

# **Evaluating Website Data Leaks through Spam Collection on Honeypots**

**Oghenerukevwe Elohor Oyinloye**

**A Thesis**

**in**

**The Department**

**of**

**Concordia Institute for Information Systems Engineering**

**Presented in Partial Fulfillment of the Requirements**

**for the Degree of**

**Master of Applied Science (Information Systems Security) at**

**Concordia University**

**Montréal, Québec, Canada**

**March 2025**

**© Oghenerukevwe Elohor Oyinloye, 2025**

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Oghenerukevwe Elohor Oyinloye**

Entitled: **Evaluating Website Data Leaks through Spam Collection on Honeypots**

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

**Master of Applied Science (Information Systems Security)**

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

Signed by the Final Examining Committee:

\_\_\_\_\_ Chair  
*Dr. Lingyu Wang*

\_\_\_\_\_ Examiner  
*Dr. Amr Youssef*

\_\_\_\_\_ Supervisor  
*Dr. Carol Fung*

Approved by

\_\_\_\_\_  
Dr. Chun Wang, Chair  
Department of Concordia Institute for Information Systems Engineering

\_\_\_\_\_ 2025

\_\_\_\_\_  
Dr. Mourad Debbabi, Dean  
Faculty of Engineering and Computer Science

# Abstract

## Evaluating Website Data Leaks through Spam Collection on Honeypots

Oghenerukevwe Elohor Oyinloye

People increasingly rely on online services for communication, education, shopping, entertainment etc., but this convenience comes with escalating spam volumes. Prior studies have linked spam primarily to user behavior or data leaks but lacked effective methods to distinguish these causes. Our research proposes using spam as a forensic indicator of data leaks and introduces a low-complexity fingerprinting technique to trace leak sources, while evaluating consent and subscription practices of websites and telecom.

We deployed 148 honeypots with 740 accounts across 370 websites in 12 communities over 12 months, analyzing 12,490 spam emails to assess forensic indicators and evaluate exposure, privacy policy exploitation seen with legitimate websites under the Canadian Anti-Spam Law (CASL). Our method was bench-marked against traditional leak detection models and fingerprinting techniques.

Our findings reveal that many legitimate websites exploit CASL consent practices by automatically enrolling users in mailing lists via implicit consent, while sites requiring explicit consent often violate their own policies, highlighting enforcement gaps. We recommend that regulators mandate a clear separation between subscription agreements and privacy policies and require explicit third-party consent at sign-up.

Additionally, our analysis shows that men aged 48–57 receive the highest spam volumes, with peak activity between 00-04 minute of each hour, with peaks on Thursdays. These insights offer valuable guidance for enhancing spam filtering models.

Using our analysis engine, we achieved 100% leak detection and 99.29% source attribution accuracy. Compared to network intrusion detection, log analysis, machine learning, and traditional fingerprinting, our method more effectively identifies compliance violations, traces leaks to their source, and estimates exposure impact. Furthermore, to address telecom issues from phone number reassignments, we proposed a nonce-based de-association method that promises significant spam reduction.

# Acknowledgments

I give all thanks and glory to the Almighty Prantokantor for making this dream a reality. I would like to express my deepest gratitude to my supervisor, Prof. Carol Fung. She is truly one of a kind—always available, even at odd hours, and incredibly supportive and helpful throughout this journey. I sincerely thank Prof. Aiman Hanna and Prof. Morales Rodrigo for their support and for giving me the opportunity to serve as a teaching assistant under their guidance.

My sincere appreciation goes to my colleagues at NGNSec Laboratory—Fareed, Waleed, Rasool, Rambod, Himel, and my dear Vaishnavi. Your support and encouragement have been invaluable. A heartfelt thank you to my dear friend Azeez and his family for their unwavering support. Special thanks also to Divine and Jumoke for their support.

I am deeply grateful to my parents, Elder Okpako & Mrs. Justina Onoruvie, and my beloved mother-in-law, Prof. Adebola Oyinloye, for their immense support. Your sacrifices and encouragement have meant the world to me.

To my church family, thank you for your love, prayers, and unwavering support. A special appreciation to Deacon Albert, Grandma Joanne, Grandma Claire, Elder Legaire, Grandma Roberts, Grandma Greta, and Grandma Benjamin (RIP)—thank you for welcoming my family and me as your own from the moment we arrived in Canada. Your kindness and love will always be cherished.

To my dear siblings, thank you for your encouragement, prayers, and steadfast support throughout this journey. To a dear brother and my pastor, Adediran Olalekan, thank you so much for having my back from start to finish. Your prayers and care have been a great source of strength.

To my dear husband, Mofoluwaso Adedeji Oyinloye, thank you so much for your unwavering support. Through it all, we made it to the finish line together. Lastly, but most importantly, to my wonderful children—my gorgeous Oyinbolade Ajirioghene Daniel, my eyecandy Oluwadamisi Oghenefega Nathaniel, and my cutie Oluwamumise Ogheneruno Jathniel—thank you for your patience, cooperation, and unconditional love. You have been incredibly understanding and supportive. You are the best I could have ever asked for.

# Contents

<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Contributions . . . . .	4
1.3 Thesis Organization . . . . .	5
<b>2 Literature Review</b>	<b>6</b>
2.1 Online Services . . . . .	6
2.2 Benchmark Policies & Law . . . . .	7
2.2.1 Canadian Law/PIPEDA . . . . .	9
2.2.2 NIST . . . . .	9
2.3 Related Works . . . . .	10
2.3.1 Email Spam Classification . . . . .	10
2.3.2 Honeypots in Leak Detection . . . . .	12
2.3.3 Fingerprinting Techniques . . . . .	13
2.3.4 User Behavior to spam . . . . .	13
2.3.5 The Distinction of our Work . . . . .	14
2.4 Research Hypothesis . . . . .	14
<b>3 Experimental Design and Deployment</b>	<b>16</b>
3.1 Experimental Design . . . . .	16
3.2 Honeypots . . . . .	16
3.3 Websites and accounts creation . . . . .	18

3.4	Spam Logging and Analysis Engine . . . . .	19
3.5	Ethical Consideration & Scope of the Study . . . . .	20
3.6	Deployment . . . . .	20
3.7	Deployment observations . . . . .	23
3.7.1	Email service Provider . . . . .	23
3.7.2	Honeypot Phone number . . . . .	24
3.7.3	Websites . . . . .	25
3.8	Summary . . . . .	26
<b>4</b>	<b>Experimental Result and Statistical discussion</b>	<b>27</b>
4.1	E-mail Spam Dataset . . . . .	27
4.2	Spammers . . . . .	28
4.3	Timeline Analysis . . . . .	29
4.3.1	Aggregate Monthly Spam . . . . .	29
4.3.2	Leak Spam Monthly Analysis . . . . .	30
4.3.3	Spam Hourly Analysis . . . . .	31
4.3.4	Spam Minutes- Interval Analysis . . . . .	33
4.3.5	Day-of-the-week Spam Analysis . . . . .	34
4.4	Nature of the Spam . . . . .	35
4.5	Demographic Analysis . . . . .	36
4.6	Addressing the Base Rate Fallacy in Our Study . . . . .	39
4.7	SMS Spam Analysis . . . . .	40
4.8	Summary . . . . .	42
<b>5</b>	<b>REGULATORY COMPLIANCE AND DATA LEAK FORENSICS</b>	<b>43</b>
5.1	Canadian Anti-Spam Law . . . . .	43
5.2	Consent Exploitation . . . . .	45
5.3	Our Recommendation . . . . .	47
5.4	Analysis Engine . . . . .	49
5.5	Account-based Analysis: . . . . .	49
5.6	Mapping-based Analysis: . . . . .	49
5.7	Leak Analysis Results: . . . . .	50
5.8	Extent of Exposure . . . . .	51

