A., Vahdatikhaki, F., Albahri, A., Salimi, S., Ghadiri Moghaddam, D., Nejati, A. (2013). "Framework for Life-cycle Infrastructure Information Modeling and Management." *CSCE Annual Conference,* Montreal, Canada.

Conference Posters

- Vahdatikhaki, F., Hammad, A. (2015). "Towards the Smart Earthmoving Site: Project-level Location-based Guidance Systems." *Fiatech Technology Conference & Showcase,* Boca Raton, Florida, U.S. (Scholarship Recipient).

- Hammad, A., El Ammari, K., Langari, S. M., Vahdatikhaki, F., Soltani, M., AlBahnassi, H., Paes, B. (2014). "Simulating Macro and Micro Path Planning of Excavation Operations Using Game Engine." *Winter Simulation Conference*.