Modelling of a Hybrid MAC Protocol for M2M Communications

Farshad Keyvan Ghazvini

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science (Electrical Engineering) at

Concordia University

Montreal, Quebec, Canada

April 2017

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Farshad Keyvan Ghazvini

Entitled: Modelling of a Hybrid MAC Protocol for M2M Communications

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

Master of Applied Sceince in Electrical 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. Rabin Raut ____________ Chair

Dr. Amir Aghdam ____________ Examiner

Dr. Chadi Assi ____________ Examiner

Dr. Mustafa Mehmet-Ali ____________ Supervisor

Approved by ________________________________
Chair of Department or Graduate Program Director

____
Dean of Faculty

Date ________________________________

Abstract

Modelling of a Hybrid MAC Protocol for M2M Communications

Farshad Keyvan Ghazvini

Machine-to-machine (M2M) communications is one of the enabling technologies for connecting massive number of devices to the Internet of Things (IoT). M2M communications have different characteristics than human-to-human (H2H) communications. In this work, we propose a scalable, hybrid MAC protocol that will satisfy user quality-of-service (QoS) requirements.

We model both periodic and nonperiodic traffic. The proposed MAC protocol organizes transmissions into superframes consisting of a number of frames. A machine is assumed to generate a one or zero packet per its period. The machines have been divided into several types according to their packet generation probabilities. The generated packets are classified into different traffic classes according to their tolerance to packet losses and served by a subframe. Further, each subframe is divided into two sub-periods one serving contention and the other reserved traffic of that traffic class. We formulated an optimization problem that minimizes frame length subject to QoS user requirements. Then, we derived packet loss probability for each class as well as total packet loss probability for the optimization. Formulation resulted in a nonlinear optimization problem, but numerical results show that an LP approximation provides a nearly optimal solution.

The work also considered the proposed protocol under user mobility. The packet arrival process under user mobility has been derived. Then the performance of the protocol has been evaluated with the contention service under this arrival process. The contention service with and without packet losses have been considered. A priority queueing mechanism also has been studied for M2M communication. The results of this thesis may be useful in the design of M2M communication system.

Acknowledgement

I would like to express my deep gratitude to my supervisor Prof. Mehmet Ali for his valuable comments and continuous support toward completion of my graduate program. Furthermore, I would like to extend my sincerest appreciations to Dr. Doughan for his respected advices and contributions.

Also, I must express my deepest appreciation to my wife, Bahareh, who has believe in me and encouraged me to pursue my graduate degree.

Likewise, I am very grateful to my family, including my in-laws. This success would not have been achievable without them.

Table of contents

List of Figures

List of Tables

List of Symbols

b: number of slots in a frame

C: total number of reservation slots

$\overline{D}$: average packet delay in number of slots

$F_x(n)$: probability that a packet will be transmitted during the first n frames since its generation

G: normalization constant

g_j: proportionality factor for contention sub-periods to be adjusted to meet the QoS requirements

$\hat{g}_j$: minimum values of g_j satisfying the packet loss requirements in the optimization solution

g : total number of packets that have not been successfully transmitted during a frame

K: number of machine types

k_j: number of class j packets that may be successfully transmitted during class j contention sub-period

k: number of packets that may be successfully transmitted during a frame

ℓ_j: number of packet losses in class j contention subperiod in a frame

ℓ: total number of packets losses in contention subperiods in a frame

M: number of frames in each superframe

$\overline{N}$: mean number of packets in the system

N: total number of machines being served during a frame

$\overline{N}_{ij}$: average number of type i machines that belong to class j

N_i: number of type i machines being served during a frame

n_i: number of the packets in the queue at the end of i'th frame

n_i: number of type i machines in the system

n_{T_j}: total number of class j packets generated during a frame

n_T : total number of packets generated during a frame

P_s: probability of a successful transmission during a contention slot

$P(z)$: PGF of the number of packets in the system at the steady-state

P_{B_i}: probability that reservation request of a type i machine will be blocked

P_k: probability that k successful packet transmissions will occur during a frame

P_g: probability that a packet willnot be successfully transmitted during a frame

P_{ℓ_j}: probability that a packet will not be successfully transmitted during j'th frame, j=1,2,3…

P_ℓ : probability that a packet will not be successfully transmitted during a frame at the steady-state

$P_x(n)$: probability that a packet will be transmitted during the n^{th} frame since its generation

$P_{\omega_i}(j)$: probability that packet interarrival time of a type-i machine is j frames

P_{r_j}: probability that r_j packets will be generated during a frame

P_{k_j}: probability that k_j successful packet transmissions will occur during the class j contention sub-period

P_{ℓ_j}: probability of class j packet loss

P_ℓ: probability of packet loss

p_{ij}: probability that packet generated by a type i machine will belong to class j

p : probability of successful packet transmission during a slot

Q_i : probability that a packet of a type-i machine will be discarded

$q_{i,k}$: packet generation probability of type i machines in step k

q_i: probability that a type i machine will generate a packet during a frame

r : number of successful packet transmission during a frame

r_i: number of packets generated by type-i machines during a frame

r: total number of packets generated by all the machines during a frame

r_{ij}: number of class j contention packets generated by type i machines during a frame

r_j: number of class j contention packets generated during a frame

T: durations of a frame in number of slots

T_F: maximum frame duration in number of slots

T_j: durations of j'th subframe in number of slots

T_{Cj}: the durations of contention subperiods in the j'th subframe

T_{Rj}: the durations of reservation subperiods in the j'th subframe

$u(\cdot)$: steady-state probability distribution of the number of busy slots

X_{ij}: number of type i machines that will receive service in the contention subperiod for class j traffic

X_{Cj}: number of slots allocated for class j contention sub-period in a frame

$\hat{X}_{ij}$: values of the decision variables for the contention traffic

x : random variable denoting the number of the frames that elapsed from the generation of a packet to its successful transmission

Y_{ij}: number of type i machines that will receive service in reserved transmission subperiod for class j traffic

Y_{Rj}: number of slots allocated for class j reserved sub-period in a frame

$\hat{Y}_{ij}$: values of the decision variables for the reservation traffic

z_j: number of packet losses during a frame

α_i: total number of packets generated during the i'th frame

β: the period of a machine in number of frames

β_i: residency time of type i machines in the network

γ_i: probability that new arriving machine belongs to type i class

δ_j: maximum packet loss probability

δ: total packet loss probabilities

$\mathcal{E}_{i,j}$: type-i machines packet arrival rate during the j'th frame

$\mathcal{E}_i$: packet arrival rate of type i machines at the steady-state

Λ: arrival rate of the machines requesting reservation service in machines/sec

λ: arrival rate of machines requesting contention service in machines/sec

ρ_i: traffic load of type i machine in contention

τ_C: average number of slots it takes to transmit a packet successfully during a contention period

ω_i : interarrival time of the packets to a type-i machine in number of frames

Chapter 1

Introduction and Literature Review

1.1. Chapter overview

In our time, we have numerous technological devices surrounding us. One can gain variety of benefits by using these devices. For having more comfort and benefits from these devices there is a need of communication among them. This communication named Machine to Machine communication, is also referred to as Machine Type communication (MTC).

At first, the definition, applications, and characteristics of Machine to Machine (M2M) communication will be presented. Then, the remainder of chapter will provide a literature survey of M2M communications.

1.1.1. Definition of M2M communication

M2M communication is the system of communication for devices without human involvement. This communication can be directly device to device communication or through a network. For scalability and security, devices need to have connectivity with no human interaction. Connecting all the devices in Internet of Things (IoT) would be achievable by using M2M communication. It is expected that M2M communication will experience a fast growth in the next few years.

1.1.2. M2M applications

It is envisioned that M2M communications will have wide range of applications in different industries, including healthcare, manufacturing, transportation systems, smart grids, etc. [20, 24]. Smart metering is expected to be an important component of smart grid systems [23, 26]. In M2M applications, sensors will be used to collect all kinds of data, which will be forwarded to servers for processing through IoT. The heterogeneity and variety of devices in M2M communication increases challenges of meeting user requirements when designing this kind of network.

1.1.3. M2M communication characteristics

M2M communications have different characteristics than Human-to-Human (H2H) communications, in terms of the traffic generation, quality of service (QoS) requirements and device characteristics [21]. Integrating heterogeneous multitude of devices in the same network is a challenging task and needs a lot of research. Having a reliable network management and service management is important. Also, low latency and high security are critical criteria in applications such as smart grid and eHealth. In addition, embedded smart meters and smart devices with limited functionality need low power consumption protocol [30].

M2M devices usually generate small amounts of data infrequently, while others generate periodically [25]. Some M2M devices need to send packets periodically such as health care devices and other may have event driven scheme that need to send a packet each time that an event is triggered by the server. M2M devices generate more traffic on the uplink, while H2H generates more on the downlink. M2M devices are usually low-cost, battery operated with limited energy resources.

1.2. Categorization of M2M communications

M2M communications may be categorized based on the MAC protocols and the type of communication infrastructures.

As our communication systems was mainly based on the Human-to-Human (H2H) communication, protocols was designed to satisfy the requirements of the H2H communications. Many M2M applications require reliable and secure communications and some of them require low latency. As a result, current communication networks are not suitable for M2M communications as they have been optimized for H2H communications. Therefore, to find a suitable medium access control (MAC) protocol for M2M communication, researchers proposed new protocols by combining and enhancing the existing ones. Further, communication may also be categorized depending on the type of infrastructure such as wired, wireless, cellular or cognitive based communication.

1.2.1. Infrastructure based classification of M2M communications

Among various network technologies that have been considered for M2M communications are WiFi and cellular networks [6]. WiFi networks lack widespread coverage, reliability and security. While cellular networks based on the Long Term Evolution (LTE) standard do not have those drawbacks, they have been optimized for H2H communications. LTE systems are connection-oriented which require connection set up through a physical random access channel (PRACH) before data transmission. The overhead of connection set up is too large since the amount of transmitted data is quite small in M2M communications. It has been reported in [1] that the signaling overhead of such a connection is about 59 bytes on the uplink and 136 bytes on the downlink, which may be more than the amount of data to be transmitted.

Next, we will describe various types of infrastructure based communication systems.

1.2.1.1. Wired based systems for M2M communication

In home and building automation, smart grid, and smart metering Powerline Communication (PLC) may be used to provide network connectivity for sensors, smart meters, and M2M devices. For instance, PLC may be used in distributed service operator (DSO) network as a communication infrastructure between electricity meters and data concentrator [30]. Wired systems have lack of device mobility and its infrastructure for huge network connectivity may not be cost worthy. While, it can be used along with other M2M network communications system.

1.2.1.2. Wireless based systems for M2M communication

As the uplink transmission in M2M communication is very important, in [4] for increasing the uplink access efficiency they proposed a new MAC enhancement. They used IEEE 802.11ah network that consist of sensor stations and cellular traffic. In this network there are two access windows to reach the Access Point (AP). First one is restricted access window (RAW) that is used by sensor devices and the other is common window that is accessible by all devices.

In [4], devices randomly select a slot in RAW for accessing the channel and they sleep in other slots. AP transmits an ACK or downlink data frame when a device succeeds in accessing the channel. They used the Maximum Likelihood (ML) estimation method from [14] to find the approximate number of devices that can access uplink slots.

For determining the size of the uplink RAW in [4], they used success probability. They showed that their proposed algorithm can determine the size of the RAW using total success probability with a good accuracy. So, by finding the optimal size of the RAW for next interval more devices can access the uplink channel in an efficient way. This work determines the optimal value for the RAW size without optimization wrt cellular traffic.

1.2.1.3. Wireless cellular systems for M2M communication

Next, we will consider different systems that have been proposed to utilize the existing cellular networks as infrastructure for M2M communication. From the variety of cellular networks for M2M communication that we will present few of them such as Code-Expanded Random Access, and enhanced LTE-A.

1.2.1.3.1. Adaptive MAC for cellular network

As mentioned in [5], the slotted-ALOHA has maximum throughput of 38.6% and CSMA needs carrier sensing module that GSM/GPRS module does not have it. So, for the cellular based M2M solution in [5] starts with Slotted Multiple access collision avoidance (S-MACA) that has been proposed in [12]. Then, they designed Adaptive control system based on traffic load statistic estimation. They only use contention based system without any device and added request and acknowledgment to reduce the data collisions.

The specific details of S-MACA is shown in Fig. 1.1. In [3] they statistically estimate the traffic load to control the RTS contention. In S-MACA when the optimum traffic load is expected, the throughput becomes unstable. To maintain the traffic load constant in ATL S-MACA, BS calculates the RTS transmission probability by estimating the traffic load. As a result, in ATL S-MACA the access delay will increase linearly with increasing the traffic load because of the need to maintain a constant throughput.

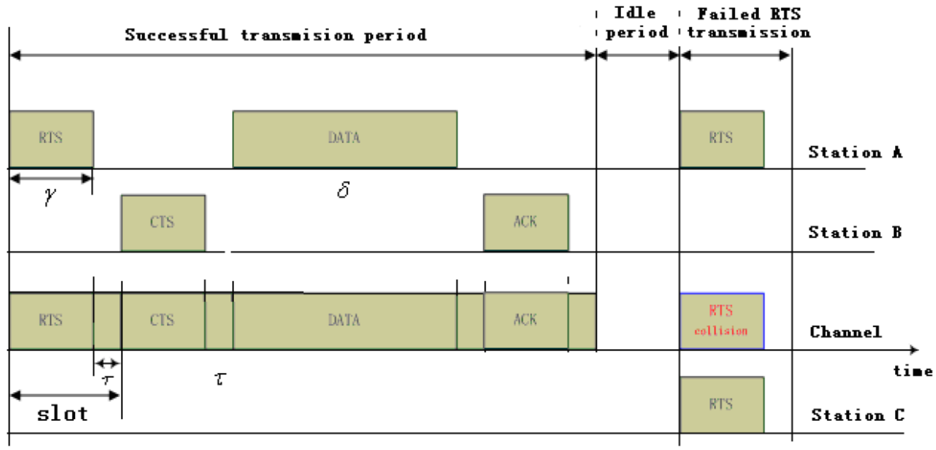


Fig. 1.1. Slotted MACA mechanism [5]

Proposed protocol in [3] is not very complex because it does not have the carrier sensing model and only needs the traffic load measurement and can be considered as a cellular based M2M module. They proposed "Pair-and-Go" for better channel utilization but this may cause problem when there is be an early departure in the channel.

1.2.1.3.2. Code-Expanded Random Access

Based on LTE random access and dynamic Random Access Channel (RACH) resource allocation approach, in [9] they proposed a random access method. In this method they used code expanded random access instead of traditional resource allocation. An example of this can be seen in Fig. 1.2.

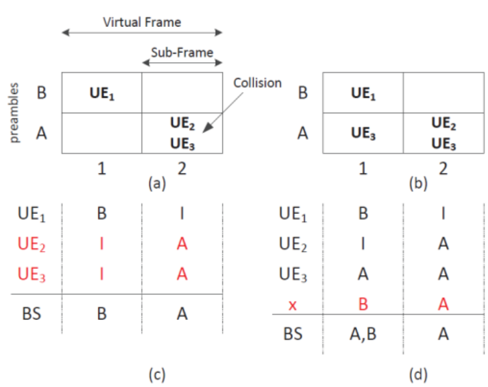


Fig. 1.2. (a) Reference random access, (b) Code-expanded ransom access, (c) Reference random access codewords, with collisions in red, (d) Code-expanded codewords, within phantom codeword in red. [9]

Contention phase efficiency have been calculated by finding expected value of the number of codewords used by single or multiple devices [9]. Also, they calculated code-expanded random access efficiency by finding number of single codewords and expected number of codewords observed by BS.

In [9] have been shown that their proposed scheme has high efficiency over large loads and adaptive random access scheme from combining reference and code-expanding random access can be efficient in high and low data load. Further, M2M devices using low amount of preambles and usage of other preambles by human-centric communications, these two traffics can coexist. Reaching optimal operation by estimating the user load is left for further research. In [9], there is no estimation of the uses load or any class for the devices. Also, for coexistence of H2H and M2M communication there would be a need of optimization.

1.2.1.3.3. LTE-A systems in M2M communications

In [11] to use LTE-A system, which had been designed for high throughput and low access delay, for M2M communication they present improvement in physical layer, MAC layer, and core network. They state that modules like GSM/GPRS are designed for screenless mobile phone are not that suitable for sensors. They proposed a solution for machines that can access the cellular network with minimum functionalities to reduce cost and size of the devices.

There are wide variety of M2M devices as mentioned in [11] and one solution or hardware platform cannot be suitable for all devices. For example, video surveillance device cannot use the protocol that is used for metering devices in an efficient way. Although, a lot of M2M devices have fix places and user mobility management systems are not suitable for them.

For efficient use of the physical layer in M2M communication, in [11], they suggest that each vendor may choose one option in each block of physical layer that is shown in Fig. 1.3, to have the customized physical layer. Also, it is required to pick at least one of the configurations in each block to complete the chain of their customized physical layer.