5.9	Discussion . . . . .	51
5.9.1	Improper Use of Information . . . . .	52
5.10	Comparative Analysis Approach for Data Leak Detection . . . . .	53
5.11	Comparative Analysis Approach for Fingerprinting . . . . .	56
<b>6</b>	<b>SUMMARY AND CONCLUSIONS</b>	<b>62</b>
6.1	Summary . . . . .	62
6.2	How the study addresses the research Questions . . . . .	64
6.3	Research Hypothesis: . . . . .	66
6.4	Conclusion . . . . .	67
6.5	Limitation of the Work . . . . .	68
6.6	Future Work & Recommendations . . . . .	69
	<b>Appendix A Appendix</b>	<b>70</b>
A.1	Definition of terms: . . . . .	70
A.2	Publication . . . . .	70
A.3	Research Data . . . . .	71
A.4	Mathematical Formula for Confidence Intervals . . . . .	76
A.5	Data Leak Incident Report: . . . . .	77
.1	SpammerID by Spammer Email Address: . . . . .	85
.2	Spam Count by ID . . . . .	92
.3	Demography & Honeypot Spam Count . . . . .	98
.4	Timeline Counts . . . . .	103
	<b>References</b>	<b>107</b>

# List of Figures

Figure 2.1	Community-Websites of interest . . . . .	8
Figure 3.1	Component Interaction Sequence Diagram . . . . .	17
Figure 3.2	Honeypot Attributes . . . . .	17
Figure 3.3	Website Table Schema . . . . .	18
Figure 4.1	Spam Categories and Counts . . . . .	28
Figure 4.2	Monthly Spam Count aggregate over 12 Months . . . . .	31
Figure 4.3	Growth Curve of Leak Spam Across Months . . . . .	31
Figure 4.4	Hourly Spam Count Over 12-months . . . . .	32
Figure 4.5	5 Minute Interval Spam Capture . . . . .	33
Figure 4.6	Distribution of Dates by Day of the Week . . . . .	35
Figure 4.7	Distribution of Spam by Classification . . . . .	37
Figure 4.8	Demography of Spam . . . . .	39
Figure 5.1	Exploitation of Implicit Consent . . . . .	45
Figure 5.2	Lack of Privacy policy Enforcement . . . . .	46
Figure 5.3	Sample explicit implied consent . . . . .	47

# List of Tables

Table 3.1	Email provider-honeypot distribution	21
Table 3.2	Address Assignment	21
Table 4.1	Top 10 Legitimate Website Spammers and Their Spam Volume	29
Table 4.2	Monthly Spam Aggregation	30
Table 4.3	Distribution of Spams by Day of the Week	34
Table 4.4	Spam Counts by Age Group and Gender	39
Table 5.1	Leaking Source Websites and Leak Recipients	50
Table 5.2	Comparison of Different Leak Detection Approaches	55
Table 5.3	Comparative Analysis of Data Leak Fingerprinting Techniques	58
Table A.1	Honeypot - Website Assignment Table	71
Table .2	Summary of Spam Incidents and Leaks	78
Table .3	Spam Sender Information: TLP	85
Table .4	Spam Sender Information: (UTL=ULS)	85
Table .5	Spam Sender Information (SWT-TSP)	86
Table .6	Spam Sender Information:(SWU=USP)	88
Table .7	Leak Data Table Count	93
Table .9	Service Provider Spam Data: TSP and USP	93
Table .10	Spam Demographic Information	98
Table .11	UTL Table: Hourly Spam Receipt	103
Table .12	TLP Hourly Spam Count	104
Table .13	SWU Hourly Spam Count	105
Table .14	SWT Hourly Spam Count	105

# Chapter 1

## Introduction

### 1.1 Overview

Spams refers to mails or messages received but not subscribed/solicited by a recipient [1]. Spams can be received as emails, text messages (SMS), phone calls, or mails. Spam emails may contain damaging implications that result in data leaks if responded to (e.g., phishing emails), or they may not contain damaging implications, but overwhelm its recipients by flooding their email accounts. *Griffiths* [2] reported that a daily estimate of 3.4 billion spam emails were received globally by Internet users in 2021. He further noted that nearly one billion email addresses were exposed to spammers in the same year, affecting one in five Internet users. In addition, *Kaspersky* [3] reported in 2023 that 45.6% of emails received were spam mails and workers spent an average of 5 to 18 hours yearly sorting out spam emails, which affects their productivity.

Spam emails may be general/untargeted to specific victims from legitimate websites, or can also be targeted to a recipient based on private information connected to data leakage [4]. The appearance of spam emails may be as a result of data leaks, where sensitive information is leaked from an organization to unauthorized parties unintentionally or intentionally [5]. For example the leak cases of, 2016 *CrowdStrike report* [6] emphasized the role of spear-phishing in exploiting system vulnerabilities, the 2022 Twitter breach, where data from 400 million users, including email addresses, was exposed [7], etc. Data leaks are often overlooked, which increases a victim's risk of suffering a data breach, where unauthorized individuals gain access to private information, such as names, addresses, date of birth, contact address, emails, medical records, or financial information [8].

Leaked data amplifies the risk of cyberattacks, including identity theft, fraud, and phishing. According to *Alkhalil et al., Crane, Tunggal* [9–11], these attacks often rely on social engineering techniques,

with the human factor being a critical vulnerability. Although spams received by users are often linked to data leaks, there is little work in the literature attributing spams to specific leaks or differentiating it from spam caused by user behavior.

This lack of reliable indicators creates significant challenges for regulators, organizations, and individuals in holding websites accountable for privacy violations or substantiating claims of data misuse. Additionally, the limited evidence regarding when and how websites might intentionally or unintentionally leak user data exacerbates the problem, contributing to the persistence of spam and privacy breaches. These issues underscore a critical gap in cybersecurity practices linking spam receipt to potential data leakage and leveraging spam analysis as a proactive tool for breach detection.

Several studies have explored spam classification [12], fingerprinting techniques [13, 14], and the use of honeypots for spam detection [15, 16]. However, these approaches have predominantly focused on post-mortem analyses, but not addressing the extent of data leaks. This limitation underscores a significant gap in the development of real-time and proactive leak detection systems. Our research aims to bridge this gap by introducing a novel, low-complexity fingerprinting technique, robust against false positives that leverages spam analysis as an innovative indicator of data leaks. Unlike existing methods that primarily focus on user behavior or direct anomaly detection, our approach emphasizes the analysis of spam receipts, categorizing them based on their frequency, patterns, and nature.

This novel work is vital for differentiating spam caused by data leaks from that stemming from user behavior, assessing how websites handle user privacy, and informing the development of more robust spam filtering models. Additionally, implementing novel early detection methods for data leaks can significantly enhance cybersecurity resilience by mitigating associated risks and providing actionable insights for organizations and regulators. Addressing these challenges not only protects users but also strengthens the integrity of the online ecosystems.

Our technique provides a proactive mechanism for identifying early warning signs of data misuse. Moreover, it offers actionable insights into the spam ecosystem, including the extent of data exposure and its propagation. By enhancing early detection capabilities, our approach not only strengthens data breach prevention efforts but also contributes to a deeper understanding of spam-related data leaks, paving the way for more robust cybersecurity practices.

To expound on our position we ask the following questions:

- (1) Understanding the Link Between Data Leaks and Spam (Novelty of Spam as a Forensic Tool):**

- Can spam serve as a direct indicator of data leaks?