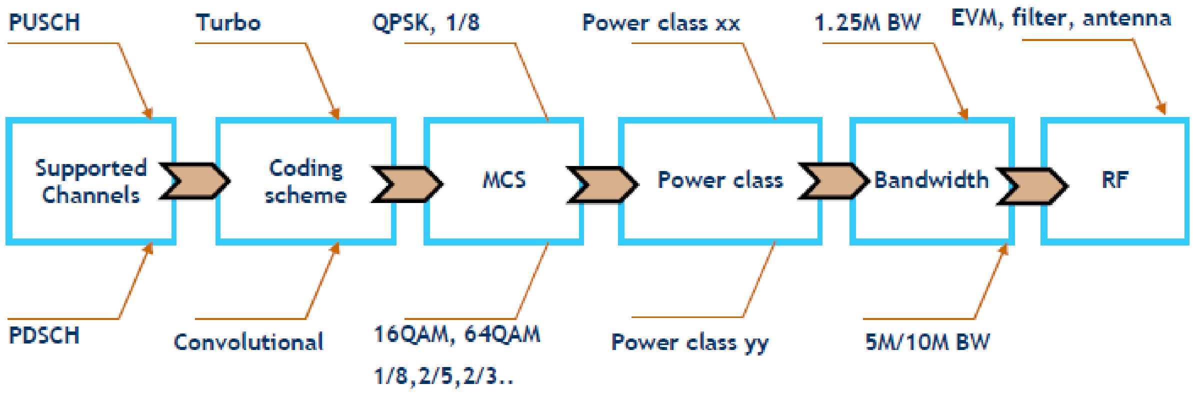


Fig. 1.3. Blocks of Physical Layer [11]

For optimizing the MAC layer they proposed two options in [11]. In the first option they used MAC PDU which does not carry a RLC PDU to limit the data transmission to MAC layer. In the other option they reduced the protocol by using the special preamble and coded data transmission.

In [11], for M2M communications, they suggest that sensors send the data to a proxy and the proxy combines all the received data from devices in to desirable packages to reduce the signaling overhead. Then, by using TCP/IP flow connect to the network and send the data to the dispatcher. The dispatcher then will connect to the server according to the server ID and will send the data to the server. At the end, they analyzed the signaling over head of the proxy-dispatcher system. Implementation of this system would be complex and signaling overhead would be an issue. The usage of proxy-dispatcher for large distributed network may not be sufficient.

1.2.1.3.4. Using enhanced LTE-A

The characteristics of M2M communication that has been mentioned by [8] are: high device density in a cell, little data in payload, and low traffic machine-originated communication. They mentioned that the important issue in M2M communication is the Physical Random Access CHannel (PRACH) overhead due to attempting to access the channel by large number of devices at the same time. So, they proposed a self-optimizing overload control (SOOC) mechanism that consists of congestion monitoring, decision-making and PRACH resource adjustment.

In [8] the solutions proposed by 3GPP have been described. These schemes are: backoff, slotted-Access, access class barring (ACB), Pull-based, PRACH resource separation, and dynamic PRANCH resource allocation. Then they proposed a SOOC mechanism for adding or reducing PRANCH resources and to prevent overloading they used separation, access class barring, slotted-access, and p-persistent schemes. In their protocol they used priority based on the time sensitivity of the packets and also they put a counter to change the probability of attempting to access based on the number of previous attempts. So, according to RACH load condition this self-optimizing control mechanism can allocate RACH resources.

1.2.1.3.5. Cognitive based schemes

In [1], to overcome the spectrum scarcity issue of M2M communication they proposed a cognitive MAC protocol with the usage of TV white spaces (TVWS) as the physical resource. They also refer to sensing and geo-location methods for finding the TVWS spectrum and used Packet Reservation Multiple Access (PRMA) for their cognitive MAC protocol. PRMA is a combination of S-ALOHA, TDMA and reservation scheme to support both periodic and bursty traffic.

As mentioned in [15] the cognitive system consist of one primary network and one or more secondary networks. The secondary networks try to use the opportunistic access to the spectrum resources without preventing proper working of the primary network. They used the cognitive system with the OFDMA based primary network and the secondary network is used for Machine to Machine (M2M) communication. The M2M system can use spectrum sensing to find the spectrum holes in OFDMA transmission that are not occupied.

In [15] they state that each User Equipment (UE) can determine the occupancy of subcarriers in OFDMA. So, they used a Cognitive Station (CS) that can detect OFDMA signals and use the unoccupied resources by the primary network to send M2M data in the same carrier. In their protocol they used pooling based system to organize M2M devices for accessing the network.

Five different methods have been considered in [15] for pooling based scenario, which are: exhaustive, gated, limited, fixed, and adaptive approaches. For performance evaluation they used measures related to the ZigBee system and they considered Poisson distribution for packet generation in their CSs. Lastly, for the performance evaluation they showed numerical results for average delay, overall throughput, resource wastage, and average queue length for these five methods. Structuring M2M communication network as a secondary network on the unused spectrum may have lack of scalability.

1.2.2. Protocol based categorization

Categorization of the M2M communication can be based on the protocol that they are using like reservation and contention. In [25], a survey of suitability of existing MAC protocols as well as review of MAC protocols specifically developed for M2M communications have been presented. The surveyed protocols have been classified as contention-based, contention free and hybrid protocols. Hybrid protocols usually divides the frame into two sub-periods: contention-based and contention-free. The contention-based transmission is used to send resource allocation requests for the contention free period [10]. There have been also hybrid protocols that switches between contention-based and contention-free transmission depending on the load [25].

1.2.2.1. Contention oriented protocol

As a contention based protocol, [2] presents Distributed Point Coordination Function-M (DPCF-M) as a novel "duty-Cycled energy-efficient Medium Access Control (MAC) protocol", which uses CSMA/CA. They considered two kind of M2M devices: (a) single radio access technology (RAT) that equipped with a low-power short-range RAT and (b) dual RAT that equipped with one short-range RAT and another RAT with cellular connectivity. The dual-RAT devices can be used as M2M gateways and provide cellular connectivity for single-RAT devices. To overcome the lack of energy efficiency in DPCF protocol because of periodic polling messages and lack of duty-cycled operation they rotate the role of M2M gateway among different devices dynamically.

[6] proposes a fast adaptive slotted Aloha (FASA), a contention-based protocol, which estimates the number of backlogged devices to assign node transmission probability in a slot. They mentioned that the M2M communication is event-driven and in a short time there are a lot of devices attempting to access to the server. So, they state that mechanism like ATL S-MACA in

[13] that uses packet sensing and uses the Poisson traffic is not suitable for this kind of communication and instead it is better to use Fast Adaptive S-ALOHA (FASA).

In FASA scheme [6] they considered S-ALOHA based random access control with a base station (BS) and huge number of M2M devices with a single RACH data attempting. Also, they assumed ideal collision channel mechanism that device with packet to send will go to backlogged state and transmit with the calculated probability from the estimation that the BS broadcast at the beginning slot. To estimate the number of backlogged devices they used drift analysis of fixed step-size estimation scheme. When estimation of the backlogged devices is far away from actual number, adjustment of step-size in FASA is needed.

They showed that FASA [6] scheme can estimate number of backlogged devices better than Pseudo Bayesian ALOHA (PB-ALOHA) and with comparison to multiplicative schemes under heavy load their simulation results show that it has better stability and delay performance. In [6] they only used contention based system and they did not consider bursty and Poisson arrival together.

In [16] they state that schedule-based protocols have complex nodes and contention-based protocol with single sink node causes throughput reduction due to the contention between devices.

They proposed hierarchical M2M network with cluster nodes as shown in Fig. 1.4. The relationship in this hierarchical network is based on child-parent scenario and data generated by the nodes will go to the sink node through their parent nodes. Data interval is based on beacon interval and there are three modes as: Rx mode (data reception), Tx mode (data transmission), and sleep mode. The beginning of the Rx mode at the higher level will be synchronized with the Tx mode of the lower level and the MAC protocol is based on CSMA/CA.

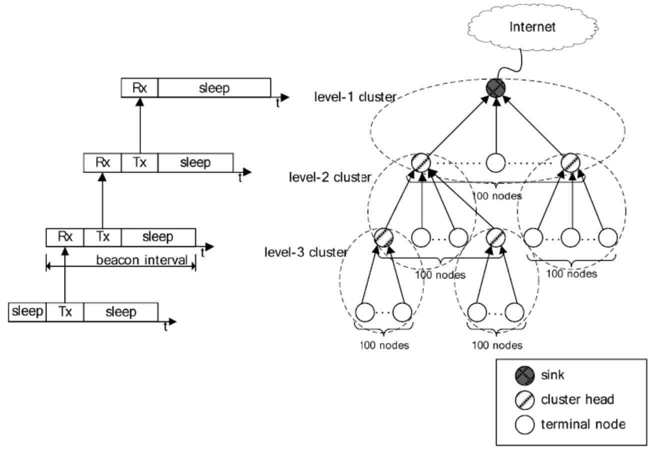


Fig. 1.4. DMAC and hierarchical architecture of M2M networks [16].

In [16] they proposed a backoff time decision rule for the CSMA/CA protocol based on the length of the data. They simulated the proposed protocol by MATLAB and analyzed the success access probability. The comparison of this protocol with SMAC and DMAC showed that proposed protocol and DMAC have less average latency. Also, this protocol improved energy consumption with comparison to DMAC.

1.2.2.2. Hybrid (contention and reservation)

In [10], they present a hybrid MAC protocol that consist of two parts: contention period and transmission period. They propose an optimization formula to maximize the system throughput by balancing the tradeoff between transmission and contention period.

As mentioned in [10], MAC protocols could be contention-based or reservation-based. In contention-based schemes, probability of collision accordance will increase whenever huge number of M2M devices are attempting to access the base station. On the other hand, in reservation-based schemes when number of devices that have information to transmit is low, then utilization is not efficient. As a result, they propose a hybrid MAC protocol which is combination of the two schemes. In their protocol each frame consist of Contention Only Period (COP) that is based on CSMA/CA access method and Transmission Only Period (TOP) that is based on TDMA data communication. They formulate an optimization problem to find the optimal contention probability for COP and the optimal number of devices in TOP that will maximize the throughput.

In the system model in [10] there are 4 parts in a frame. Frame starts with Notification Period (NP) during which base station (BS) notifies the beginning of the contention to all the devices. Then, the p-persistent CSMA based contention period begins. Afterward, in Announcement Period (AP) the slots belong to each device is broadcast and then the TOP begins.

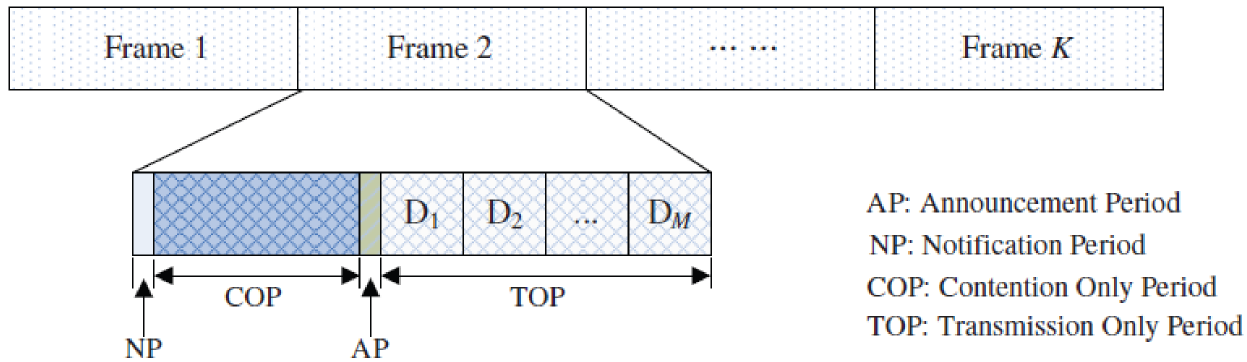


Fig. 1.5. The Frame structure [10]

In [10] the aggregate throughput is defined as the sum of the throughput of all devices which have transmission slots allocated to them during each frame. Then, by solving a convex optimization problem they attempt to maximize the aggregate throughput subject to the constraint that the summation of the COP and TOP durations should not be more than the duration of the frame. They used the result of this optimization problem in the NP to broadcast the duration of the contention period and the contention probability by the base station to all the devices after broadcasting the advertisement message (ADV) and estimating the number of the devices that have information to transmit.

In the COP a device that has data to transmit sends its transmission request (Tran-REQ) using the p-persistence CSMA protocol and if the Tran-REQ is received successfully without any collisions the BS increments the counter of devices by one. The number of the devices will be controlled by the optimal number of devices and optimal duration of COP from the optimization problem results. Then in AP, BS broadcasts the announcement message with successful devices' IDs and the transmission schedule mentioned in the message. The devices that don't have data or their ID is not verified will go to sleep mode.

In TOP the devices that have slots assigned to them for transmitting data will turn on their radio module and send their data to the BS.

At the end they compare their protocol with the contention-based protocol – slotted-ALOHA and reservation-based protocol – TDMA. They showed that the aggregate throughput and utility of their protocol will be higher than slotted-ALOHA as the number of devices increase and is always higher than TDMA.

In [10] the transmission will be only in the TOP and they did not consider the QoS of the devices or any class of devices.

The main concern of [17] was energy efficiency of the cellular network infrastructure. They modeled MAC with two parts of CSMA and TDMA. Also, they used the partial clustering in order to reduce energy consumption by having less number of devices contending for channel access. In this clustering the devices far away from BS that have transmission power higher than a threshold gathered into clusters and device with lowest transmission power will be the cluster head (CH). Non-clustered devices considered as CHs that have no other member in their cluster.

Frame formation in [17] consists of two parts. First part is for intra-cluster communication that uses CSMA/CA protocol to communicate within cluster members. The second part uses TDMA for cluster heads and BS communication to tackle heterogeneous traffic pattern.

To achieve lowest power-wasting in [17] protocol, they used non-persistent CSMA that has less implementation cost and is suitable for M2M communication that has small packet size transmission. In addition, they used multi-phase CSMA protocol, which has less collisions due to division of contending duration, shown in Fig. 1.6.

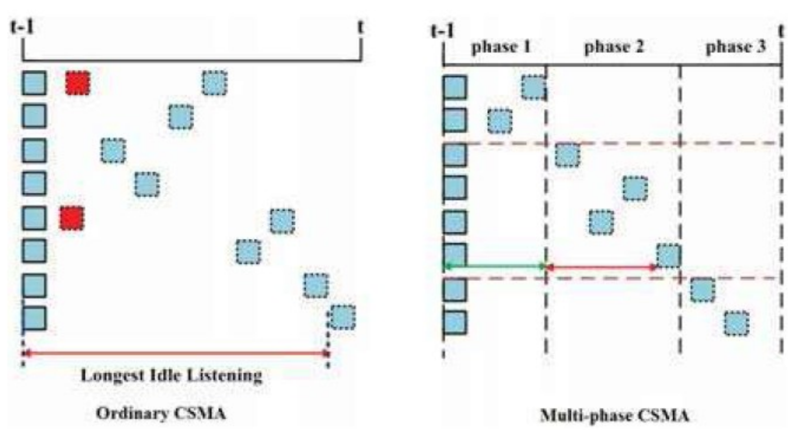


Fig. 1.6. Ordinary CSMA and Multi-Phase CSMA [17]

Results of the simulation showed that the proposed protocol consumes less energy in comparison to dynamic (TDMA). They used the advantages of both contention-based and contention-free MAC protocols and also improved energy efficiency through the use of partial clustering.

For robust hybrid protocol in [18] they used CSMA, TDMA, and IEEE 802.11 DFC protocol. They used DFC for Physical Carrier Sensing and Virtual Carrier Sensing with use of four-way handshake mechanism (RTS-CTS-DATA-ACK) to optimize the transmission period in each TDMA slot. By adding this mechanism they prevent communication failure in case of clock synchronization failure in TDMA protocol.

The system model in [18] consist of M2M domain, network domain, and application domain. In M2M domain, devices communicate with each other through gateways and network domain is responsible for providing reliable and secure channel. The use of DFC mode increases the packet overhead in comparison with TDMA scheme.

They optimized the previous hybrid-MAC protocols by use of the DFC mode and they compared their protocol with s-ALOHA and TDMA and provided simulation results for throughput and average transmission delay. In [18] there is no optimization for the frame structure and there is overhead of acknowledgments.

1.3. Contribution of the Thesis

In this thesis we propose a hybrid MAC protocol for M2M communications. The machines have been divided into different types according to their packet generation processes. The generated traffic has been divided into multiple classes according to their QoS requirements.

We have chosen packet loss probabilities as the main QoS requirement. The transmissions have been organized into frames which have been divided into subframes each serving a class. Each subframe is divided into two sub-periods one serving reservation the other contention traffic of that class. The frame structure of the protocol has been optimized such that QoS requirement of each class is met. The main contributions of this thesis are,

- The proposed MAC protocol handles both periodic and nonperiodic traffic.
- The model captures heterogeneity of the traffic through multiple classes.
- The QoS requirements have been included in the model through packet loss rate constraints.
- The main parameters of the protocol are determined through optimization.
- The proposed protocol is flexible enough to be adapted under changing traffic conditions.
- The model also has considered the effect of the user mobility.

It is believed that the above contributions set apart this research from previous work presented in the literature, survey.

1.4. Organization of the Thesis

The remainder of the thesis is organized as follows.

Chapter 2 describes the proposed MAC protocol. Then, it develops a nonlinear optimization problem to determine the frame length subject to packet loss constraints. It derives per class packet loss rete constraints. It develops linearization of the optimization problem and proposes an ad-hoc algorithm to determine the accuracy of the LP approximation.

In Chapter 3, the proposed MAC protocol is studied under user mobility. The packet arrival process is determined under the assumption that users spent random amount of time in the system. Then the performance of the protocol subject to this arrival process is studied with contention service with and without packet loss. In the former case a packet is lost only when a machine generates a new packet. In the latter case, it is assumed that the packets are served by a global queue with infinite storage and then derived mean packet delay.

In Chapter 4, the service of traffic in a global priority queue has been studied. In this model, the server serves to the traffic according to their relative priorities.

Chapter 5 presents the conclusions of the thesis and discusses future work.

Chapter 2

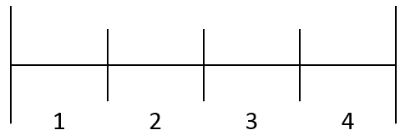
Modelling of a Hybrid MAC Protocol for M2M Communications with Finite User Population

2.1. Introduction

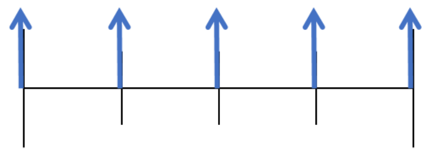
In this chapter we propose a hybrid MAC protocol that differs in a number of ways from the existing hybrid protocols. First of all, we divide traffic into classes, which will enable traffic differentiation according to their packet loss tolerances. Further, each class is divided into two groups as those accessing channel through contention and the others having contention-free access. Splitting of the traffic into classes and channel access types reduces the number of users attempting to use the contention channel simultaneously, thereby increasing system scalability. Here we used TDMA for contention-free channel access and slotted Aloha for the contention-based channel access. Reservation of resources to the users that generates packets almost periodically also improves the service and reduces the contention. Further, optimization of the frame structure ensures that the system is not overloaded and that the user's QoS requirements are satisfied.

2.2. The proposed MAC protocol

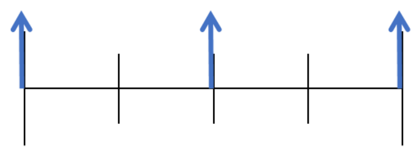
In this section, we describe the proposed MAC protocol for machine access to the channel. We assume that a dedicated channel has been assigned for M2M communications and communications is between machines and serving base station in a cell. We consider only the uplink communications and propose a hybrid MAC protocol that provides both contention and reservation based communications. We assume that transmissions are organized into superframes. A superframe consists of M frames, which are numbered as $m = 1, \ldots, M$. A frame is divided into slots with frame length equaling to T_F slots. The packets have fixed length and the packet transmission time equals to a slot. We assume that each machine may generate one or zero packets periodically. Machines are divided into K types according to their packet generation probabilities during their periods. The superframe length is an integer multiple of machine periods which are also an integer multiple of frame length. The frame and superframe lengths are chosen to correspond to the minimum and maximum of machine periods respectively. Let β denote the period of a machine in number of frames. Fig.2.1i) shows the superframe structure for $M = 4$ frames. In this case, the possible machine periods are $\beta = 1, \ldots, 4$ frames. Fig.s 2.1ii-v) show the packet generation corresponding to machine periods $\beta = 1, \ldots, 4$ respectively. We note that an arrow in the figures corresponds to generation of a packet. However, since the packet generation is probabilistic, a machine may not always generate a packet during the scheduled frame. Fig. 2.1v) we shows only one possibility of arrival for $\beta = 4$ as the single packet could have arrived during any frame of the superframe. We assume that a type i machine will generate a new packet according to an independent Bernoulli trial with probability q_i at the end of its period, which will be transmitted during the next frame. A small value of q_i indicates a non-periodic traffic and a large value, a periodic traffic. A packet that cannot be transmitted during a frame will be discarded, as it is considered to have old information. This also conserves the energy of the machines by limiting the packet transmission attempts as machines may be battery powered. If the measured value by a sensor has not changed from the previous period, then the machine may choose not to generate a packet, which explains the probabilistic packet generation of periodic traffic. The base station will assign the frames to machines according to their periods in a manner that the load is uniformly distributed over a superframe. In the following, without loss of generality, we will analyze the performance of the system in a random frame.



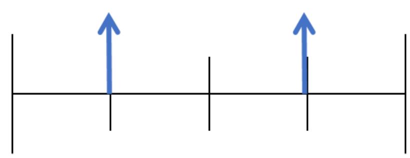
i) Superframe structure with four frames, M=4.



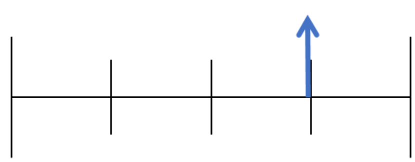
ii) Packet generation with machine period of $\beta = 1$ frames.



iii) Packet generation on odd-frames with machine period of $\beta = 2$ frames.



iv) Packet generation on even frames with machine period of $\beta = 2$ frames.



v) Packet generation with machine period $\beta = 4$ frames.

Fig.2.1 Superframe structure with different machine periods for superframe length of M=4 frames.

The generated traffic will have heterogeneous QoS requirements according to their tolerance to packet loss. As a result, the traffic will be classified into J classes according to their acceptable maximum packet loss probabilities. We assume that a packet generated by a type i machine will belong to class j with probability p_{ij}, which will have maximum packet loss probability of δ_j. We will also constrain total packet loss probability to δ.

From the frame structure shown in Fig. 2.2, each frame will be divided into J subframes one per class with j'th subframe serving j'th class traffic. Further, each subframe will be divided into two sub-periods, one to serve contention and the other one to reserved traffic of that class. During the contention sub-period, the packets will be directly transmitted in the channel according to the slotted Aloha random access protocol. In general, CSMA protocol is preferred to Aloha protocol in random access as it is supposed to have higher throughput. However, this may not be true in M2M communications, since while Aloha throughput is independent of ratio of propagation time to packet transmission time, throughput of CSMA is dependent on this ratio. As this ratio increases CSMA throughput decreases and it may drop below Aloha throughput. In M2M communications, this ratio will be high due to short packet lengths and as a result CSMA throughput advantage over Aloha will disappear [13]. Assuming that there are n users contending to the channel and each one transmitting its packet with probability p, then, the probability of a successful transmission in a slot is given by,

$$P_s = np(1-p)^{n-1} \tag{2.1}$$

The above probability is maximized when each user transmits with probability $p = 1/n$. The lower bound of the maximum value of P_s is $1/e$, which is approached rapidly with increasing value of n. We note that the base station will estimate n from the successfully received packets in each frame and new estimate will be passed to users whenever it significantly differs from the old value to be used in the corresponding frame of the next superframe. In this analysis, we will use this lower bound of the success probability, to obtain a lower bound on the performance of the system. As a result, the number of slots for a successful packet transmission will be geometrically distributed with a mean of $\tau_C = e$ slots.

A machine designated for service in a reserved sub-period will be assigned a slot. Reserved sub-periods will provide TDMA like service. A machine requesting service in the reserved channel will transmit its first packet through the contention subchannel and will also submit its request. If a slot is available in the reserved subchannel, then it will be assigned by the system, and will be used in the transmission of its subsequent packets. It is assumed that users, generating periodic traffic for significant amount of time, will request service in the reserved sub-period.

Let T, T_j denote the durations of a frame and its j'th subframe in number of slots respectively, and T_{Cj}, T_{Rj}, the durations of contention and reserved subperiods in the j'th subframe also in number of slots respectively. The frame length may be expressed as follows,

$$T = \sum_{j=1}^{J} T_j, \qquad \text{where } T_j = T_{Cj} + T_{Rj} \tag{2.2}$$

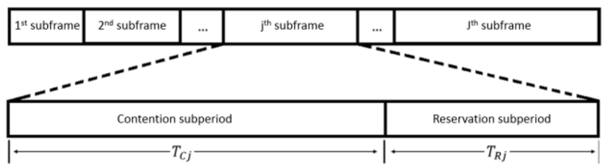


Fig. 2.2. Frame Structure

Table 2.1. summarizes the notation used in this chapter.

The total number of machines expressed in terms of N_i is given by,

$$N = \sum_{i=1}^{K} N_i \tag{2.3}$$

Symbol	Definition
K	number of machine types.
$\overline{N}_{ij}$	average number of type i machines that belong to class j.
N_i	number of type i machines being served during a frame.
N	total number of machines being served during a frame.
q_i	probability that a type i machine will generate a packet during a frame.
τ_c	average number of slots it takes to transmit a packet successfully during a contention period.
p_{ij}	probability that a packet generated by a type i machine will belong to class j.
X_{ij}	number of type i machines that will receive service in the contention subperiod for class j traffic.
Y_{ij}	number of type i machines that will receive service in reserved transmission subperiod for class j traffic.
T_F	maximum frame duration in number of slots.

Table 2.1. SUMMARY of NOTATION

2.3. Optimization Problem

We would like to design the system so that the QoS requirements of the users are met and the frame length is optimized. The users' QoS requirements will be satisfied if the packet loss probabilities of traffic classes are kept below threshold values. Optimization of the frame requires that the traffic is served in a frame of minimal length. In this optimization, there is a tradeoff between serving a packet in a contention or a reserved sub-channel. Service in a contention sub-channel depends on the traffic load. The packets not served during a frame will be discarded, while in the reserved channel the service is guaranteed. However, if a machine to be served in a reserved sub-channel does not generate a packet in a frame, then that slot is wasted. As a result, it is more efficient to serve users with low packet generation probability in a contention sub-channel rather than in a reserved sub-channel.

The system will be designed such that it will be able to serve average load generated during a frame and QoS requirements of the users are met. In this design, the lengths of the contention sub-periods will be chosen such that their lengths will be able to serve average load assigned to them. If the QoS requirements are not satisfied, then one of the following two solutions can be used:

a) the lengths of the contention sub-periods may be adjusted to meet the requirements. The adjustment may be done through a proportionality factor g_j, $j = 1,2,\ldots,J$. $g_j = 1$ corresponds to a contention sub-period length set to serve average of load assigned to it. $g_j < 1$ corresponds to the contraction and $g_j > 1$ corresponds to the expansion of the contention sub-period relative to its $g_j = 1$ length.

b) the lengths of the reservation sub-periods are increased and so, some of the traffic is transferred from the contention to the reservation sub-periods.

Next, we present the optimization problem that minimizes the frame length subject to the QoS requirements of user traffic.

$$\text{Minimize} \quad [\tau_C \sum_{j=1}^{J} g_j \sum_{i=1}^{K} q_i X_{ij} + \sum_{j=1}^{J} \sum_{i=1}^{K} Y_{ij}] \tag{2.4}$$

subject to,

$$X_{ij} + Y_{ij} \geq N_i p_{ij}, \qquad i = 1,2,\ldots,K,\ j = 1,2,\ldots,J \tag{2.5}$$

$$P_{\ell_j} < \delta_j\,, \quad j = 1,2,\ldots,J \quad ; \qquad P_\ell < \delta \tag{2.6a,b}$$

$$T = \tau_C \sum_{j=1}^{J} g_j \sum_{i=1}^{K} q_i X_{ij} + \sum_{j=1}^{J} \sum_{i=1}^{K} Y_{ij} \tag{2.7}$$

$$T \leq T_F \tag{2.8}$$

In the above optimization g_j, X_{ij}, Y_{ij} are decision variables to be determined by the optimization.

In the above, equation (2.4) gives the objective function of the optimization problem which corresponds to the frame length. The first term of the objective function corresponds to the total amount of time, in a frame, allocated to serve contention traffic and the second term corresponds to the time allocated to the reserved traffic. The constraints in (2.5) ensure that the average number of packets that may be generated during a frame will be served and the constraints in (2.6a,b) guarantee that the packet loss probabilities of each class as well as the total packet loss probabilities are kept below their upper bounds, δ_j and δ respectively. The equation (2.7) determines the needed minimum frame length, T, for the transmission of the traffic load such that QoS requirements are met. The constraint in (2.8) ensures that the needed frame length is smaller than the minimum machine traffic period.

The solution of the above optimization problem will determine the frame structure: frame duration, durations of subframes allocated to each traffic class, and sub-period durations in each subframe. Each user will know the superframe structure, the frames that had been scheduled to transmit, as well as, the subframe and the sub-period structure of those frames. The system will not change the frame structure unless changes occur in machine population size, in machine packet generation probabilities or in packets' loss probabilities. The system will monitor the traffic characteristics to determine if the frame structure needs to be re-optimized. In the next section, we will derive the packet loss expressions for the constraints in (2.6a,b).

2.4. Derivation of Packet Loss Probability

While there cannot be any packet losses in the transmission of reserved traffic, contention traffic may experience packet losses. Since durations of the contention subperiods have been determined based on the expected traffic load, it is possible that machines generate more packets than the transmit capacities of the contention subchannels. This will result in packets that could not be transmitted during a frame, which will be lost. In this section, we will derive the packet loss probability of each traffic class as well as the total packet loss probability needed in the constraints (2.6a, b). Let us introduce the following notation for this derivation,

X_{Cj}: number of slots allocated for class j contention sub-period in a frame.

Y_{Rj}: number of slots allocated for class j reserved sub-period in a frame.

r_{ij}: number of class j contention packets generated by type i machines during a frame.

r_j: number of class j contention packets generated during a frame.

k_j: number of class j packets that may be successfully transmitted during class j contention sub-period.

n_{T_j}: total number of class j packets generated during a frame.

n_T : total number of packets generated during a frame.

P_{r_j}: probability that r_j packets will be generated during a frame.

P_{k_j}: probability that k_j successful packet transmissions will occur during the class j contention sub-period.

ℓ_j: number of packet losses in class j contention subperiod in a frame.

ℓ: total number of packets losses in contention subperiods in a frame.

P_{ℓ_j}: probability of class j packet loss.

P_ℓ: probability of packet loss.

First, we derive the packet loss probability of class j packets, defined as the ratio of the average number of packet losses to average number of class j packets generated during a frame,

$$P_{\ell_j} = \frac{E[\ell_j]}{E\left[n_{T_j}\right]} \ , \tag{2.9}$$

where, $E[n_{T_j}] = \sum_{i=1}^{K} q_i N_i \, p_{ij}$

Next, we determine the average number of class j packet losses, $E\left[\ell_j\right]$. We note that r_{ij} has Binomial distribution given below,

$$P_{r_{ij}} = \binom{X_{ij}}{r_{ij}} (q_i)^{r_{ij}} (1 - q_i)^{X_{ij} - r_{ij}} \tag{2.10}$$

Let us use normal approximation for the above Binomial distribution with mean and variance given by,

$$\mu_{r_{ij}} = X_{ij}\, q_i\,, \qquad \sigma^2_{r_{ij}} = X_{ij}\, q_i(1 - q_i) \tag{2.11}$$

From the above definitions,

$$r_j = \sum_{i=1}^{K} r_{ij} \tag{2.12}$$

Since r_{ij} are independently distributed normal random variables, then r_j will also have a normal distribution with mean and variance given by,

$$\mu_{r_j} = \sum_{i=1}^{K} \mu_{r_{ij}}, \qquad \sigma^2_{r_j} = \sum_{i=1}^{K} \sigma^2_{r_{ij}} \tag{2.13}$$

Since the probability of successful transmission during a contention slot is given by P_s, the distribution of the number of packets that may be transmitted during a class j contention subperiod is given by the following Binomial distribution,

$$P_{k_j} = \binom{X_{Cj}}{k_j} P_s^{k_j} (1 - P_s)^{X_{Cj} - k_j}\,, \qquad k_j = 0,1,\ldots,X_{Cj}$$

(2.14)

where,

$$X_{Cj} = \left\lceil g_j \tau_C \sum_{i=1}^{K} q_i X_{ij} \right\rceil\,, j = 1,2,\ldots,J \tag{2.15}$$

The mean and variance of the random variable k_j is given by,

$$\mu_{k_j} = X_{Cj} P_s, \quad \sigma^2_{k_j} = X_{Cj} P_s (1 - P_s) \tag{2.16}$$

Let us define z_j as,

$$z_j = r_j - k_j \tag{2.17}$$

Note that the positive values of z_j indicates the packet losses. We Let $f_{z_j}(x)$ denote the pdf of the random variable z_j, then z_j will have a normal distribution with mean and variance given by,

$$\mu_{z_j} = \mu_{r_j} - \mu_{k_j}, \qquad \sigma^2_{z_j} = \sigma^2_{r_j} + \sigma^2_{k_j} \tag{2.18}$$

In the above substituting from (2.13, 2.16) as well as setting $\tau_C = e$ and $P_s = 1/e$, result in,

$$\mu_{z_j} = (1 - g_j) \sum_{i=1}^{K} q_i X_{ij} \tag{2.19}$$

$$\sigma_{z_j}^2 = \sum_{i=1}^{K} X_{ij}\, q_i [(1 - q_i) + g_j (e - 1)/e] \tag{2.20}$$

We can now determine the expected packet loss for class j

$$E[\ell_j] = E[z_j, z_j > 0] = \int_0^{\infty} x f_{z_j}(x) dx \tag{2.21}$$

$$E[\ell_j] = \frac{1}{\sqrt{2\pi}\sigma_{z_j}} \int_0^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu_{z_j}}{\sigma_{z_j}}\right)^2} dx \tag{2.22}$$

The above may be written as,

$$E[\ell_j] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} (\mu_{z_j} + \sigma_{z_j} u) e^{-\frac{1}{2}u^2} du \tag{2.23}$$

where,

$$u = \frac{x-\mu_{z_j}}{\sigma_{z_j}}, \qquad \alpha = -\frac{\mu_{z_j}}{\sigma_{z_j}} \tag{2.24}$$

Following the integration in (2.23)

$$E[\ell_j] = \mu_{z_j}[1 - \Phi(\alpha)] + \frac{1}{\sqrt{2\pi}} \sigma_{z_j} e^{-\frac{1}{2}\alpha^2} \tag{2.25}$$

where,

$$\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{1}{2}u^2} du \tag{2.26}$$