**(2) Effectiveness of Honeypots in Leak Detection:**

- How can honeypots reveal patterns in the way spammers access and use leaked data?

**(3) Analyzing Spam Ecosystems:**

- To what extent does leaked data spread through the spam ecosystem?
- How many entities use the same leaked data to send spam, and what does this reveal about the nature of spam networks?

**(4) Ethics and Intent Behind Data Leaks:**

- Is spam an unintended consequence of negligent data practices, or do some websites intentionally leak user data?

**(5) Novelty of Spam as Forensic Tool:**

- Can analyzing the nature and frequency of spam help detect and attribute intentional data leaks? How can spam forensics enhance current data breach detection and investigation methods?

**(6) Impact on Privacy and Security:**

- Can patterns in spam behavior inform new privacy safeguards or security protocols for websites handling user data?

**(7) Quantifying the Reach of Leaked Data:**

- How can we measure the spread of leaked data based on the number of different senders in spam communications?
- What metrics can be used to assess the scale and impact of a leak through spam patterns?

To address these questions, we setup our honeypots and conducted a real-time evaluation over a 12-month period. We deployed our experiment using 148 honeypots on 370 websites. We created 740 accounts on websites using unique honeypot name fingerprinting. Our honeypots received a significant volume of spam and we applied account-based and mapping-based analyses to identify suspect legitimate websites to leak, extent of data exposure, privacy issues associated with legitimate websites.

The results revealed that some websites failed to adhere strictly to privacy policies, either due to negligence or distributive intent, leading to sharing of honeypot data. By correlating spam trends with our legitimate websites, we identified patterns indicative of both intentional and negligent data misuse. For instance, legitimate websites such as datemyage.com, cheryl.com, dating.com, and zoosk.com leaked our honeypot private information to unknown websites like amolatina.com, amaldate.com, and usadatingz.com etc., resulting in our honeypots receiving spam directly linked to these leaks. Further analyses uncovered distribution patterns and extensive exploitation of honeypot data within the spam ecosystem, with dating.com alone being associated with eight different spammers. Additionally, time-line and frequency analysis provided valuable insights into spam propagation, revealing that the peak of spam receipt occurred between 00:00 and 04:00 minutes of each hour, significantly on Thursdays and the male gender predominately targets of high spam volume. These findings offer a novel layer of intelligence that can enhance existing spam filtering models.

## 1.2 Contributions

The major contributions of this work can be summarized as follows, we:

- introduce a novel data leak tracing system using honeypot profiles and spams, which offers a proactive and real-time method of leak detection. With the increasing public and regulatory (e.g. GDPR, CCPA, PIPEDA) attention on data privacy regulations. There's a growing concern about data handling practices, and our study is offering new evidence or insights into how websites might be indirectly contributing to spam by mishandling data. To substantiate our approach we created 148 honeypot profiles with unique name variants, leading to 740 accounts on 370 websites.
- provide a deeper insight into the spam ecosystem. This detailed forensic approach helped unravel how far leaked data travels and the data-sharing practices behind the scenes. By using spam sender IDs, we measured the spread of data, providing a fresh way to quantify the extent of a data leak. This added an extra dimension to the analysis that goes beyond simply measuring the volume of spam. Additionally, we identified the peak periods of spam activity—at the levels of minutes, hours, and days. This method introduces a new approach for organizations and security professionals to detect leaks and enhance spam filtering models.

- offer a comprehensive understanding of how legitimate websites exploit anti-spam laws, contributing to the increase in spam volume.
- offer a substantial real-time dataset of 12,490 email spam messages over a 12-month period, received from 177 senders across various spam types including phishing, scams, and subscriptions. This dataset can be used by regulators to substantiate privacy policy refinement. Also, the dataset can be shared publicly to aid other researchers requiring email spam dataset.

### **1.3 Thesis Organization**

The rest of the thesis is organized as follows:

- chapter 2 presents related works, applicable Canadian laws & NIST privacy guide recommendation, website of interest and the research our hypothesis,
- chapter 3 describes our experimental design and deployment
- chapter 4 contains our experimental results and statistical analysis discussions,
- chapter 5 presents Canadian Anti-Spam Law (CASL), its impact, drawing insights from the results of the experiment conducted, additionally we also present the forensic analysis of the data leak and the comparison of our approach to existing leak detection & fingerprinting methods.
- chapter 6 presents a summary of our research, how the research questions & hypothesis were addressed by the experiment, our conclusion, limitation recommendations & future work.

## Chapter 2

# Literature Review

This chapter contains a discussion on online services used in the research, applicable laws, related works and the research hypothesis used to expound on our postulation of spam as a forensic indicator for website leak, consent and subscription practices. In section 2.1 we described online services and those used in this research. Applicable law and policy were summarized in sectioned 2.2, while section 2.3 contained related works, and section 2.4 contained the research hypothesis. Content of this chapter was accepted for publication as a paper in the Conference ACM CODASPY25 (Appendix [A](#)).

### 2.1 Online Services

A significant and growing number of consumers today engage in various online services, including e-commerce, education, job searches, transportation, news, social activities, health-related bookings or information, family connections, and forum discussions. These services often include marketplaces that connect users with third-party sellers or service providers. Consumers are drawn to these platforms due to benefits such as the convenience of searching a wide range of products, flexible payment mechanisms, and a variety of product and seller choices [17].

The operations of online businesses vary depending on their specific business models; however, they all share the common characteristic of facilitating the exchange of valuable resources, mostly involving monetary transactions. To crawl the websites selected in this research, we focus on several key areas that are essential to daily life, including e-commerce, accommodation, transportation, employment, health, sports, recreation (such as dating, gaming, video editing, and movie streaming), social networks, education, news, and forums. To effectively utilize these services, all websites required some form of

registration, which often involves the collection of PII. To facilitate reporting similar information on privacy statements—given that some websites share similar business models—we organized the websites into groups we refer to as “communities,” as illustrated in Figure 2.1.

Our e-commerce community includes websites for food, clothing, shoes, home appliances, and general superstores where consumers can purchase home necessities online. The family community encompasses all dating websites, while the transportation community consists of platforms for searching and booking flights, rides, and travel services. The accommodation community features websites where users can search for apartments or houses for rent or purchase, as well as resorts for vacations and hotel bookings. The recreation community is further divided into three categories: audio/video streaming, gaming, and audio/video creation/editing. The education community comprises websites dedicated to learning. In the health community, users can gather health information and make appointments for themselves and their pets. The sports community includes websites that provide information and updates on sports. The forum community consists of websites where participants can ask questions and receive responses, while the news community features websites for reading or viewing news. Appendix A provides a list of the websites and the associated communities, along with links to their privacy policies regarding PII data collection and usage for defaulting websites.

## **2.2 Benchmark Policies & Law**

In light of the increasing incidence of data theft, identity theft, and other cybercriminal activities stemming from personally identifiable information (PII) leaks, data regulation has become essential. Online businesses are now required to provide policy statements that disclose how they use collected data and the level of privacy they offer to users. Regulatory frameworks such as Canada’s Personal Information Protection and Electronic Documents Act (PIPEDA) [18] and the NIST framework [19] have established recommended standards for protecting PII. For this research, we considered the privacy policies of our target websites against the standards set by PIPEDA and NIST considering that the honeypots are located in either Canada or the United States, but specific emphasis was on the Canadian Anti-Spam Law. In this context, we provide a summary of the Canadian PIPEDA and the NIST framework for information privacy handling.

Figure 2.1: Community-Websites of interest

communities	website-of-interest
E-commerce	5,11,14,15,20,24,25,27,33,35,40,43,45,47,52,53,55,57,59,62,65,67,68,70,71,72,74,75,78,80,82,83,84,88,92,93,94,96,98,99,102,105,107,112,116,120,121,125,126,127,128,131,135,138,142,143,144,145,146,147,149,150,151,156,158,159,160,161,162,164,166,167,171,172,173,174,175,176,180,182,183,184,185,186,187,190,194,195,196,200,201,203,210,211,213,214,215,216,218,219,220,221,222,223,224,225,226,227,228,230,232,236,238,242,243,246,247,248,250,251,253,255,256,259,260,263,267,268,271,272,273,275,276,277,278,279,285,286,287,289,290,292,296,297,300,301,303,305,306,309,312,313,314,316,318,319,320,322,325,236,327,328,331,333,334,335,336,337,338,339,340,344,345,346,347,349,353,355,358,359,360,362,363,365,366,367,368,369,370
accommodation	39,78,87,119,137,240,262,282,288,341
education	1,2,3,4,17,23,34,36,56,58,76,90,100,101,108,109,114,129,130,131,154,159,168,179,217,311,323
health	218,257,264,270,295,302,305,308,329,342,343
sport	6,44,51,73,97,146,203,207,265,269
forum	50,60,85,106,261,307
social networks	8,89,199,239,258,291,350
family(dating)	22,30,32,38,46,49,112,113,124,165,178,205,208,235,245,254,324,330
recreation	115,123,124,133,134,136,140,157,163,170,174,181,192,198,206,229,233,240,293,299,357
gaming	7,9,13,18,26,37,66,122,148,189,244,284,304,321,354
movies/audio streaming	12,28,31,41,48,49,60,69,111,117,118,132,139,153,188,197,212,241,281,294,298,307,332,356
video/audio creation/editing/creativity	10,21,79,104,234,280,317,348
job	19,29,63,86,95,141,177,193,249,266
transport	16,43,54,64,103,119,152,169,191,202,209,252,274,283,315,352
News	62,77,81,91,231,237,310

### **2.2.1 Canadian Law/PIPEDA**

In Canada, information privacy laws are divided into two main categories: the Privacy Act and the Personal Information Protection and Electronic Documents Act (PIPEDA). The Privacy Act governs how federal institutions handle personal information and grants individuals the right to access or correct their information. PIPEDA applies to how personal information is managed by federally regulated organizations and private sector businesses operating across provincial or national borders. Key aspects of PIPEDA include obtaining consent for data collection, limiting the collection and use of personal information, ensuring data accuracy, and safeguarding information. Additionally, individuals have the right to access and challenge the accuracy of their data. Three provinces—Alberta, British Columbia, and Quebec—have their own privacy laws for businesses operating exclusively within their borders, offering protections similar to PIPEDA. Provincial laws like the Personal Information Protection Act (PIPA) in British Columbia and Alberta, and Quebec’s Act Respecting the Protection of Personal Information in the Private Sector, further regulate privacy in those regions [18].

### **2.2.2 NIST**

The NIST framework used in the USA and accepted by most industries is akin to the Canadian privacy act PIPEDA. It is a voluntary tool designed to aid enterprise risk management focused on providing data privacy in five core areas: identifying, protecting, detecting, responding, and recovering [20]. The practice of minimizing the use, collection, and retention of PII is a basic privacy principle. By limiting PII collections to the least amount necessary to conduct its mission, the organization may limit potential negative consequences in the event of a data breach involving PII. Organizations should consider the total amount of PII used, collected, and maintained, as well as the types and categories of PII used, collected, and maintained. This general concept is often abbreviated as the —minimum necessary principle. PII collections should only be made where such collections are essential to meet the authorized business purpose and mission of the organization. If the PII serves no current business purpose, then the PII should no longer be used or collected.

Both bodies and other internationally accepted data regulating bodies state that the following PII are to be protected: name, age, ID numbers, financial details, ethnic,origin, blood type,education, medical records, social status, employment information etc. They further require that PII be confidential and as such cannot be shared/sold/used with/to/by third parties except the business has expressly received consent from the individual. Also businesses are required to provide broad content on their websites in

a privacy policy statement which should be easily accessible on how the data is collected, used, stored, share and transferred when required [18–20].

## 2.3 Related Works

The volume of spam is estimated to be as high as 45.6% of global email traffic in these days [3]. The problem of spams have led to significant efforts in the literature focused on spam identification and classification. [12].

### 2.3.1 Email Spam Classification

Today, various machine learning techniques have been used for email spam classification, focusing on the effectiveness and efficiency of different models. A comprehensive review by [12, 21–25], analyzed the application of algorithms like Naive Bayes, SVM, Decision Trees, and K-Means for spam detection, noting the dominance of supervised learning and deep learning models with accuracy exceeding 90%. Also, *Karthika et al.* [26] study aimed to identify the most effective supervised machine learning algorithms for distinguishing spam emails from legitimate ones based on email content features. The researchers applied and compared three classifiers—Naive Bayes, J48 (a decision tree classifier), and Multilayer Perceptron (MLP, a neural network model)—on a dataset of emails collected from the UCI Machine Learning Repository. Using the WEKA machine learning tool, the authors trained these classifiers on a dataset with attributes such as word and character frequencies and capital letter usage. They used a 10-fold cross-validation approach to evaluate model performance based on criteria like prediction accuracy, processing time, and the number of correctly and incorrectly classified instances. The study compared multiple classifiers using the same dataset, providing a fair and robust analysis. Their result showed that J48 achieved an accuracy of 92%, with a balance between speed and correctness. Naive Bayes was the fastest in model building but slightly less accurate (89%). MLP had the highest accuracy at 93% but required significantly more time for training. The dataset, while real-world, may not fully represent the diversity of spam tactics in practice. The MLP classifier's high computational requirements could limit its application to larger datasets. Some methods like Naive Bayes required modifications (e.g., FBL) to handle dependent attributes effectively. while MLP offers superior accuracy, Naive Bayes is efficient for quick model building, and J48 provides a good balance for practical applications. The authors suggested the research to enhance scalability and adaptability for evolving spam techniques. Furthermore, *Erdélyi et al.* [27] investigated methods for improving the classification of

web spam. It highlighted the effectiveness of machine learning techniques combined with cost-effective feature sets. The research focused on improving spam filtering accuracy while reducing computational requirements. It evaluates various feature classes for web spam detection and assesses the trade-offs between computational effort and classification performance. The study evaluated several machine learning models, including Random Forest and LogitBoost, using data from publicly available datasets such as WEBSpam-UK2007 and DC2010. The authors examined both computationally expensive and inexpensive features to identify those that contribute most significantly to classification accuracy. The authors utilized ensemble classification techniques, combining predictions from various classifiers trained on different subsets of features. They tested these methods on two datasets, comparing results based on the area under the ROC curve (AUC) and normalized discounted cumulative gain (NDCG). Their classifier ensemble achieved a 5% improvement in AUC compared to the best results from the Web Spam Challenge 2008. Using only inexpensive content features and a small vocabulary-based bag-of-words representation, the study achieved a 3.5% improvement in AUC. On the DC2010 dataset, they improved the best NDCG for spam classification by 7.5%. Incremental processing of new data, such as link features, remains computationally challenging. Focused on computational efficiency, making the method suitable for real-time applications like web crawling. Achieved high performance across multiple datasets using robust machine learning techniques. The integration of machine learning classifiers with genetic algorithms, as shown in the study by *Fareed and Kumar* [28] research paper outlined a comprehensive approach to email spam classification using ensemble learning, featuring a unique integration of machine learning classifiers and genetic algorithms. The study aimed to improve email spam detection systems by analyzing various classification techniques and introducing an ensemble-based spam classifier that leverages machine learning and genetic algorithms for improved accuracy and efficiency. The authors developed an ensemble classification system incorporating Logistic Regression and Genetic Algorithms. Their methodology involved feature extraction, preprocessing, frequency analysis, and classification to distinguish spam from legitimate emails (ham). Logistic Regression was employed as the primary classifier, with a one-versus-rest (OvR) approach to enhance class prediction, Genetic Algorithms were used to optimize feature selection and improve classification performance by evolving solutions over successive iterations. Initial Logistic Regression classification achieved an accuracy of 86.73%. After integrating the Genetic Algorithm, accuracy improved to 88.93%. The system demonstrated scalability and adaptability, showing increased accuracy as the training dataset size expanded. Genetic Algorithms require significant computational resources, which may limit real-time application.