Substituting the above in (2.9) determines the class j packet loss probability,

$$P_{\ell_j} = \frac{1}{\sum_{i=1}^{K} q_i N_i p_{ij}} \left\{ \mu_{z_j}[1 - \Phi(\alpha)] + \frac{1}{\sqrt{2\pi}} \sigma_{z_j} e^{-\frac{1}{2}\alpha^2} \right\} \tag{2.27}$$

Next, we determine the expression for the total packet loss probability defined below,

$$P_\ell = \frac{E[\ell]}{E[n_T]}, \tag{2.28}$$

where,

$$E[\ell] = \sum_{j=1}^{J} E[\ell_j]\ , \quad E[n_T] = \sum_{i=1}^{K} q_i N_i \tag{2.29}$$

Substituting from (2.25), we conclude that,

$$P_\ell = \frac{1}{\sum_{i=1}^{K} q_i N_i} \ \sum_{j=1}^{J} \left[\mu_{z_j}[1 - \Phi(\alpha)] + \frac{1}{\sqrt{2\pi}} \sigma_{z_j} e^{-\frac{1}{2}\alpha^2} \right] \tag{2.30}$$

2.5. Linearization of the optimization problem

The optimization problem formulated in the previous two sections is nonlinear. The objective function in (2.4) and the expressions determined in (2.27, 2.30) for packet loss constraints in (2.6a,b) are nonlinear because they contain products of the decision variables as well as other nonlinear terms. Since the solution of nonlinear optimization problems is difficult, we would like to linearize our problem to benefit from the widely available linear programming (LP) solution tools such as IBM ILOG CPLEX Optimization Studio. For constant values of $g_j,\ j = 1,2,\ldots,J$ the objective function in (2.4) becomes linear. As a result, we linearize the objective function in (2.4) by setting $g_j = 1,\ j = 1,2 \ldots, J$. This means that the length of the contention subperiods will be chosen such that the average contention traffic load may be served. As a result of this choice, μ_{z_j} given by (2.19) and as a consequence of this, α defined in (2.24) become zeros. These reduce (2.27, 2.30) to,

$$P_{\ell_j} = \frac{1}{\sum_{i=1}^{K} q_i N_i p_{ij}} \left[\frac{1}{\sqrt{2\pi}} \sigma_{z_j} \right] \tag{2.31}$$

$$P_\ell = \frac{1}{\sum_{i=1}^{K} q_i N_i} \ \sum_{j=1}^{J} \left[\frac{1}{\sqrt{2\pi}} \sigma_{z_j} \right] \tag{2.32}$$

The above expressions are still nonlinear, because of the square-root operation in determining σ_{z_j}. First, let us consider further simplification of (2.31). Next, we square both sides of the equation (2.31),

$$P_{\ell_j}^2 = \frac{\sigma_{z_j}^2}{2\pi\left[\sum_{i=1}^{K} q_i N_i p_{ij}\right]^2} \tag{2.33}$$

$P_{\ell_j}^2$ is linear in the decision variables. Thus in the optimization we will use the following linear constraint,

$$\frac{\sigma_{z_j}^2}{2\pi\left[\sum_{i=1}^{K} q_i N_i p_{ij}\right]^2} \leq \delta_j^2 \tag{2.34}$$

to satisfy the constraint $P_{\ell_j} < \delta_j$ in (2.6a). We note that the above linearization does not involve any approximation.

Next, we will determine the bound for the total packet loss rate. We note that since in the Binomial distributions in (2.10) and (2.14), the number of trials are very large, we can assume that $\sigma_{r_j}^2 \geq 1$, $\sigma_{k_j}^2 \geq 1$, and as a result from (2.18), $\sigma_{z_j}^2 \geq \sigma_{z_j} \geq 1$. Taking the square of the total packet loss rate in (2.6b), we have,

$$P_\ell^2 \leq \delta^2 \tag{2.35}$$

From (2.32), we have,

$$P_\ell^2 = \frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[\sum_{j=1}^{J} \sigma_{z_j}\right]^2 \tag{2.36}$$

$$P_\ell^2 = \frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[\sum_{j=1}^{J} \sum_{k=1}^{J} \sigma_{z_j} \sigma_{z_k}\right] \tag{2.37}$$

The above may be bounded as follows,

$$P_\ell^2 \leq \frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[\sum_{j=1}^{J} \sum_{k=1}^{J} (\max(\sigma_{z_j}, \sigma_{z_k}))^2\right] \tag{2.38}$$

As a result, (2.35) will be satisfied if the following is satisfied,

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[\sum_{j=1}^{J} \sum_{k=1}^{J} (\max(\sigma_{z_j}, \sigma_{z_k}))^2\right] \leq \delta^2 \tag{2.39}$$

Next we give the set of constraints that will be equivalent to the above for the case of $J = 3$, that will be used in the numerical results section. There are six cases to be considered depending on the relative values of $\sigma_{z_1}, \sigma_{z_2}, \sigma_{z_3}$

i) $\sigma_{z_1} \geq \sigma_{z_2} \geq \sigma_{z_3}$ ii) $\sigma_{z_1} \geq \sigma_{z_3} \geq \sigma_{z_2}$

iii) $\sigma_{z_2} \geq \sigma_{z_1} \geq \sigma_{z_3}$ iv) $\sigma_{z_2} \geq \sigma_{z_3} \geq \sigma_{z_1}$

v) $\sigma_{z_3} \geq \sigma_{z_1} \geq \sigma_{z_2}$ vi) $\sigma_{z_3} \geq \sigma_{z_2} \geq \sigma_{z_1}$

Each of the above cases is taken care of by the corresponding constraints below,

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_1}^2 + 2\sigma_{z_2}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40i}$$

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_1}^2 + 2\sigma_{z_3}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40ii}$$

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_2}^2 + 2\sigma_{z_1}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40iii}$$

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_2}^2 + 2\sigma_{z_3}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40iv}$$

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_3}^2 + 2\sigma_{z_1}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40v}$$

$$\frac{1}{2\pi\left[\sum_{i=1}^{K} q_i N_i\right]^2} \left[4\sigma_{z_3}^2 + 2\sigma_{z_2}^2 + \sum_{j=1}^{J} \sigma_{z_j}^2\right] \leq \delta^2 \tag{2.40vi}$$

The above gives us a linear set of constraints in the decision variables. This will ensure that the total packet loss constraint $P_\ell < \delta$ in (2.6b) is satisfied.

Next, we will determine the accuracy of setting $g_j = 1,\ j = 1,2,\ldots,J$ in the optimization, through the following ad-hoc algorithm. The ad-hoc algorithm solves the nonlinear optimization problem in the three steps described below,

i) In the first step, the optimization problem in (2.4-2.8) is solved by setting $g_j = 1,\ j = 1,2,\ldots,J$ and without the packet loss constraints given by (2.6a, b). Let $\hat{X}_{ij}, \hat{Y}_{ij}$ denote the values of the decision variables for the contention and reservation traffic in this solution.
ii) Calculate per class and total packet loss rates and frame length P_{ℓ_j}, P_ℓ, T for increasing values of $g_j,\ j = 1,2,\ldots,J$ for the values of decision variables $\hat{X}_{ij}$ obtained in step i) until $P_{\ell_j} < \delta_j$ for $j = 1,2,\ldots,J$ are satisfied.
iii)Let $\hat{g}_j, j = 1,2,\ldots,J$ denote the minimum values of g_j satisfying the packet loss requirements in the desired solution. Let also $\hat{P}_{\ell_j}, \hat{P}_\ell, \hat{T}$ denote, the per class and total packet loss rates and the frame length of this solution.

The optimality of the LP approximation will be determined through comparison of the packet loss probabilities and the frame length of LP solution with $\hat{P}_{\ell_j}, \hat{P}_\ell\ \hat{T}$ obtained from the ad-hoc algorithm.

2.6. Numerical Results

In this section, we present some numerical results regarding the analysis of the proposed MAC protocol. We assume a system with 5 types of machines and 3 classes of packets. Table 2.2 shows the number of machines per type and the packet generation probability of each type of machine. Table 2.3 presents the classes of the packets generated by the machines of each type. Table 2.4 presents the average number of machines from each type in each class, $\overline{N}_{ij} = N_i p_{ij}, i = 1,2, \ldots, K, j = 1,2, \ldots, J$. We note that the total number of the machines in the system equals to $N = 3350$. Also, maximum frame length equals to N, which corresponds to the assignment of one reservation slot to each machine.

	Machine types				
	i=1	i=2	i=3	i=4	i=5
N_i	500	750	450	1000	650
q_i	0.35	0.8	0.3	0.1	0.2

Table 2.2. Number of machines per type and packet generation probability of each type of machine.

Packet class probability	Machine types				
	i=1	i=2	i=3	i=4	i=5
p_{i1}	0.3	0.1	0.1	0.2	0.3
p_{i2}	0.1	0.4	0.1	0.1	0.1
p_{i3}	0.6	0.5	0.8	0.7	0.6

Table 2.3. Class probabilities of packets generated by machines of each type.

j	$\overline{N}_{1j}$	$\overline{N}_{2j}$	$\overline{N}_{3j}$	$\overline{N}_{4j}$	$\overline{N}_{5j}$
1	150	75	45	200	195
2	50	300	45	100	65
3	300	375	360	700	390

Table 2.4. Average number of type i machines generating class j traffic, $\overline{N}_{ij}, i = 1,2, \ldots, K, j = 1,2, \ldots, J$.

Table 2.5 presents LP solution of the optimization problem for per class packet loss rate constraints. The classes 1, 2 and 3 have been assigned maximum packet loss rates of $\delta_1 = 0.1\%, \delta_2 = 0.5\%, \delta_3 = 1\%$ respectively. The table presents the number of per type machines assigned to receive service from each class in the contention and reservation subperiods with one row per class. From row 1, all the machines in class 1 have been assigned to reservation service because of the very low packet loss rate requirement. From row 2, the machines of all types in class 2 also will receive the reservation service except the type 4 machines. Finally, in class 3, the machines of types 1 and 2 will receive the reservation service, types 4 and 5, the contention service while machines of type 3 have been split between the contention and the reservation services. We note that sum of the machines in the two subperiods for each machine type equals to the average number of machines for that type per class given in Table 2.3, $\overline{N}_{ij} = X_{ij} + Y_{ij}, i = 1,2, \dots, K, j = 1,2, \dots, J$. Table 2.6 presents the packet loss rate of each class as well as the total packet loss rate from the LP solution. It may be seen that the solution satisfies the maximum packet loss rate constraints of each class, $P_{\ell_j} < \delta_j, j = 1,2, \dots, J$. Solution results in total packet loss rate of $P_\ell = 0.7119$. The table also presents the frame length determined by the solution, $T = 2572$ slots, which is significantly lower than the worst case frame length of $N = 3350$ slots.

j	$\delta_j(\%)$	X_{1j}	Y_{1j}	X_{2j}	Y_{2j}	X_{3j}	Y_{3j}	X_{4j}	Y_{4j}	X_{5j}	Y_{5j}
1	0.1	0	150	0	75	0	45	1	199	0	195
2	0.5	0	50	0	300	0	45	88	2	0	65
3	1.0	0	300	0	375	139	221	700	0	390	0

Table 2.5. Number of machines from each type and class assigned to receive service by LP in contention and reservation subperiods subject to class packet loss requirements, $\delta_j, j = 1,2,3$.

j	δ_j(%)	$P_{\ell j}$ (%)	P_ℓ (%)	T(slots),
1	0.1	0.0634	0.7119	2572
2	0.5	0.4937		
3	1.0	0.9906		

Table 2.6. Per class and total packet loss rate and frame length in slots corresponding to LP assignment given in Table 2.5.

Tables 2.7 and 2.8 present the solution corresponding to Tables 2.5 and 2.6 respectively for the same system but subject to not only per class packet loss rate constraints but also to total packet loss requirement, $\delta = 0.5$. Since the value of δ is smaller than the total packet loss rate $P_\ell = 0.7119$ in Table 2.6, this new requirement will force the transfer of more machines from the contention service to the reservation service in order to meet the lower total packet loss requirement. The comparison of X_{i3} values for i=3, 4, 5 in Tables 2.5 and 2.7 confirm these machines transfer. As may be seen, this solution satisfies the total packet loss requirement, $P_\ell = 0.336 \leq \delta$ but at the expense of higher frame length of 3131 slots.

j	δ_j(%)	X_{1j}	Y_{1j}	X_{2j}	Y_{2j}	X_{3j}	Y_{3j}	X_{4j}	Y_{4j}	X_{5j}	Y_{5j}
1	0.1	0	150	0	75	0	45	1	199	0	195
2	0.5	0	50	0	300	0	45	88	12	0	65
3	1.0	0	300	0	375	0	360	213	487	0	390

Table 2.7. Number of machines from each type and class assigned to receive service by LP in contention and reservation subperiods subject to both class packet loss requirements, δ_j, as well as total packet loss requirement $\delta = 0.5\%$.

j	δ_j (%)	$P_{\ell j}$ (%)	P_{ℓ}(%)	T(slots)
1	0.1	0.0634	0.336	3131
2	0.5	0.4937		
3	1.0	0.3421		

Table 2.8. Per class and total packet loss rates and frame length corresponding to LP assignment given in Table 2.7.

Fig. 2.3 plots the packet loss rates of each class as well as the total packet loss rate as a function of the maximum total packet loss rate, δ, for the maximum per class packet loss rates of $\delta_1 = 0.1\%, \delta_2 = 0.5\%, \delta_3 = 1\%$. The regions of almost linearly increasing per class packet loss probabilities correspond to receiving contention service, while flat regions to receiving reserved service. The total packet loss probability is initially close to δ but it underestimates it as δ increases because of the approximation in the bound. Fig. 2.4 plots the resulting frame length as a function of δ for this system. We note that the results presented in Tables 2.7 and 2.8 correspond to the point $\delta = 0.5$ in Fig.s 2.3 and 2.4. For the small values of δ, the total packet loss constraint is binding and for large values, the per class packet loss constraints are binding. In Fig. 2.4, frame length decreases as δ increases and it is determined initially by δ and later on by δ_j.

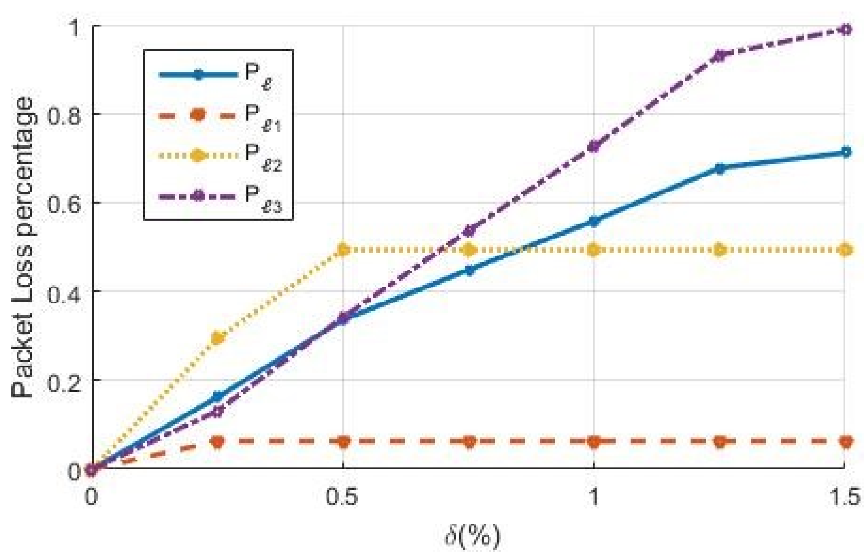


Fig. 2.3. Per class and total packet loss rate as a function of maximum total packet loss rate, $\boldsymbol{\delta}$, for values of $\delta_j = [0.1\%, 0.5\%, 1\%]$.

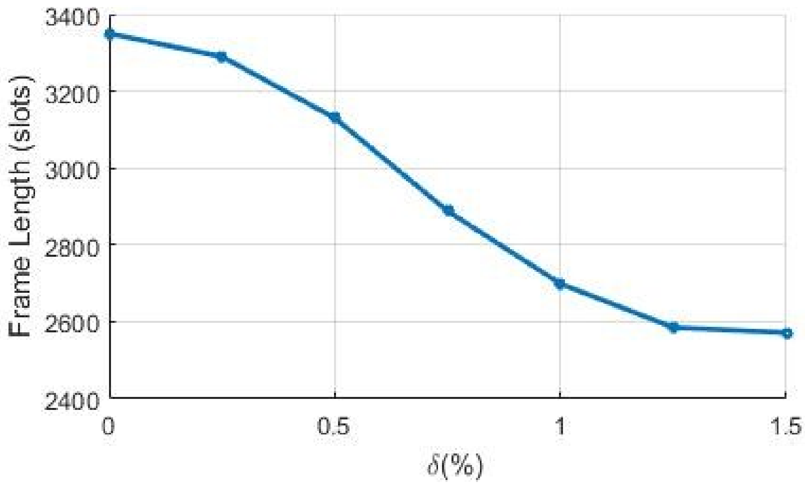


Fig. 2.4. Frame length as a function of total packet loss rate, $\boldsymbol{\delta}$, for values of $\delta_j = [0.1\%, 0.5\%, 1\%]$.