Despite the progress with MLP and DLP, challenges remain in adaptive spam tactics [29], real-time

spam & false positive/negative management [12, 26], multilingual content [12] and high computational cost of genetic algorithms posing scalability challenges [28].

### 2.3.2 Honeypots in Leak Detection

Another area considered in leak detection is the use of honeypots. Honeypots can serve as decoy systems with the intention to attract attackers and gathering data on malicious behavior. It also enhances security insights and diverts threats from critical assets [30]. Honeypots can be classified into production honeypots, which monitor live environments, and research honeypots, which focus on attack methodologies and vulnerabilities [30]. Spam classification and malicious profile detection in online platforms using honeypot-based frameworks and deep learning techniques have been carried out. *Mendili et al.* [15] used a deep learning-based honeypot framework and achieved a high detection accuracy of 99.23% for identifying malicious social media profiles and spam tweets, although it relied on honeypots, which made it vulnerable to evasion by sophisticated attackers. Similarly, a study focused on Arabic Twitter spammers highlighted the persistence and adaptability of spammers through organized campaigns, revealing the challenges of evolving tactics and the need for continuous updates to detection methods [31]. The Social Honeypot Project, which utilized automated honeypots on MySpace and Twitter platforms, demonstrated high detection accuracy (99.21% on MySpace, 88.98% on Twitter), though it also faced limitation of evolving spammer strategies [32]. *Zhang et al.* [33] proposed the concept of pseudo-honeypots, which was to improve scalability and efficiency in spam detection, offering significant reductions in overhead compared to traditional honeypots, but potentially lacking robustness against more sophisticated attacks. *Karma et al.* [16] carried out a comprehensive survey on intelligent spam email detection, emphasizing the role of AI and ML in analyzing email components, yet highlighted challenges with adversarial spam tactics and the need for further refinement to address sophisticated forms like spear phishing.

*Bhowmick et al.* [34] provided a comprehensive review of recent advancements in content-based email spam filtering techniques, focusing on machine learning (ML)-based methods. It explored feature engineering, classification approaches, and emerging trends in spam detection. Emphasizing the need for adaptive models to handle evolving spam characteristics and challenges like snowshoe spam, which spreads messages across multiple domains to evade detection. They discussed the trade-off between precision and recall, stressing the importance of reducing false positives to avoid mis-classification of legitimate emails, highlighting the limitations of static datasets in addressing evolving spam tactics.

### 2.3.3 Fingerprinting Techniques

Furthermore, in the detection of data leaks, fingerprinting techniques have been used. Avila et al. [13], and Gaikwad et al. [14] reported that several data leak detection methods have been used to detect website data leaks. These methods include Watermarking, which embeds unique identifiers within datasets to trace leaks back to their sources; fake data injection, used to identify unauthorized data exposure, and user behavior Monitoring through log analysis and anomaly detection.

[13, 14] also noted that Watermarking methods are ineffective due to their vulnerability to modification or removal by malicious actors. Fake data injection can reduce dataset utility, limiting its use. User behavior monitoring is often reactive and faces scalability challenges, especially with large, diverse datasets. Sharpira et al. [35] introduced n-gram-based content fingerprinting, which, while useful for detecting intentional leaks, struggles with space requirements and attacker evasion. Bhandar and Kini [36], Zilberman et al. [37] effectively applied fuzzy network fingerprinting to source data leaks and detect leaks over network channels, but were computationally intensive for large organizations. Also, Nayak et al. [29] noted that leak detection is challenging as data often changes in form, complicating traditional fingerprinting techniques.

### 2.3.4 User Behavior to spam

Another area of spam analysis was carried out by Hanna et al., and Rodrigo et al. [38,39], their work focused primarily on user behavior and demographics as drivers of spam. Our work instead highlights spam as an indicator of website privacy lapses, positioning spam receipt as evidence of data leaks. Rodrigo et al. [39] reported that users' online activities, such as participation in forums and shopping sites, greatly influenced spam receipt. This observation is consistent with findings from Bhandari & Kini [36] Almaatouq et al. [40], Ezpeleta et al. [41], and Jamal et al. [42], who all reported that online behavior plays a crucial role in determining the frequency and nature of spam received.

While Studies by Hann et al. [38] and Rodrigo et al. [39] focused on how user behavior, demographics, and email provider practices influence spam distribution, our research shifts the emphasis to the inadequate enforcement of privacy policies by websites. We argue that beyond user activities and interests, spam is also generated due to data leaks and intentional sharing of user information by websites with insufficient privacy safeguards. By analyzing honeypot data, we observed that spam frequently originates from sites failing to protect user data, a factor that Hann and Rodrigo overlooked. Our research highlights the correlation between spam receipt and data breaches, presenting spam as a key indicator

of privacy policy violations and data leaks, which contributes to the spam ecosystem in ways previous studies did not address. Marco et al. [43] proposed a methodology to extract and analyze webpages linked within spam messages, aiming to identify and characterize relationships between these pages and the spam messages. By using a lazy associative classifier, they generated classification rules based on these linked pages to enhance spam detection, complementing existing classification systems like SpamAssassin. In contrast, our research shifts the focus from analyzing linked pages in spam messages to examining spam as a symptom of potential data leaks from websites. Rather than relying solely on page associations for spam classification, our model investigates spam as a forensic indicator of data leaks through intentional or unintentional disclosure by websites. Our approach emphasizes identifying patterns of spam receipt tied to specific honeypots, suggesting data breach indicators that can serve as early warnings, thus expanding the scope beyond mere message-page associations to a deeper analysis of data leaks' potential impact on users.

### **2.3.5 The Distinction of our Work**

Previous works have discussed user behavior to spam receipt, spam classification and detection strategies using static dataset. To the best of our knowledge, no academic work has directly examined the relationship between spam and data leaks. Our research differs from the existing research in that it provides early warning real-time leak spam detection on a low-complexity fingerprinting. Also, our approach allows us to link leak spams to legitimate websites within our website pool and providing a deeper insight to the spam eco-system, while revealing potential privacy violations. Our technique shifts the focus toward spam as a signal of website responsibility failures rather than just a nuisance and a user behavior factor.

## **2.4 Research Hypothesis**

We investigated how spam received via honeypots can serve as a direct indicator or consequence of data leaks, an area that is not typically explored in traditional data breach analyses. To expound our position, we determined the following: (a) spam as a forensic tool, a direct indicator of data leaks (b) Honeypots, revealing patterns in the methods spammers use to access and utilize leaked data (c) the extent to which leaked data spreads and the number of entities leveraging the same leaked data to send spam within the spam ecosystem (exploring the nature of spam networks), (d) timeline information that can inform intelligent spam filtering (e) how patterns in spam behavior can inform new privacy

safeguards or security protocols for websites that manage user data. We use the following hypothesis to establish our postulations:

- (1) Users who engage in healthy Internet activities—defined as visiting and interacting solely with their assigned websites—will still experience data leaks over time due to signed-up accounts on websites that may not effectively enforce their privacy policies.
- (2) There exists a potential link between spam receipt and websites associated with data leaks.
- (3) The spammer's intent (distribution or harvesting) can be inferred from the nature of the received messages.
- (4) The receipt of spam by users is not directly correlated with their genuine voluntary consented subscriptions at the time of account creation on websites.

## Chapter 3

# Experimental Design and Deployment

In this chapter, we present our method of utilizing honeypots for spam collection, the deployment and observations. The content of this chapter is accepted for publication as a paper in the conference ACM CODASPY25.

### 3.1 Experimental Design

Our experimental approach consists of five main components: the honeypot profiles, websites selection, accounts creation, spam logging, and analysis engine. Details of those components are explained below. The Figure 3.1 is a sequence diagram describing the interactions between the component.

In the Figure 3.1 honeypots sets up its profile, is assigned 5-unique website, creates account on the assigned websites, detects and logs the spam, performs different kinds of spam analysis, to investigates potential information leaks and privacy issues.

### 3.2 Honeypots

To track spams, we created 148 honeypot profiles. Each profile is uniquely identified by its email address. To create those email addresses, we activated 20 T-mobile phone numbers since many email services required valid phone numbers for verification and limited the number of email addresses that could be created with each phone number. To reflect Canadian's demographic distribution, 90% of the honeypot profiles were designated with Canadian nationalities, while the remaining 10% were designed with other nationalities. The age distribution of the profiles spanned from 13 to 75 years uniformly, ensuring representation across all age groups. Balanced male and female profiles were included for each age category

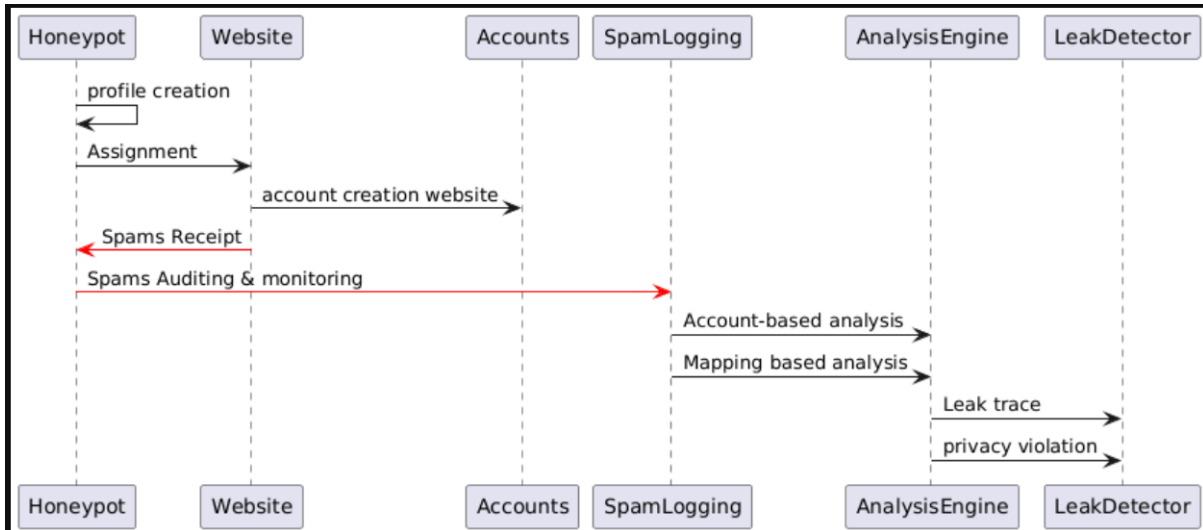


Figure 3.1: Component Interaction Sequence Diagram

Each honeypot profile was assigned with several key attributes including first name, last name, gender, date of birth, address, default email service provider, phone number, home address, educational, and employment status. These attributes provided a comprehensive identity design for each honeypot, allowing for detailed analysis of targeted spam activities as depicted Figure 3.2.

Column Name	Data Type	Constraints
id	INT	PRIMARY KEY, AUTO_INCREMENT
first_name	VARCHAR(50)	NOT NULL
last_name	VARCHAR(50)	NOT NULL
gender	VARCHAR(10)	NOT NULL
date_of_birth	DATE	NOT NULL
default_email_provider	VARCHAR(50)	NOT NULL
phone_number	VARCHAR(15)	NOT NULL
home_address	VARCHAR(255)	NOT NULL
educational_status	VARCHAR(100)	NOT NULL
job_status	VARCHAR(100)	NOT NULL

Figure 3.2: Honeypot Attributes

To further differentiate the name attributes (first and last names), unique name variants were created for each account. This was achieved by incorporating middle names, adding underscores, creating

compound names, or including/omitting letters from the default names. For example: one honeypot’s default name was Kuban Manik and its variants were Kubian Manik-lev, kubannie, kuban-greaves manike, kubanie manikc etc. Honeypot information is available in appendix A.

### 3.3 Websites and accounts creation

To mimic the behavior of real users, we selected 370 websites and grouped them into 12 communities. These communities are e-commerce, accommodation, transportation, jobs, recreation, news, education, forums, sports, health, social networks, and family(see chapter 2). To determine which websites would be included in the research, we analyzed the hit information collected for each website, giving preference to websites that had hits > 100,000 based on our search on google.com between the months of June 2023 and August 2023, noting these hit statistics fluctuates as reported by *Gao et al.* [44]. We made sure each community is represented by at least 8 websites so that each community is properly represented. Figure 3.3 show the website schema and the detailed website listing is available in Appendix A.

Table: communities

Column Name	Data Type	Constraints
id	VARCHAR(50)	PRIMARY KEY
name	VARCHAR(50)	NOT NULL, UNIQUE

Table: websites

Column Name	Data Type	Constraints
id	INT	PRIMARY KEY, AUTO_INCREMENT
url	VARCHAR(255)	NOT NULL, UNIQUE
community_id	VARCHAR(50)	NOT NULL, FOREIGN KEY REFERENCES communities(id)
description	TEXT	

Figure 3.3: Website Table Schema

For each honeypot profile, we manually created 5 unique accounts on 5 different websites. Therefore, we had 740 (148 X 5) honeypot accounts created in total. It was our interest to create as many accounts as possible, but based on the honeypots profiles we have and having to cover a large number of monitored, we decided to limit the number to 5 accounts per profile given the amount of manual work involved, while focusing on effective account management, monitoring and tractability of data leaks.

With in mind, we linked each website with 2 different honeypot accounts after the honeypot accounts creation. It is worth noting that we used unique name variants on each of the 740 website account created. For example, the profile name Jamie Russell has 5 accounts on 5 different websites with names Jamie Russell , Jamlie Ruselle, Jiamie Ruissell, Jamiee Ruseille, and Jaimie. Therefore, all the 740 accounts we created has distinct names. This helped us quickly identify which website leaked the email address to spammers if the spams received contain the names of the intended recipients. Appendix [A](#) we present the table detailing the assignment, due to privacy concerns the email address and other personally identifiable information excluding the names have been left out in the table for each honeypot. This is to ensure the honeypots data is not exploited by fraudulent actors.