Next, we present the solution of the nonlinear optimization problem through the ad-hoc algorithm for the system under consideration. Step i) of the algorithm solves the optimization problem by setting $g_j = 1, j = 1,2,3$ and without packet loss constraints. Table 2.9 presents per type per class machine assignment resulting from this solution, $\hat{X}_{ij}, \hat{Y}_{ij}$ respectively. In Table 2.10, we present per class and total packet loss probabilities for increasing values of g_j as well as the resulting frame length for the values of decision variables obtained in Table 2.9. For this example we assumed the same g_j value for all the classes. It may be seen that as g_j value increases the packet loss probabilities decrease and the frame length increases. It may be seen that the packet loss constraints $\hat{P}_{\ell_j} < \delta_j$ are satisfied at $\hat{g}_j = 1.2$ with a frame length of $\hat{T}$=2511. This frame length is slightly lower than the frame length given by the LP solution of the problem given in Table 2.6, $T = 2572$ slots. However, this solution uses longer contention subperiods than the LP solution. Longer contention periods result in higher power consumption, which is not desirable. As a result, we conclude that LP gives nearly optimal solution, which is easy to obtain.

j	$\hat{X}_{1j}$	$\hat{Y}_{1j}$	$\hat{X}_{2j}$	$\hat{Y}_{2j}$	$\hat{X}_{3j}$	$\hat{Y}_{3j}$	$\hat{X}_{4j}$	$\hat{Y}_{4j}$	$\hat{X}_{5j}$	$\hat{Y}_{5j}$
1	150	0	0	75	45	0	200	0	195	0
2	50	0	0	300	45	0	100	0	65	0
3	300	0	0	375	360	0	700	0	390	0

Table 2.9. Number of per type per class machine assignment, $\hat{X}_{ij}, \hat{Y}_{ij}$, for $g_j, j = 1,2,3$, without packet loss constraints.

g_j	j	$P_{\ell j}$ (%)	P_{ℓ}(%)	T(slots)
0.5	**1**	33.7624	23.6499	1484
	2	9.1079		
	3	27.2876		
1	**1**	2.8056	1.5231	2218
	2	1.1573		
	3	1.3269		
1.2	**1**	0.0956	0.0571	2511
	2	0.1597		
	3	0.0008		
1.5	**1**	1.7514E-05	5.0054E-04	2952
	2	0.0019		
	3	8.059E-14		

Table 2.10. Per class and total packet loss rates and frame length for different values of the contention subperiods proportionality factor, g_j, for values of decision variables $\hat{X}, \hat{Y}$ determined in Table 2.9.

Next, we present a sample calculation of frame duration for a given channel transmission rate and packet length. Assuming a channel transmission rate of 10Mbps and a packet length of 100bytes, it results in a slot duration of $80\mu sec$. The frame length of the system under consideration, $T = 2572$ slots, results in frame duration of $205.76\ msec$, which is not too long.

Fig. 2.5 presents the frame length as a function of increasing packet generation probabilities in different step sizes for all the machine types. Let $q_{i,k}$ denote the packet generation probability of type i machines in step k. Then, $q_{i,k}$ is given by,

$$q_{i,k} = \begin{cases} q_{i,k-1} + i\Delta q, & q_{i,k-1} < 1, \quad k \geq 1 \\ 1 & q_{i,k-1} = 1, \quad k \geq 1 \end{cases}$$

where $q_{i,0} = 0$ for i=1…5, $\Delta q = 0.05$ and $(i\Delta q)$ is the step size for type i machines. The frame length has been determined using the LP approximation and without the packet loss constraints. As may be seen from the figure, the frame length increases with the increasing number of steps. Initially, all the traffic types are being served by the contention subchannels. As the packet generation probability increases, the traffic of each type switches from contention to reserved subchannels. Thus eventually, all the traffic is served by the reserved subchannels and the frame length does not increase anymore with the step number value, k.

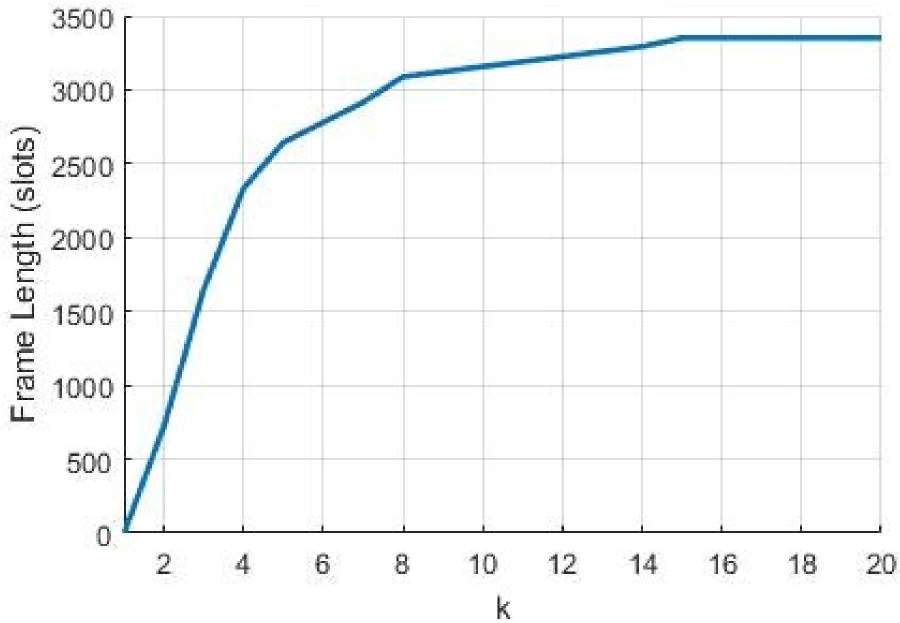


Fig. 2.5. Frame length as a function of packet generation probabilities of all the machine types without packet loss constraints.

2.7. Conclusion

In this chapter, we proposed a scalable hybrid MAC protocol for M2M communications. The protocol is designed to serve multiple priority traffic classes and meet their QoS requirements. The protocol is able to serve both periodic and nonperiodic traffic. Each traffic class is served in a dedicated subframe during a frame, and each subframe is divided into contention and reservation subperiods. Contention subperiod serves nonperiodic traffic, while reservation subperiod serves periodic traffic and traffic with low packet loss requirements. We developed a nonlinear optimization problem that determines the frame structure of this protocol subject to packet loss requirements. We have derived packet loss probabilities of different traffic classes as well as total packet loss probability. We have provided an LP approximation to this optimization problem. Numerical results show that LP provides nearly optimal solution and the system will be able to serve large machine populations.

Chapter 3

Modelling of the Hybrid MAC Protocol for M2M Communications with Mobile Users

3.1. Introduction

In this chapter we consider modelling of the hybrid MAC protocol for M2M communications introduced in the previous chapter with mobile users. We assume that the users (machines) arrive at the system according to Poisson process and spend random amount of time in the system. We consider multiple types of users with different packet generation probabilities.

As in the previous chapter, we assume different classes of machines. A machine may request either reservation or contention service. It is assumed that transmissions are organized into frames. A frame will be divided into reservation and contention interval. For the sake of simplicity, we assume that reservation and contention interval lengths are given and we will not optimize frame length. A machine requiring reservation service will be either assigned a slot of service or will be denied if no bandwidth is available. For this serve, we determine blocking probability of reservation request.

The contention traffic may be divided into different classes according to their packet loss tolerances, but for simplicity we will assume a single class of contention traffic. We will consider two service disciplines for contention traffic, which are

i. A packet remains in the system until it is either successfully transmitted or the machine generates a new packet. This is an extension of the contention service of the previous chapter, where a packet transmission was attempted only in a single frame. For this service discipline, we determine packet loss probability.
ii. A packet remains in the system until it is successfully transmitted. For this service discipline we determine mean packet delay.

3.2. Reservation Service

We assume that a machine requesting service will be assigned a slot if slots are available for the duration of machine's residency time in the network and otherwise request will be denied. We assume that the arrival of the machines requesting reservation service is according to a Poisson process with parameter Λ machines/sec. The machines are divided into K types and a new arriving machine belongs to type i with probability γ_i and the distribution of its residency time in the network has a rational Laplace transform with mean $1/\beta_i$. Let C denote the total number of reservation slots and $u(\cdot)$ denote the steady-state probability distribution of the number of busy slots. Since this is a blocking network, from [32], then $u(\cdot)$ satisfies the following recursive equation,

$$\sum_{i=1}^{K} a_i u(j-c_i) = ju(j) \qquad j = 0, 1, \dots, C \qquad (3.1)$$

where a_i is the traffic load of type i machines and is given by,

$$a_i = \frac{\gamma_i \Lambda}{\beta_i} \qquad (3.2)$$

and c_i is the number of slots assigned to each accepted request. Since, $c_i = 1$, (3.1) results in

$$\sum_{i=1}^{K} a_i u(j-1) = ju(j) \qquad j = 0, 1, \dots, C \qquad (3.3)$$

In the above recursion we need to also use the normalization condition, $\sum_{j=0}^{C} u(j) = 1$, and set $u(x) = 0$ for negative values of x.

Let P_{B_i} denote probability that reservation request of a type i machine will be blocked, then it is given by,

$$P_{B_i} = \sum_{\ell=0}^{c_i-1} u(C-\ell) \qquad i = 1,2,3 \tag{3.5}$$

Since $c_i = 1$, probability of blocking for all machine types is given by,

$$P_{B_i} = u(C) \qquad i = 1,2,3 \tag{3.6}$$

3.3. Contention Service

Next, we will consider contention service with/without discarding of the packets. We assume that machines requesting contention service arrive at the network according to a Poisson process with parameter λ machines/sec. For service with packet discarding, we will determine packet loss probability and for serving without discarding we will determine mean packet delay. The machines are divided into K types and a new arriving machine belongs to type-i with probability γ_i and distribution of its residency time in the network has a rational Laplace transform with mean $1/\beta_i$ sec. Let us define ρ_i as the traffic load of type i machine.

The ρ_i is given by,

$$\rho_i = \frac{\gamma_i \lambda}{\beta_i}.$$

The network may be modeled as an infinite server type single node BCMP network [32]. Let n_i denote the number of type i machines in the system, then joint probability distribution of the number of machines in the system is given by,

$$P(\bar{n}) = P(n_1, n_2, \dots, n_i, \dots, n_K) = \frac{1}{G} \prod_{i=1}^{K} \frac{{\rho_i}^{n_i}}{n_i!} \tag{3.7}$$

Where G is the normalization constant. We evaluate G from the normalization condition as below.

$$\sum_{n_1=0}^{\infty} \dots \sum_{n_i=0}^{\infty} \dots \sum_{n_K=0}^{\infty} P(\bar{n}) = \sum_{n_1=0}^{\infty} \dots \sum_{n_i=0}^{\infty} \dots \sum_{n_K=0}^{\infty} \left[\prod_{i=1}^{K} \frac{{\rho_i}^{n_i}}{n_i!} \right] = 1 \tag{3.8}$$

Substituting from (3.7) in the above results in,

$$G = \prod_{i=1}^{K} \left[\sum_{n_1=0}^{\infty} \dots \sum_{n_i=0}^{\infty} \dots \sum_{n_K=0}^{\infty} \frac{{\rho_i}^{n_i}}{n_i!} \right] = e^{\sum_{i=1}^{K} \rho_i} \quad (3.9)$$

As a result, the joint distribution of the number of machines from (3.7) is given by,

$$P(\bar{n}) = P(n_1, n_2, \dots, n_i, \dots, n_K) = e^{-\sum_{i=1}^{K} \rho_i} \prod_{i=1}^{K} \frac{{\rho_i}^{n_i}}{n_i!} \quad (3.10)$$

From the above mean number of type-i machines in the system is given by $E[n_i] = \rho_i$. Let us define n as the total number of machines in the system, then its mean value is given by,

$$E[n] = \sum_{i=1}^{K} \rho_i \quad (3.11)$$

Next we will determine distribution of the total number of new packets generated by the machines in the system. Let us define,

r_i : number of packets generated by type-i machines during a frame.

r : total number of packets generated by all the machines during a frame.

Then, we have,

$$r = \sum_{i=1}^{K} r_i$$

Let $r_i(z)$ and $r(z)$ denote probability generating functions, PGFs, of the distributions of r_i and r respectively, then, since machines are independent of each other,

$$r(z) = \prod_{i=1}^{K} r_i(z) \quad (3.12)$$

As a type-i machine generates a packet according to independent Bernoulli trials with parameter q_i during a frame,

$$r_i(z|n_i) = (q_i z + 1 - q_i)^{n_i} \quad (3.13)$$

Then, $r(z)$ is given by,

$$r(z) = \sum_{n_1=0}^{\infty} \dots \sum_{n_i=0}^{\infty} \dots \sum_{n_K=0}^{\infty} \prod_{i=1}^{K} r_i\,(z|n_i) P(\bar{n}) \quad (3.14)$$

Substituting from (3.13) in (3.14) and interchanging the order of summation and multiplication results in,

$$r(z) = e^{-\sum_{i=1}^{K} \rho_i} \prod_{i=1}^{K} \left[\sum_{n_i=0}^{\infty} (q_i z + 1 - q_i)^{n_i} \frac{{\rho_i}^{n_i}}{n_i!} \right] \quad (3.15)$$

$$r(z) = e^{-\sum_{i=1}^{K}\rho_i}\prod_{i=1}^{K} e^{(q_i z+1-q_i)\rho_i} \quad (3.16)$$

which gives,

$$r(z) = e^{\sum_{i=1}^{K}(q_i\rho_i)(z-1)} \quad (3.17)$$

As can be seen the total number of new packets generated during a frame has a Poisson distribution with parameter $\sum_{i=1}^{K}(q_i\rho_i)$. The mean and variance of this distribution are given by,

$$\mu_r = \sum_{i=1}^{K} q_i\rho_i\,,\ \sigma_r^2 = \sum_{i=1}^{K} q_i\rho_i \quad (3.18)$$

3.3.1.Contention Service with packet losses

Next, we consider contention service that a packet is discarded if not it is transmitted successfully by the time that its machine generates a new packet. We will determine packet loss probability of this service.

3.3.1.i. Derivation of packet loss probability in a frame

First, we will determine packet loss probability for the packet arrival process determined in the above. This packet loss probability adopts the derivation in section 2.4 to the above packet arrival process.

Let us introduce the following notation for this derivation,

T_F : frame length in number of slots.

k: number of packets that may be successfully transmitted during a frame.

P_k: probability that k successful packet transmissions will occur during a frame.

g : total number of packets that have not been successfully transmitted during a frame.

P_g: probability that a packet willnot be successfully transmitted during a frame.

Since probability of successful transmission during a contention slot is given by p_s, distribution of the number of packets that may be transmitted during a frame follows Binomial distribution,

$$P_k = \binom{T_F}{k} p_s^k (1-p_s)^{T_F-k}\,, k = 0,1,\dots,T_F \quad (3.19)$$

The mean and variance of the random variable k are given by,

$$\mu_k = T_F p_s, \quad \sigma_k^2 = T_F p_s(1 - p_s) \tag{3.20}$$

Let us define *m* as the difference of the random variables *r* and *k*.

$$m = r - k \tag{3.21}$$

r has the Poisson distribution given by (3.17) and *k* the Binomial distribution by (3.19). We will use normal approximations to both of these random variables as they have high mean values. As a result *m* will have a normal distribution with mean and variance given by,

$$\mu_m = \mu_r - \mu_k, \quad \sigma_m^2 = \sigma_r^2 + \sigma_k^2 \tag{3.22}$$

$$\mu_m = \sum_{i=1}^{K} q_i \rho_i - T_F p_s \tag{3.23}$$

$$\sigma_m^2 = \sum_{i=1}^{K} q_i \rho_i + T_F p_s(1 - p_s) \tag{3.24}$$

Let $f_m(x)$ and $F_m(x)$ denote the pdf and PDF of the random variable m.

Let g denote number of packets that has not been transmitted during a frame, then,

$$g = \begin{cases} 0, & m \le 0 \\ m, & m > 0 \end{cases} \tag{3.25}$$

Then pdf of the random variable *g* is given by,

$$f_g(x) = \begin{cases} F_m(0), & x = 0 \\ f_m(x), & x > 0 \end{cases} \tag{3.26}$$

Next, we determine the following expected value of the random variable *g*,

$$\mu_g = \int_0^{\infty} x f_m(x) dx \tag{3.27}$$

Substituting normal probability density function for $f_m(x)$ in the above,

$$\mu_g = \frac{1}{\sqrt{2\pi}\sigma_m} \int_0^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu_m}{\sigma_m}\right)^2} dx \tag{3.28}$$

$$\mu_g = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} (\mu_m + \sigma_m u) e^{-\frac{1}{2}u^2} du \tag{3.29}$$

where, $u = \frac{x-\mu_m}{\sigma_m}$, $\alpha = -\frac{\mu_m}{\sigma_m}$ (3.30)

$$\mu_g = \mu_m[1 - \Phi(\alpha)] - \frac{1}{\sqrt{2\pi}}\sigma_m \int_b^{\infty} e^v dv \tag{3.31}$$

where, $\Phi(\alpha) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} e^{-\frac{1}{2}u^2} du\ ,\ v = -\frac{1}{2}u^2, b = -\frac{1}{2}\alpha^2$ (3.32)

Performing the integration in the second term,

$$\mu_g = \mu_m[1 - \Phi(\alpha)] + \frac{1}{\sqrt{2\pi}}\sigma_m e^{-\frac{1}{2}\alpha^2} \tag{3.33}$$

Finally, probability that transmission of a packet will be unsuccessful during a frame is given by,

$$P_g = \frac{\mu_g}{\mu_r} \tag{3.34}$$

3.3.1.ii. Packet loss when a machine generates a new packet

As explained above, we assume that a machine will discard a packet not successfully transmitted during a frame only when it generates a new packet. Thus a machine will keep trying to transmit a packet in subsequent frames unless it generates a new packet. Let us define,

$\mathcal{E}_{i,j}$: type-*i* machines packet arrival rate during the *j*'th frame.