### **3.4 Spam Logging and Analysis Engine**

After honeypot accounts creation, we actively monitored the 148 email addresses associated with those accounts. We systematically logged and audited those email accounts. This log has several key columns, including a unique spam ID, spam sender's name and email, spam receiver's name and email, the content of the spam message, and the receiving date/time of the spam. See the research data in Appendix A for more information on the spam dataset.

To analyze the spam received, we employed two analytical methods: account-based analysis and mapping-based analysis. Account-based analysis involves identifying the accounts associated with the spams, whereas mapping-based analysis looks for intersections of mapped websites.

For spams that are addressed to specific names (targeted spams), we applied account-based analysis to map spams to accounts through the unique name/email combo for each account. Since each account was only linked to one website, this helped link spams to the websites who leaked the account information. Account-based analysis was also employed to identify trends and interpret the timelines, frequency, and nature of all spam received by the honeypots.

In the cases where spams did not contain specific names (untargeted spams), we applied mapping-based analysis. In this approach, we collected the recipient emails for each spammer and looked for intersections in business type and third party relations of their mapped websites. Consequently, we were able to establish a relationship between the suspected websites who leaked data and the recipients of the leaked data.

The account based and mapping analysis provided insight to spammer's activity peak period- monthly, day-of-the-week, hourly, by minutes intervals and the growth curve of the spams (see Appendix [A](#) for

breakdown details). The analyses also enabled us to assess the impact of exposure, evaluate the forensic value of spam in identifying data leaks.

We also gained insights into the spam ecosystem, the intent of the spam sender and legitimate website (such as data harvesting or distribution), and the extent of the leaks through sender information and spam patterns. Thereby providing intelligent insight for spam filtering models.

### **3.5 Ethical Consideration & Scope of the Study**

In our study, informed consent was not necessary, as all profile data were designed as honeypots and did not pertain to any identifiable individuals. The websites accessed did not require any permission to utilize their services. This study focuses on evaluating online services that address fundamental human needs, explicitly excluding pornography websites, with the selected websites representing these essential needs. The number and ages of the honeypots utilized span multiple generations, reflecting the diverse demography engaged with online services across various communities. Considering the vast number of websites available on the internet, our study encompasses a wide range of parallel websites that are pertinent to key communities relevant to human existence.

### **3.6 Deployment**

We conducted the experiment between January 2024 and January 2025 for 12-months, involving 148 honeypot profiles, 370 websites, and 740 honeypot accounts.

Each honeypot was assigned one email service provider from the following email service providers - AOL, Gmail, YahooMail, outlook, Hotmail, ZohoMail, Mailfence, and GMX. Table 3.1 shows the distribution of honeypots to email service provider. The irregularities observed with the email service providers arose from the challenges encountered during distribution.

After distributing the websites, home addresses were assigned to the profiles, The allocation was skewed towards one address due to the necessity of replacing addresses for certain websites that specifically required a U.S. address; consequently, other U.S. addresses were assigned fairly sequentially to minimize the likelihood of sequential profiles sharing the same address. Noting that not all websites of interest required address information during the account registration process. The details of the assigned addresses are provided in Table 3.2.

Also, each honeypot was sequentially assigned one of 20 mobile numbers purchased and activated from service provider T-Mobile [45]. We created 740 honeypot accounts (5 accounts for each profile)

Table 3.1: Email provider-honeypot distribution

email service provider	number of honeypot
aol	17
yahoo	23
gmx	12
gmail	19
outlook	21
hotmail	19
zohomail	17
mailfence	20
total	148

Table 3.2: Address Assignment

Address	no of Honeypot Assignment
*** Bolvd. Newman LaSalle QC CA	17
*** Rue Raudot Montreal QC CA	15
*** Rue Thomas Dubuc Longueuil QC CA	17
*** Cool Springs Pt Loganville GA USA	28
*** willow beach platsburgh NY USA	17
*** Venable Street Richmond Virginia USA	15
*** Melrose Cresent Eastern Passage Nova Scotia	14
*** Evancrest Gardens NW Calgary Alberta CA	12
*** Richmond, VA, USA	13

to cover 370 websites, so that each website has two honeypot accounts. To establish a baseline, we monitored each phone number for spam over the first 30 days before linking it to any profiles. During the honeypot accounts creation and spam monitoring period, we ensured that every honeypot followed the rules below:

- (1) honeypots did not subscribe to website mailing list voluntarily provided by email service provider or websites of interest at registration
- (2) Honeypot could visit only 5 websites and could not interact outside their assigned websites
- (3) Honeypot emails could only be used as recovery or primary email as required (mailfence.com and zohomail.com).
- (4) Honeypot age have only one male and female representation with each having unique attributes excluding age, so that name, gender, email address, job, date of birth, educational status where unique to each honeypot, while home address and phone number were shared amongst honeypots.
- (5) Honeypot profile on each website was unique (2)

In the period of monitoring and auditing, our honeypots engaged in various activities on the accounts, including adding, liking, and removing items in e-commerce carts; searching for products; completing course modules; and browsing educational content. We also searched for apartments in the U.S. and Canada by type and location, as well as transport options domestically and internationally, while ensuring we did not click on advertisements/pop-ups in the course of interacting with the websites.

To accurately identify the spam received, we grouped and coded the spam sender information into four categories (See Appendix [A](#) for the spam dataset):

- (1) Targeted Leaked Spam (TLP): This category includes spam received from websites where no account was created but contains our unique primary key information from our honeypot profiles (honeypot variant names), email addresses, or other Personally Identifiable Information (PII). These emails were analyzed using account-based analysis.
- (2) Untargeted Leaked Spam (ULS): These spam emails come from unsubscribed websites and the content of the spam did not contain unique primary key information (honeypot variant names). These emails were analyzed using mapping analysis.
- (3) Untargeted Service Provider Spam (USP): These emails are received from our websites of interest and email service providers but do not contain any unique honeypot information, they were analyzed using account-based analysis.
- (4) Targeted Service Provider Spam (TSP): These emails are received from our websites of interest and email service providers and contain unique information from our honeypots.

All received spam emails were further grouped into the type of spam, which we referred to as classifications:

- Adverts/Offers: One-way communications, intended to inform potential customers about products and services, including details on how to obtain them, as well as promotional offers related to specific products, product lines, brands, or companies aimed at encouraging customer engagement or sales.
- Scams: These can take various forms, including lottery scams, advance-fee fraud, investment scams, fake job offers, and romance scams. Scams may occur through various methods, including phone calls, emails, fake websites, and in-person interactions.
- Phishing: This category encompasses fraudulent attempts to obtain sensitive information by masquerading as a trustworthy entity in electronic communications. This includes unsolicited surveys

sent without prior consent. We looked for specific language or requests indicating an intent to obtain personal information, such as requests for login credentials or personal details for supposed prize claims, update, feedback etc.