$\mathcal{E}_i$: packet arrival rate of type *i* machines at the steady-state.

P_{ℓ_j}: probability that a packet will not be successfully transmitted during *j*'th frame, *j*=1,2,3…

P_ℓ : probability that a packet will not be successfully transmitted during a frame at the steady-state.

ω_i : interarrival time of the packets to a type-i machine in number of frames.

x : random variable denoting the number of the frames that elapsed from the generation of a packet to its successful transmission.

$P_x(n)$: probability that a packet will be transmitted during the n^{th} frame since its generation.

$F_x(n)$: probability that a packet will be transmitted during the first *n* frames since its generation.

$P_{\omega_i}(j)$: probability that packet interarrival time of a type-*i* machine is *j* frames.

Q_i : probability that a packet of a type-i machine will be discarded.

The packet arrival rate of type-*i* machines in the *j*+1'th frame is given by,

$$\mathcal{E}_{i,j+1} = q_i\rho_i + (1-q_i)P_{\ell_j}\mathcal{E}_{i,j}, \quad j = 1,2,3, \dots \tag{3.35}$$

We note that $\mathcal{E}_{i,j+1}$ may be calculated recursively from its initial value of $\varepsilon_{i,1} = q_i\rho_i$. In frame #1, the contending packets consists of only newly generated packets. Thus arrival process of the packets at frame #1 is Poisson with the rate $\varepsilon_{i,1} = q_i\rho_i$. At the end of the frame #1, the contending packets will be split into two substreams randomly as successful and unsuccessful. A substream of the unsuccessful packets together with the newly generated packets will form the packet arrival process to the 2nd frame. Since random splitting of a Poisson process and merging of Poisson processes results in Poisson processes, the packet arrival process to the 2nd frame will also Poisson. With this reasoning, the packet arrival process in each frame will be Poisson. As a result, we can determine probability of unsuccessful packet transmission in each frame by using the results of the previous subsection. Probability that a packet will not be successfully transmitted during *j*'th frame, P_{ℓ_j}, may be determined by setting $\mu_r = \sum_{i=1}^{K}\mathcal{E}_{i,j}$ in (3.18) and calculating corresponding P_g from (3.34) which gives $P_{\ell_j} = P_g$.

Taking limit in the (3.35) as $j \to \infty$, the packet arrival rate of type-*i* machines at the steady-state is given by,

$$\mathcal{E}_i = \frac{q_i\rho_i}{1-P_\ell(1-q_i)} \tag{3.36}$$

where $\mathcal{E}_i = \lim_{j\to\infty}\mathcal{E}_{i,j}, \quad P_\ell = \lim_{j\to\infty}P_{\ell_j}$.

$$P_{\omega_i}(j) = q_i(1-q_i)^{j-1}, \qquad j = 1,2, \dots \tag{3.37}$$

$$P_x(n) = (1-P_\ell)P_\ell^{n-1}, n = 1,2, \dots \tag{3.38}$$

$$F_x(n) = \sum_{n=1}^{n} P_x(j) = (1-P_\ell)\sum_{j=1}^{n} P_\ell^{j-1} \tag{3.39}$$

which may be simplified as,

$$F_x(n) = 1 - P_\ell^n \tag{3.40}$$

Next we determine probability that a packet of a type$-i$ machine will be discarded, which is given by,

$$Q_i = \sum_{n=1}^{\infty} P(x>n)P_{\omega_i}(n) = \sum_{n=1}^{\infty}(1-F_x(n))P_{\omega_i}(n)$$

$$Q_i = \sum_{n=1}^{\infty} P_\ell^n q_i(1-q_i)^{n-1} \tag{3.41}$$

The above simplifies to,

$$Q_i = \frac{q_i P_\ell}{1 - P_\ell(1 - q_i)} \tag{3.42}$$

3.3.2.Contention Service without packet loss

In this scenario we assume that the packets will remain in the system until they are transmitted. It will be assumed that the packets generated by all the machines will join to a global queue which will provide random order of service. In this subsection, we will derive the probability generation function (PGF) of the distribution of the number of packets in the system. Then we will determine mean packet delay. Let us introduce the following notation,

b: number of slots in a frame.

r : number of successful packet transmissions during a frame

p : probability of successful packet transmission during a slot.

n_i: number of the packets in the queue at the end of i'th frame.

a_i: total number of packets generated during the i'th frame.

$$r_j = Prob(j\ packets\ will\ be\ served\ during\ a\ frame)$$

$$r_j = \binom{b}{j} p^j q^{b-j} \qquad j = 0,1,\dots,b \tag{3.43}$$

$$r(z) = E[z^r] = \sum_{j=0}^{b} z^j \Pr(r = j) = (pz + 1 - p)^b \tag{3.44}$$

The number of packets in the system follows a Markov chain at the end of frames, which form imbedded points. The number of packets at two consecutive imbedded points follows the equation below,

$$n_{i+1} = n_i - \min(n_i, r) + a_{i+1} \tag{3.45}$$

Let $P_i(z)$ denote the PGF of n_i,

$$P_i(z) = E[z^{n_i}] = \sum_{k=0}^{\infty} z^k \Pr(n_i = k) \tag{3.46}$$

Then $P_{i+1}(z)$ is given by,

$$P_{i+1}(z) = E\left[z^{n_i - \min(n_i, r) + a_{i+1}}\right] = E\left[z^{n_i - \min(n_i, r)}\right] A(z) \tag{3.47}$$

Where $A(z)$ is PGF of the number of packets generated during a frame. $A(z) = E[z^{\alpha_i}]$.

Next, we will determine $E\left[z^{n_i - \min(n_i, r)}\right]$,from the definition of the expectation,

$$E\left[z^{n_i - \min(n_i, r)}\right] = \sum_{k=0}^{\infty} \sum_{j=0}^{b} z^{k - \min(k, j)} \Pr(n_i = k, r = j) \tag{3.48}$$

We note that since number of new packet arrivals are independent of the queue length, $\Pr(n_i = k, r = j) = \Pr(n_i = k) \Pr(r = j)$

Next, we determine the above expectation by conditioning on the value of r.

$$E\left[z^{n_i - \min(n_i, r)} | r = j, j > 0\right] = \sum_{k=0}^{\infty} z^{k - \min(k, j)} \Pr(n_i = k)$$

$$= \sum_{k=0}^{j-1} z^0 \Pr(n_i = k) + \sum_{k=j}^{\infty} z^{k-j} \Pr(n_i = k)$$

$$= \sum_{k=0}^{j-1} \Pr(n_i = k) + z^{-j}\left[P_i(z) - \sum_{k=0}^{j-1} z^k \Pr(n_i = k)\right] \tag{3.49}$$

$$E\left[z^{n_i - \min(n_i, r)} | r = j, j = 0\right] = \sum_{k=0}^{\infty} z^{k - \min(k, j)} \Pr(n_i = k) = \sum_{k=0}^{\infty} z^k \Pr(n_i = k) = P_i(z)$$

combining the above conditional expectations,

$$E\left[z^{n_i - \min(n_i, r)}\right] = \sum_{j=0}^{b} E\left[z^{n_i - \min(n_i, r)} | r = j\right] \Pr(r = j)$$

$$= \sum_{j=1}^{b} \sum_{k=0}^{j-1} \Pr(n_i = k) \Pr(r = j) + \sum_{j=1}^{b} z^{-j} \left[P_i(z) - \sum_{k=0}^{j-1} z^k \Pr(n_i = k)\right] \Pr(r = j)$$
$$+ P_i(z) \Pr(r = 0)$$

$$= \sum_{j=1}^{b}\sum_{k=0}^{j-1} \Pr(n_i = k)\Pr(r = j)$$

$$+ \sum_{j=0}^{b} z^{-j}\Pr(r = j)\, P_i(z) - \sum_{j=1}^{b}\sum_{k=0}^{j-1} z^{k-j}\Pr(n_i = k)\Pr(r = j)$$

$$E\left[z^{n_i - \min(n_i, r)}\right]$$
$$= \sum_{j=1}^{b}\sum_{k=0}^{j-1} \Pr(n_i = k)\Pr(r = j) + (pz^{-1} + 1 - p)^b P_i(z)$$
$$- \sum_{j=1}^{b}\sum_{k=0}^{j-1} z^{k-j}\Pr(n_i = k)\Pr(r = j) \tag{3.50}$$

Substituting the above expectation in (3.47) gives,

$$P_{i+1}(z) = \left[\sum_{j=1}^{b}\sum_{k=0}^{j-1} \Pr(n_i = k)\Pr(r = j) + (pz^{-1} + 1 - p)^b P_i(z) - \sum_{j=1}^{b}\sum_{k=0}^{j-1} z^{k-j}\Pr(n_i = k)\Pr(r = j)\right] A(z) \tag{3.51}$$

We assume a steady-state solution where $lim_{i\to\infty} P_i(z) = P(z)$. Taking the limit of the above equation as, $i \to \infty$ gives,

$$P(z) = \frac{\left[\sum_{j=1}^{b}\sum_{k=0}^{j-1} \Pr(n_i = k)\Pr(r = j) - \sum_{j=1}^{b}\sum_{k=0}^{j-1} z^{k-j}\Pr(n_i = k)\Pr(r = j)\right] A(z)}{1 - (pz^{-1} + 1 - p)^b A(z)} \tag{3.52}$$

Let $p_k = \Pr(n_i = k)$ and $r_j = \Pr(r = j)$, we have,

$$P(z) = \frac{\left[\sum_{j=1}^{b}\sum_{k=0}^{j-1} p_k\, r_j (z^b - z^{b+k-j})\right] A(z)}{z^b - (p + (1-p)z)^b A(z)}$$

Finally, we have PGF of the number of packets in the system at the steady-state as,

$$P(z) = \frac{\left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j (z^b - z^{b+k-j})\right] A(z)}{z^b - (p+(1-p)z)^b A(z)} \tag{3.53}$$

we note that $A(z)$ is given by $r(z)$ in equation (3.17).

The above PGF has b unknown probabilities $p_k, k = 0, \ldots, b-1$. These unknowns may be determined through the application of Rouche's theorem using the standard procedure [33]. Next, we determine the mean number of packets in the system, $\bar{N}$, which is given by,

$$\bar{N} = \left. \frac{dP(z)}{dz} \right|_{z=1}$$

Let us rewrite (3.53) as follows,

$$[z^b - (p + (1-p)z)^b A(z)] P(z) = \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(z^b - z^{b+k-j}\right)\right] A(z) \tag{3.54}$$

Next differentiating the both side of the above wrt z gives,

$$\begin{aligned} bz^{b-1} - b(p + (1-p)z)^{b-1}(1-p)A(z) - \left[p + (1-p)z^b A'^{(z)}\right] P(z) \\ + [z^b - (p + (1-p)z)^b A(z)] P'^{(z)} \\ = \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(bz^{b-1} - (b+k-j)z^{b+k-j-1}\right)\right] A(z) \\ + \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(z^b - z^{b+k-j}\right)\right] A'(z) \end{aligned} \tag{3.55}$$

The second differentiation of the above results in,

$$\begin{aligned} [b(b-1)z^{b-2} - b(b-1)(p + (1-p)z)^{b-2}(1-p)A(z) \\ - 2[b[p + (1-p)z]^{b-1}(1-p)A'(z)] - [p + (1-p)z]^b A''(z)] P(z) \\ + 2[bz^{b-1} - b(p + (1-p)z)^{b-1}(1-p)A(z) - (p + (1-p)z)^b A'(z)] P'(z) \end{aligned}$$

$$= \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(b(b-1)z^{b-2} - (b+k-j)(b+k-j-1)z^{b+k-j-2}\right)\right] A(z)$$

$$+ 2\left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(bz^{b-1} - (b+k-j)z^{b+k-j-1}\right)\right] A'(z) \qquad (3.56)$$

Substituting $z = 1$ in the above results in,

$$[b(b-1) - b(b-1)(1-p)^2 - 2[b(b-1)A'(1)] - A''(1)] + 2[b - b(1-p) - A'(1)]P'(1)$$

$$= \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(b(b-1) - (b+k-j)(b+k-j-1)\right)\right]$$

$$+ 2\left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j \left(b - (b+k-j)\right)\right] A'^{(1)} \qquad (3.57)$$

Solving the above for $\overline{N} = P'(1)$ gives,

$$\overline{N} = \frac{\begin{bmatrix} \left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j\left(b(b-1)-(b+k-j)(b+k-j-1)\right)\right] \\ +2\left[\sum_{k=0}^{b-1} p_k \sum_{j=k+1}^{b} r_j\left(b-(b+k-j)\right)\right]A'^{(1)} \\ -\left[b(b-1)-b(b-1)(1-p)^2-2\left[b(b-1)A'(1)\right]-A''(1)\right] \end{bmatrix}}{2[b-b(1-p)-A'(1)]} \qquad (3.58)$$

Next we use the Little's formula to obtain the average packet delay in number of slots as follows,

$$\overline{D} = \frac{\overline{N}}{\lambda} \qquad (3.59)$$

3.4. Numerical Results

Next, we present some numerical results for the analysis presented in this chapter. In Fig 3.1 we present probability distribution of the number of busy slots in reservation subframe with $C = 100$ slots. From this figure, probability of reservation request equals to $u(C) = 0.005$.

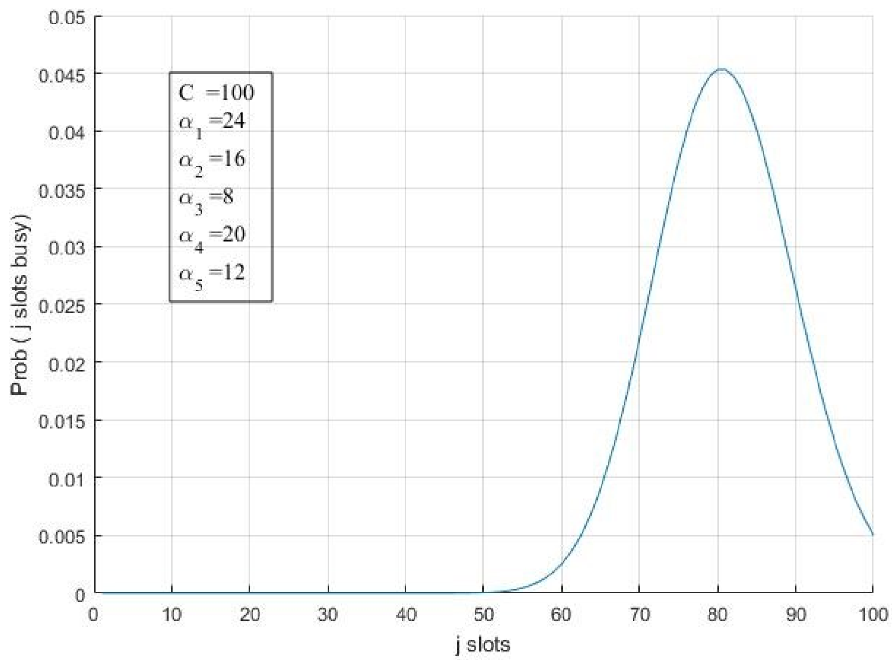


Fig 3.1. Probability distribution of the number of busy slots in reservation subframe

In Fig. 3.2, we present average number of machines in the system as a function of the total machine arrival rate for a given values of γ_i, $\beta_{i,}$ q_i. It may be observed that average number of machines in the system will increase linearly with the total arrival rate.

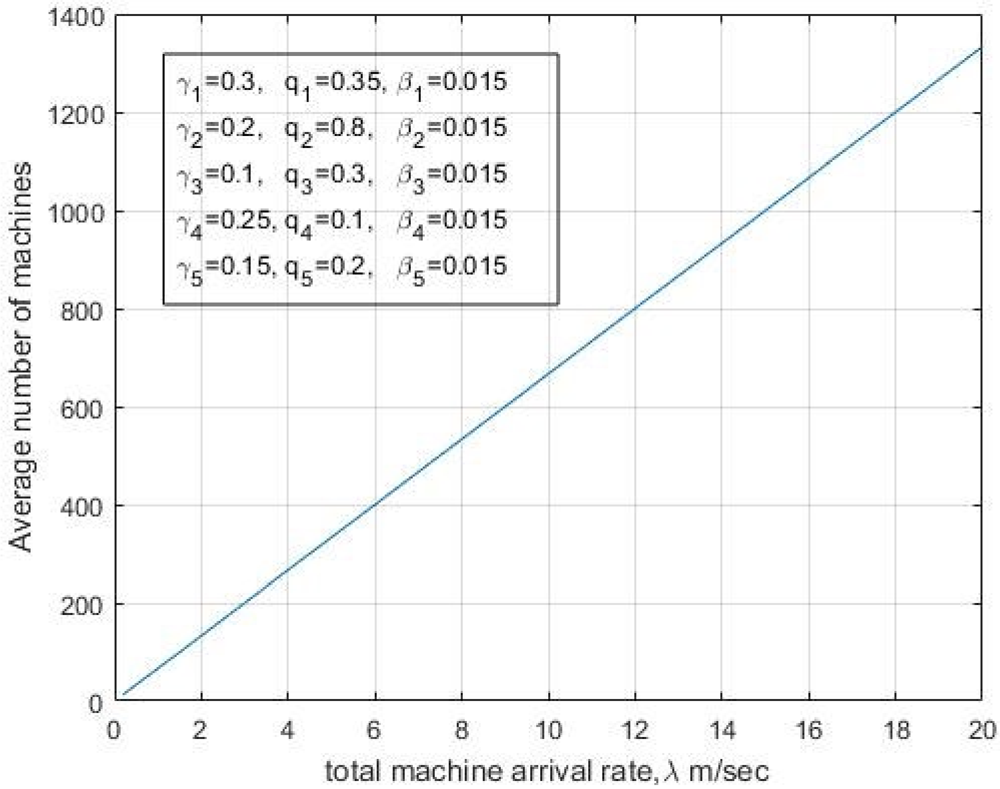


Fig. 3.2. Average number of machines in the system as a function of the total machine arrival rate for the given type parameter values, $\gamma_i,\ q_i,\ \beta_i, i = 1..5$

In Fig. 3.3, we present packet loss probability as a function of the frame number with total arrival rate as a parameter for the system in Fig.3.2. It may be seen that packet loss probability quickly reaches to its steady-state value.

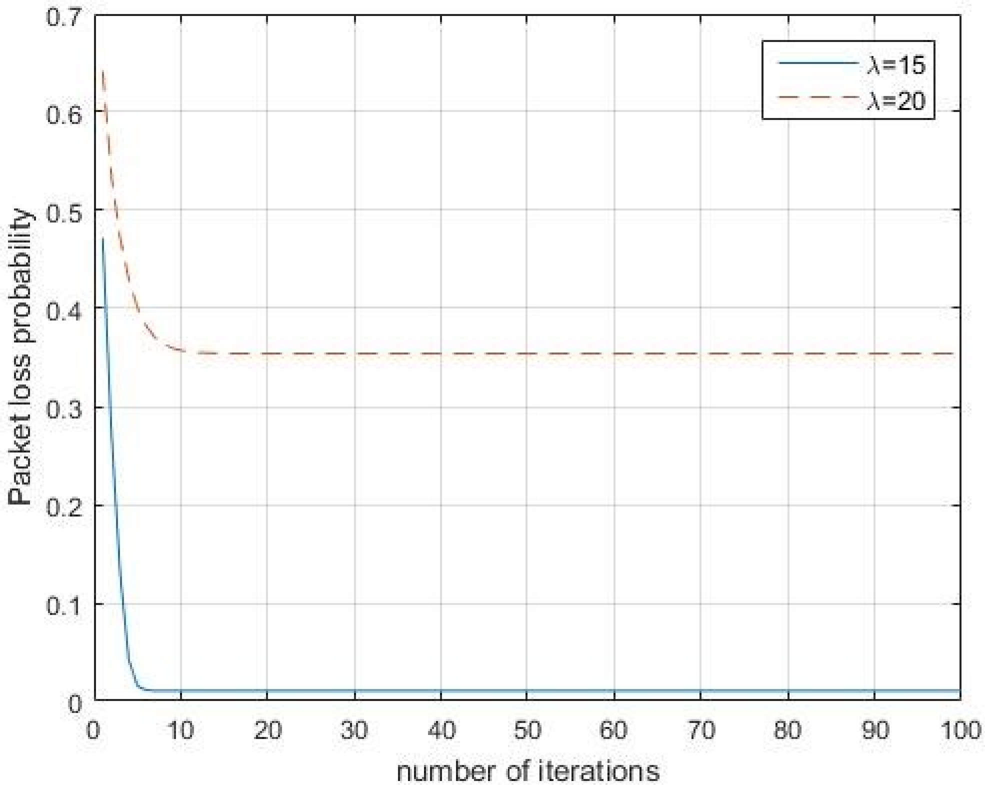


Fig. 3.3. Packet loss probability as a function of the frame number with total machine arrival rate as a parameter.

Fig. 3.4 presents initial and steady-state packet loss probabilities P_{l_0} and P_{l_∞} respectively as a function of the total machine arrival rate for the parameter values of Fig. 3.2. Lower packet loss probability may be observed in the system that allows a packet transmission attempts until a machine generates a new packet.

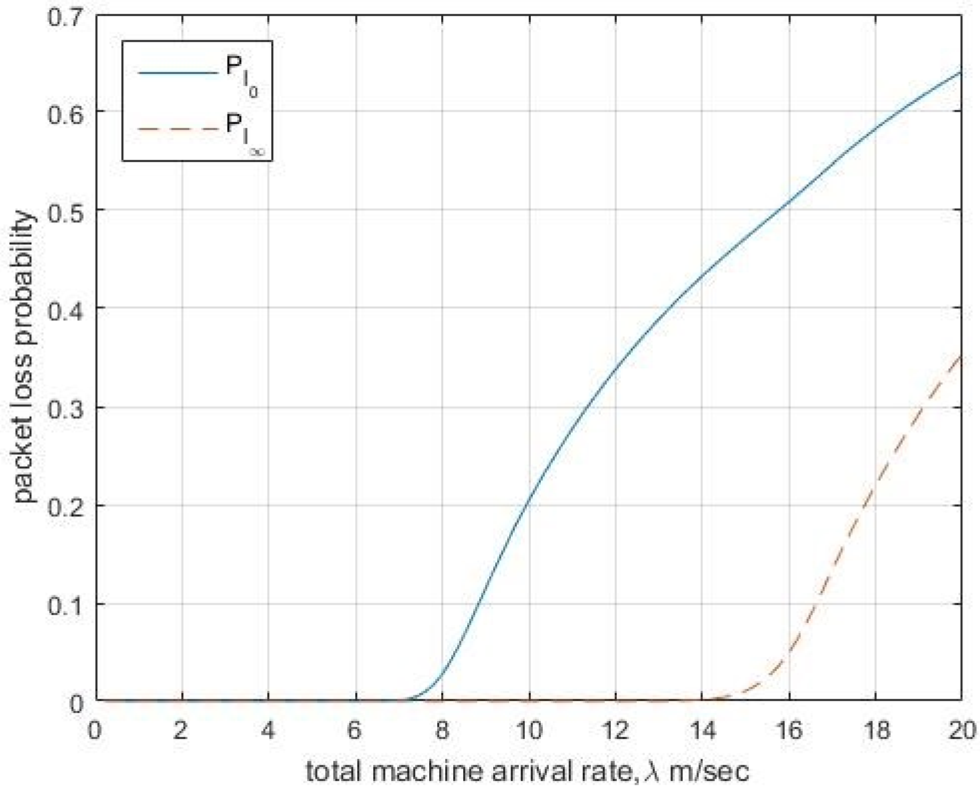


Fig. 3.4. Initial and steady-state packet loss probabilities as a function of the total machine arrival rate.

The Fig. 3.5, presents packet loss probability of each machine type as a function of machine arrival rate, we may observe that packet loss probabilities have nearly zero values for the wide range of machine arrival rate. Also, it may be seen that higher packet generation probabilities in machine types result in higher packet loss probabilities.

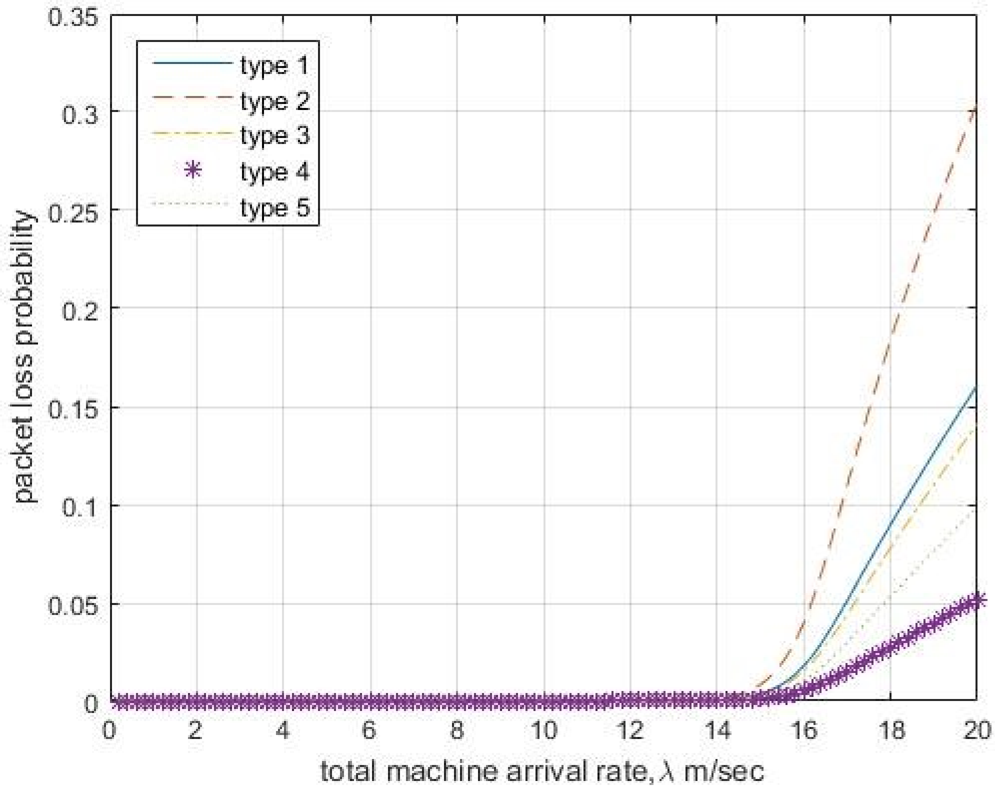


Fig. 3.5. Per machine type packet loss probability as a function of total machine arrival rate.

In Fig. 3.6 the mean packet delay as a function of total packet arrival rate is presented for contention serve without packet discarding. It may be seen that with no packet loss discipline the mean packet delay increases exponentially with the total packet arrival rate.

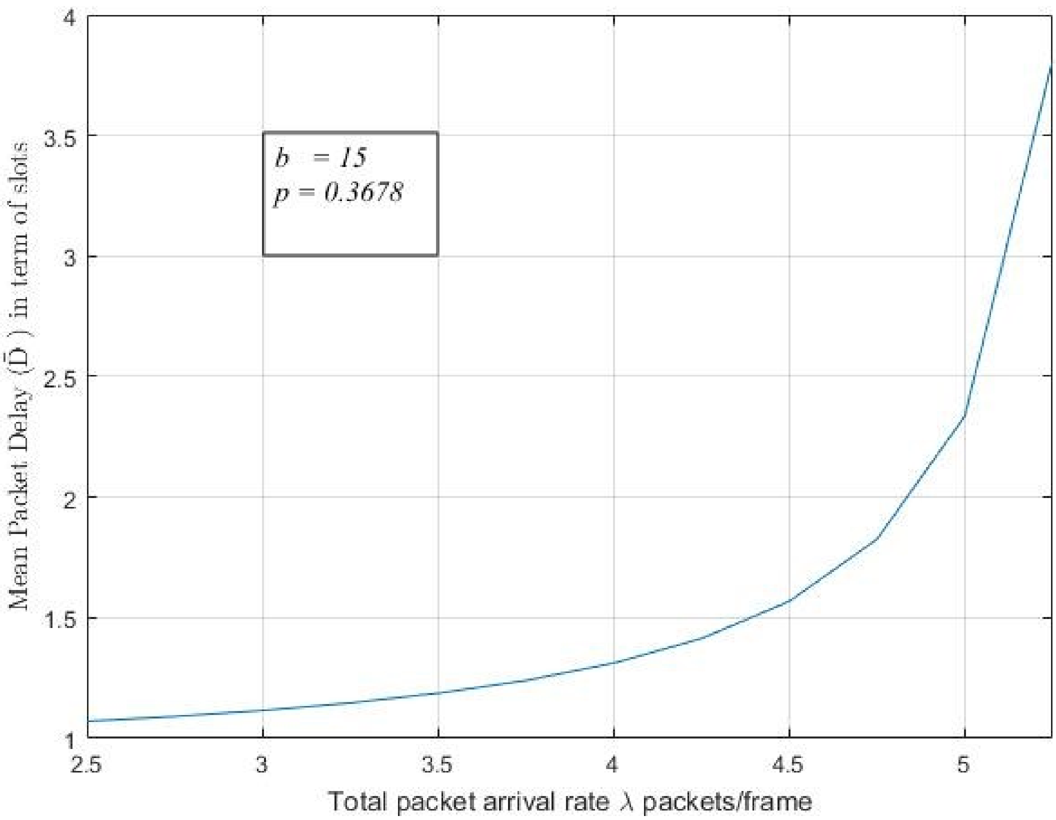


Fig. 3.6. Mean packet delay as a function of total packet arrival rate for a given frame length in number of slots and packet success probability in a slot

3.5. Conclusion

In this chapter, we considered the performance of our MAC protocol for M2M communication in the presence of mobile users. We assume that the users (machines) arrive to the system according to a Poisson process and spend random amount of time in the network. Machines have been classified into K types where each machine may generate a single packet per frame during the time that it's in the system according to a different probability. We considered reservation service and contention based service with and without packet loss. For reservation service, we have determined blocking probability of a reservation request. For contention service, we have determined mean packet delay and packet loss performance measures. We showed numerical results for blocking probability, packet loss probability and mean packet delay.

Chapter 4

Modelling of a MAC Protocol with Probabilistic Service for M2M Communications

4.1. Introduction

In this chapter, we consider a MAC protocol that serves the packets according to their priorities. We assume that traffic is divided into classes with each assigned a priority. All the packets are served by a single server with an infinite queue according to their relative probabilities.

4.2. System Model

We assume that there are J classes of traffic. The packets arrive at the network according to a Poisson process with the rate λ packets/slot. An arriving packet belongs to class j with probability q_j. Letting λ_j denote arrival rate of class j packets. Then,

$$\lambda_j = \lambda q_j \qquad j = 1,2,\ldots,J \tag{4.1}$$

$$\sum_{j=1}^{J} q_j = 1$$

The notation of the analysis is given in Table 4.1.

Symbol	Definition
n_i	number of class i packets in the system.
q_j	probability that a packet generated belongs to class j
B_j	transmission time of a class j packet
ρ_j	traffic intensity of class j traffic
ρ	total traffic intensity
P_s	Probability of successful packet transmission
p_j	weight of selecting a class j packet for service

Table 4.1. Definition of the symbols

The arriving packets join to an infinite queue which is served by a single server. We assume that the service is given according to discriminatory random order service (DROS) discipline defined in [31]. In this service discipline, the server chooses a class j packet to be served with probability given by,

$$\frac{p_j}{\sum_{i=1}^{J} n_i p_i} \qquad j = 1,2,\ldots,J \tag{4.2}$$

Where n_i and p_i are defined in Table 4.1. The packets of class with higher weight parameter will be chosen with higher probability.

4.3. Mean Packet Waiting Time

Next, we present packet waiting time for DROS discipline from the results in [31]. From [31], mean waiting time of a packet in slots is given by,

$$E[W_j] = \sum_{k=1}^{J} \frac{\lambda_j E[{B_j}^2]}{2} \sum_{k=1}^{J} (X^{-1})_{jk} \qquad 1 \leq j \leq J \tag{4.3}$$

Where $E[{B_j}^2]$ is the second moment of the transmission time of a packet and X^{-1} is the inverse of X matrix with $J \times J$ components defined as below.

$$X_{jk} = \begin{cases} -\frac{\rho_k p_k}{p_j + p_k} & if\ j \neq k \\ 1 - \sum_{\ell=1}^{J} \frac{\rho_\ell p_\ell}{p_j + p_\ell} - \frac{\rho_j}{2} & if\ j = k \end{cases} \tag{4.4}$$

Where $\rho_j = \lambda_j E[B_j]$ is traffic intensity for class j traffic.

Total traffic intensity may be defined as (4.5) and for the stability $\rho < 1$.

$$\rho = \sum_{j=1}^{J} \rho_j \tag{4.5}$$

We will assume that packets have constant transmission times, which equals to a single slot.

In the following, we will consider two service scenarios, which will be referred to as constant and geometric service time scenarios. In constant service scenario packet of each class will be served according to first-come-first-served (FCFS) discipline, in the geometric service scenario packets of each class will be served randomly. Next, we present mean packet waiting time for each scenario.

4.3.1. Mean Packet Waiting Time with Constant Service Time

In this case we assume that service time of a packet equals to its transmission time. Since packet transmission times are equal to a single slot, $E[B] = 1$. So, traffic intensity for class j machine will be $\rho_j = \lambda_j E[B] = \lambda_j$ and second moment of packet transmission time will be $E[B^2] = 1$.

4.3.2. Mean Packet Waiting Time with Geometric Service Time

In this case, we assume that server chooses to serve a packet of a class j, then all packets in that class contend for service according to slotted Aloha protocol. In this case, service time of a packet will be geometric, and service time in number of slots given by,

$$P_K = \Pr(\textit{service time of a packet is k slots})$$

$$P_K = P_s(1 - P_s)^{k-1} \qquad k = 1,2,3,\ldots \tag{4.6}$$

Then PGF of the service time of a packet is given by,

$$B(z) = E[z^B]$$

$$B(z) = \frac{P_s z}{1-(1-P_s)z} \tag{4.7}$$

$$E[B] = \frac{dB(z)}{dz}\bigg|_{z=1} = \frac{1}{P_s}$$

$$\mathrm{E}[\mathrm{B}^2] = \frac{d^2B(z)}{dz^2}\bigg|_{z=1} = \frac{1+(1-\mathrm{P_s})}{\mathrm{P_s}^2} \tag{4.8}$$

4.4. Numerical results

In this section, we present some numerical result for the two scenarios considered in this chapter. We assumed that we have 3 classes of traffic. Table 4.2. shows the traffic intensity of each class of traffic.

q_1	q_2	q_3
0.54	0.29	0.17

Table 4.2. Probability of an arriving packet belonging to class *j* traffic

It can be observed that $\sum_{j=1}^{J} q_j = 1$. By increasing λ from lower to the upper limit, that $\rho < 1$ for stability, we will obtain the mean packet waiting time for each scenario.

We assume the weight parameters given in Table 4.3 for each class of traffic.

p_1	p_2	p_3
5	2	1

Table 4.3. Weight of selecting a class *j* packet for service

Fig.s 4.2 and 4.3 plot mean packet waiting time as a function of λ for parameters value given in Tables 4.2 and 4.3 for constant and geomatric service times respectively.

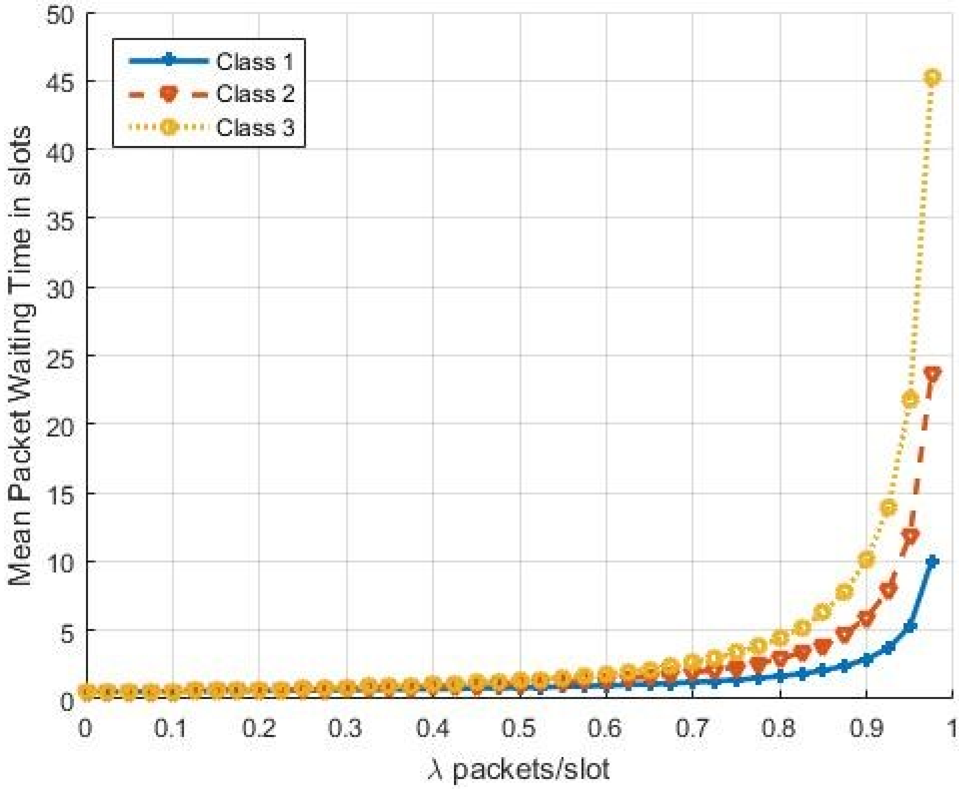


Fig. 4.1. Mean packet waiting time for constant service time as a function of λ

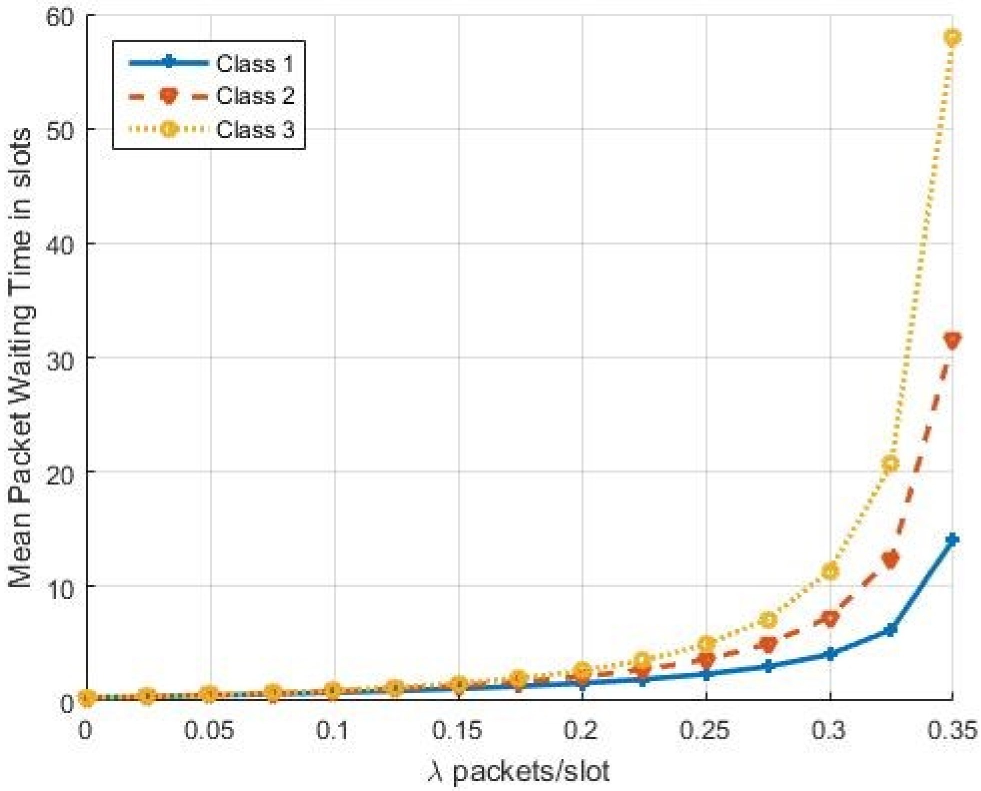


Fig. 4.2. Mean packet waiting time for geometric service time as a function of λ

As may be seen mean waiting time of class 1 traffic is lower than classes 2 and 3 because of the higher selection probability of class 1 traffic for service.

Fig. 4.3 plots mean packet waiting time for constant and geometric service times for class 1 traffic as a function of λ for parameter values given in Table 4.1 and 4.2. As expected, constant service time results in lower mean waiting time than geometric service time. Also, the system sarurated faster under geometric than sonstant service time.

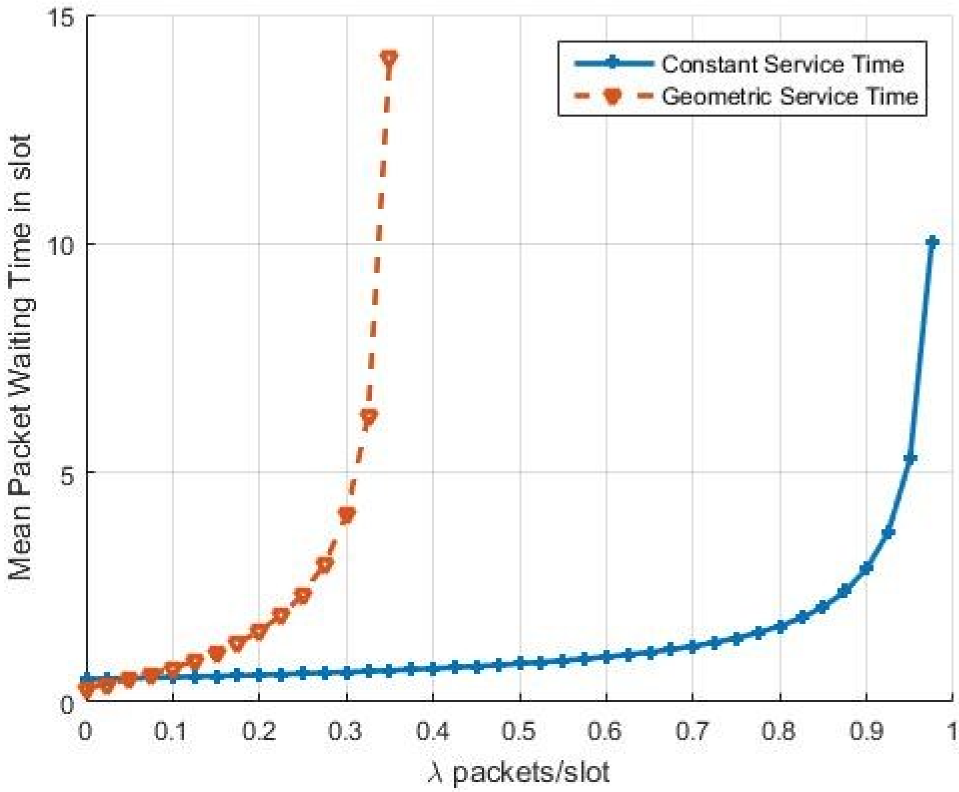


Fig. 4.3. Mean packet waiting time as a function of λ

4.5. Conclusion

In this chapter we modeled a MAC protocol that serves the packets according to their priorities. We considered constant and geometrically distributed service time and obtained the results for each scenario separately. In addition, from the comparison between two scenarios we showed the deference of mean packet waiting time as a function of λ. From the numerical results it can be seen that the mean packet delay with geomatric service time will be higher than the constant service time.

Chapter 5

Conclusion and Future Work

M2M communications is to play an important role in realization of IoT. However so far M2M communications lacks simple, efficient and robust MAC protocol. In this thesis, we propose a scalable hybrid MAC protocol for M2M communications. This protocol will be able to serve multiple classes both bursty and periodic traffic by heterogeneous users. Each traffic class will be served in a dedicated period during the frame, further bursty traffic is served by the contention channels and periodic traffic by the reserved channels. We developed an LP optimization problem that determines the frame structure of this protocol that meets QoS requirements of the users. Numerical results show that the system will be able to serve large machine populations.

Then the analysis of the proposed protocol has been extended to capture user mobility. It's assumed that the users arrive to the system according to a Poisson process and spend random amount of time in the system. Under this arrival process we considered both reservation and contention service. Contention service has been studied with and without packet discarding. We determined blocking probability of reservation requests, packet loss probability and mean packet delay for contention service schemes. We have also studied a MAC protocol that serves the packets according to their relative probabilities.

The proposed protocol is simple, meets the QoS requirements of the users and adoptable to changing traffic conditions in the system. The results of this thesis will be helpful in the design of M2M communications.

As future work we propose optimization of the system that will include energy consumption of the machine communications and packet delay constraints of the traffic.

References

[1] A. Aijaz and A. H. Aghvami. “A PRMA based MAC protocol for cognitive machine-to-machine communications”. Presented at *Communications (ICC), 2013 IEEE International Conference* on. 2013 DOI: 10.1109/ICC.2013.6654955.

[2] F. V. Azquez-Gallego, J. Alonso-Zarate, I. Balboteo and L. Alonso. “DPCF-M: A medium access control protocol for dense machine-to-machine area networks with dynamic gateways”. Presented at *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop* on. 2013, . DOI: 10.1109/SPAWC.2013.6612098.

[3] Chen-Yu Hsu, Chi-Hsien Yen and Chun-Ting Chou. “An adaptive multichannel protocol for large-scale machine-to-machine (M2M) networks”. Presented at *Wireless Communications and Mobile Computing Conference (IWCMC)*, 2013 9th International. 2013, . DOI: 10.1109/IWCMC.2013.6583731.

[4] Chul Wan Park, Duckdong Hwang and Tae-Jin Lee. “Enhancement of IEEE 802.11ah MAC for M2M communications”. *Communications Letters, IEEE* 18(7), pp. 1151-1154. 2014. . DOI: 10.1109/LCOMM.2014.2323311.

[5] Guo Wang, Xiaofeng Zhong, Shunliang Mei and Jing Wang. “An adaptive medium access control mechanism for cellular based machine to machine (M2M) communication”. Presented at *Wireless Information Technology and Systems (ICWITS)*, 2010 *IEEE International Conference* on. 2010, . DOI: 10.1109/ICWITS.2010.5611820.

[6] Huasen Wu, Chenxi Zhu, R. J. La, Xin Liu and Youguang Zhang. “Fast adaptive S-ALOHA scheme for event-driven machine-to-machine communications”. Presented at *Vehicular Technology Conference (VTC Fall), IEEE*, 2012. DOI: 10.1109/VTCFall.2012.6398917.

[7] Y. Liu, C. Yuen, X. Cao, N. U. Hassan, and J. Chen, “Design of a scalable hybrid MAC protocol for heterogeneous M2M networks,” *IEEE Internet Things* J., vol.1, no.1, pp. 99–111, 2014.

[8] A. Lo, Y. W. Law, M. Jacobsson and M. Kucharzak. “Enhanced LTE-advanced random-access mechanism for massive machine-to-machine (M2M) communications”. Presented at *27th World Wireless Research Forum (WWRF) Meeting*, 2011.

[9] N. K. Pratas, H. Thomsen, C. Stefanovic and P. Popovski. "Code-expanded random access for machine-type communications". Presented at *Globecom Workshops (GC Wkshps), 2012 IEEE*. 2012, . DOI: 10.1109/GLOCOMW.2012.6477838.

[10] Yi Liu, Chau Yuen, Jiming Chen and Xianghui Cao. "A scalable hybrid MAC protocol for massive M2M networks". Presented at *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. 2013, . DOI: 10.1109/WCNC.2013.6554572.

[11] Yu Chen and Wei Wang. "Machine-to-machine communication in LTE-A". Presented *at Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010. DOI: 10.1109/VETECF.2010.5594218.

[12] C.L. Fullmer and J.J. Garcia-Luna-Aceves, "Floor acquisition multiple access (FAMA) for packet-radio networks", in: Proc. ACM SIGCOMM 95*, Cambridge, MA* (August 28–September 1, 1995).

[13] G. Wang, X. Zhong, S. Mei, and J. Wang, "An adaptive medium access control mechanism for cellular based Machine to Machine (M2M) communication," in Proc. of *IEEE ICWITS* 2010, 2010.

[14] G. Khandelwal, K. Lee, A. Yener, and S. Serbetli, "ASAP: A MAC protocol for dense and time-constrained RFID systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, no. 2, pp. 18 730-1–18 730-13, Jan. 2007.

[15] D. Tarchi, R. Fantacci, and D. Marabissi, "Proposal of a cognitive based MAC protocol for M2M environments," in *Proc. 24th IEEE Int. Symp. Personal Indoor Mobile Radio Commun. (PIMRC)*, 2013, pp. 1609– 1613.

[16] I. Park, D. Kim, and D. Har. "MAC achieving low latency and energy efficiency in hierarchical M2M networks with clustered nodes." *IEEE Sensors Journal 15.3,* 2015, 1657-1661.

[17] A. Azari and G. Miao. "Energy efficient MAC for cellular-based M2M communications". Presented at *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference* on. 2014, . DOI: 10.1109/GlobalSIP.2014.7032092.

[18] P. K. Verma, R. Tripathi and K. Naik. "A robust hybrid-MAC protocol for M2M communications". Presented at *Computer and Communication Technology (ICCCT), 2014 International Conference* on. 2014, . DOI: 10.1109/ICCCT.2014.7001503.

[19] Boswarthick D, Elloumi O, Hersent O, editors. M2M communications: a systems approach. *John Wiley & Sons*; 2012 Mar 14.

[20] H. Shariatmadari, et al., "Machine-Type Communications: Current Status and Future Perspectives Toward 5G Systems", *IEEE Communications Magazine, Comm. Standards Supplement*, pp. 10-17, Sept. 2015.

[21] C. Kahn, H. Viswanathan, "Connectionless Access for Mobile Cellular Networks", *IEEE Communications Magazine, Comm. Standards Supplement*, pp. 26-31, Sept. 2015.

[22] S. Andrew et al., "Efficient Small Data Access for Machine-Type Communications in LTE", *Proceedings of ICC*, pp. 3569-74, June 2013.

[23] J.J Nielson et al., " What Can Wireless Cellular Technologies Do About the Smart Metering Traffic?", *IEEE Communications Magazine, Comm. Standards Supplement*, pp. 41-47, Sept. 2015.

[24] C. Pereira and A. Aguiar,"Towards Efficient Mobile M2M Communications: Survey and Open Challenges", *Sensors*, 14, 19582-19608; doi:10.3390/s141019582, 2014.

[25] A. Rajandekar and B. Sikdar,"A Survey of MAC Layer Issues and Protocols for Machine-to-Machine Communications", *IEEE Internet of Things Journal for Machine-to-Machine Communications*, Vol. 2, No. 2, pp.175-186, April 2015.

[26] C. Karuupongsiri, K. Munasinghe, and A. Jamalipour, " Smart Meter Packet Transmission via the Control Signal of LTE Networks", *IEEE ICC-Mobile and Wireless Networking Symposium*, pp. 2291-2996, 2015.

[27] G. M. Madueno et al.,"Massive M2M Access with Reliability Guarantees in LTE Systems", *IEEE ICC-Mobile and Wireless Networking Symposium*, pp. 2297-3002, 2015.

[28] A. Gotsis, A. Lioumpas and A. Alexiou, "M2M Scheduling Over LTE", *IEEE Vehicular Technology Magazine*, pp. 34—39, Sept. 2012.

[29] A. Leon-Garcia and I. Widjaja, "Communication Networks",*McGraw Hill*, 2nd ed., 2004.

[30] Boswarthick, David, Omar Elloumi, and Olivier Hersent, eds. M2M communications: a systems approach. *John Wiley & Sons*, 2012.

[31] J. Kim, J. Kim, & B. Kim, "Analysis of the M/G/1 queue with discriminatory random order service policy" *Performance Evaluation*, 68(3), 256-270, Mar. 2011.

[32] J. Kaufman, ''Blocking in a shared resource environment'' *IEEE Transactions on communications*, 29(10), 1474-1481, 1981.

[33] J.F. Hayes, & T.V.G. Babu, Modeling and analysis of telecommunications networks. *John Wiley & Sons*, 2004.