## **3.7 Deployment observations**

### **3.7.1 Email service Provider**

The research aimed to exploit twelve email service providers: ProtonMail, Yandex Mail, Tutanota Mail, Gmail, Mail.com, Outlook, Yahoo Mail, AOL Mail, GMX Mail, Hotmail, Mailfence, and Zoho Mail. We reviewed their features and tested their sign-up criteria, subsequently creating email addresses for all our honeypots across these providers. The following observations were made:

**ProtonMail:** We were unable to access the honeypot profiles created on this service because ProtonMail locked all associated addresses, citing suspected misuse or abuse. While the account creation management system effectively flagged potential issues with multiple account creation, this resulted in unintended consequences for our legitimate research activities. We submitted an appeal to review the suspension decision but have yet to receive a response. This led us to exclude this service provider from our study. As a result, we redistributed the honeypots to more reliable email service providers

**Yandex Mail:** All Yandex Mail addresses created for our research failed to receive any emails. Testing with a randomly generated address confirmed the issue, leading us to exclude this service provider from our study. As a result, we redistributed the honeypots to more reliable email service providers

**Tutanota:** Profiles created on Tutanota could not be accessed, even with valid usernames and passwords. Attempts to recover the email addresses using the recovery key provided during profile creation were also unsuccessful. We observed that the lengthy recovery keys, consisting of 64-bit hexadecimal values in lowercase, could be vulnerable to security risks if exposed through social engineering. Despite these concerns, we were unable to recover the profiles. To address this issue, we recommend implementing a multi-factor authentication system that combines a physical token with recovery codes. Leading us to exclude this service provider from our study. As a result, we redistributed the honeypots to more reliable email service providers.

**Yahoo Mail, AOL Mail:** These providers allowed profile creation upon providing a valid verification code via SMS. Different phone numbers were used for various honeypots under these services.

**Mailfence:** This provider permits each recovery email to be used only once for creating a honeypot profile. Unlike other email service providers in this study, Mailfence does not allow the reuse of

the same emails for more than one profile account. In contrast, Gmail, Outlook, and Hotmail. AOL, yahoo.com, gmx, etc., accept one email address as an alternate recovery address for multiple profile accounts creation. We also observed that Mailfence limited the number of profile addresses that could be created on the same network, even when different machines were used. The daily limit for profile creation decreased from four to one before alerting us that the account creation limit was reached. This commendable approach helps track network information effectively, thus protecting the system from potential abuse and slowing down the rate of fake account creation.

**GMX.com and Mail.com:** Despite being provided by the same company, we could not access the profiles created on Mail.com and redistributed the affected honeypots to other email service providers. For GMX.com profiles, SMS verification was requested after account and website creation, but we did not receive the verification codes for about three weeks, despite our SIM cards having valid subscriptions. After this period, we successfully received the codes upon retrying, indicating potential technical challenges with the request verification service at the initial time.

**Outlook and Hotmail:** We observed that these services permanently deleted emails in the spam folder after 10 days. To address this, we moved the spam content to the inbox. Additionally, we observed that subsequent emails from the same senders were not placed in the spam folder, a trend consistent across all email service providers tested.

**Zohomail:** This is a secondary email service provider, which required a primary email address to create accounts.

### **3.7.2 Honeypot Phone number**

**Google Voice Accounts:** To increase our pool of mobile numbers while reducing the cost of maintaining a large number of phone numbers for the one-year duration of the research experiment, we created Google Voice accounts and were assigned numbers. However, we encountered several issues:

**Incompatibility with Email Providers:** We found that the Google Voice numbers could not be used with any email service provider, including Gmail, with which Google Voice shares a parent company. To resolve this, we reassigned our physical phone numbers to all honeypots. This redistribution increased the number of honeypots sharing the same phone numbers, but it did not affect the results regarding SMS spam received during the experiment because each honeypot had different names.

**Registration Issues with SIM Card:** We were unable to register the physical SIM card number labeled 9 for Google Voice as it did not receive the SMS verification code. We suspect that a previous owner of SIM number 9 still had an active Google Voice account associated with that number. Despite the

number being reissued to a new user, it was impossible for us to use it on this service. Additionally, we encountered a similar issue with Uber Eats, where the SIM card number was still reported to be in use by another user named Austin.

**Telecommunication Company Interaction:** Interactions with the telecommunication company representatives indicated that there is currently no way to monitor where users register their phone numbers, making it difficult to clean such associations at the point of exit or release of the number. We believe it is crucial to implement additional measures to ensure that any data connected to phone numbers is properly destroyed or cleaned before reissuing them.

**Spam Receipt Inspection:** The 30-day observatory period of the SIM card, we inspected it for spam receipt as a result of previous ownership. The details of SMS spam received are reported in Chapter 4.

### **3.7.3 Websites**

The honeypots created for this research are categorized as high-interactive honeypots; however, we encountered several challenges, as outlined below:

**Credit Card Requirement:** The Hellofresh website required a credit card as part of its registration process. Although we did not use this website as one of our primary sites, we attempted to replace it while keeping a record of the information provided for spam and leak tracking. Notably, despite the incomplete registration process due to the inability to provide a credit card, the website retained the honeypots personal information and subsequently sent several spam emails to the associated honeypot email addresses. This observation led us to conclude that while the key information required for account registration was provided, the additional credit card requirement was enforced by the service provider as compulsory, albeit not explicitly required at sign-up.

**Home Address Requirements:** During account creation, we observed that all communities, except for news, forums, and recreation (video/gaming/movies/audio streaming), did require a home address at the point of account creation. The other communities, including health, e-commerce, education, family, transportation, and accommodation, required a home address. Some required it at the time of registration, while others permitted entry during subsequent interactions.

## 3.8 Summary

This study, conducted between January 2024 and January 2025, aimed to evaluate online services addressing fundamental human needs, excluding pornography websites, by deploying 148 honeypot profiles across 370 websites and utilizing 740 honeypot accounts. The selected websites represented key communities relevant to human existence, and the honeypots reflected diverse demographics spanning multiple generations. We explored email services such as AOL, Gmail, YahooMail, Outlook, Hotmail, ZohoMail, Mailfence, and GMX, noting irregularities and challenges during distribution, which led to the exclusion of several providers like ProtonMail, Yandex, and Tutanota due to issues with account access and verification. The honeypots were assigned unique demographic attributes, including gender, name, and job, with shared addresses and phone numbers, while strictly adhering to rules that prevented voluntary subscriptions and restricted website interactions. Each honeypot engaged in various activities, including product searches and browsing, while spam received was categorized into four groups: Targeted Leaked Spam, Untargeted Leaked Spam, Untargeted Service Provider Spam, and Targeted Service Provider Spam. The spam was further classified into adverts, scams, and phishing attempts. Despite the challenges encountered, such as difficulties with phone number registration and website-specific address requirements, the research provided a comprehensive examination of online service interactions and spam receipt across different email providers, highlighting privacy risks and system vulnerabilities.

## Chapter 4

# Experimental Result and Statistical discussion

In this chapter, we present the results obtained from our experiments. We begin with a summary of spam collection, followed by analyses of spam data, including timelines with confidence intervals, nature, and demographic distribution based on age group and gender. The content of this chapter is accepted for publication as a paper in the conference ACM CODASPY 2025 (Refer to appendix [A](#)).

### 4.1 E-mail Spam Dataset

Over a 12-month period, we audited a total of 12,490 spam messages originating from 177 distinct senders. It is also important to note that some spam messages may have been filtered out by the email service provider, which implements its own spam filtering mechanisms. Notably 9,406 (75.3%) were received from 147 (39.7%) legitimate websites, coded as USP & TSP (please [3.6](#) for the definitions), while 3,084 stemmed from leak of honeypot Personal information to 30 leak recipient websites, coded as (TLP & ULS) (please [3.6](#) for the definitions) and shown in figure [4.1](#). Our honeypots did not interact or have accounts with these leak recipient websites. The 12,490 spams were received by 107 honeypots (72%), while the remaining 41 honeypots (28%) did not receive any spam emails. (See Appendix [A](#) for detailed information)

The high volume of untargeted spams at 7061 (56.5%) (USP) compared to targeted spam at 2345 (18.8%) (TSP) from legitimate websites, as shown in figure [4.1](#), suggests that websites where users already have accounts tend to generate more untargeted spam than targeted spams. This is likely because users are existing account holders on these online platforms.

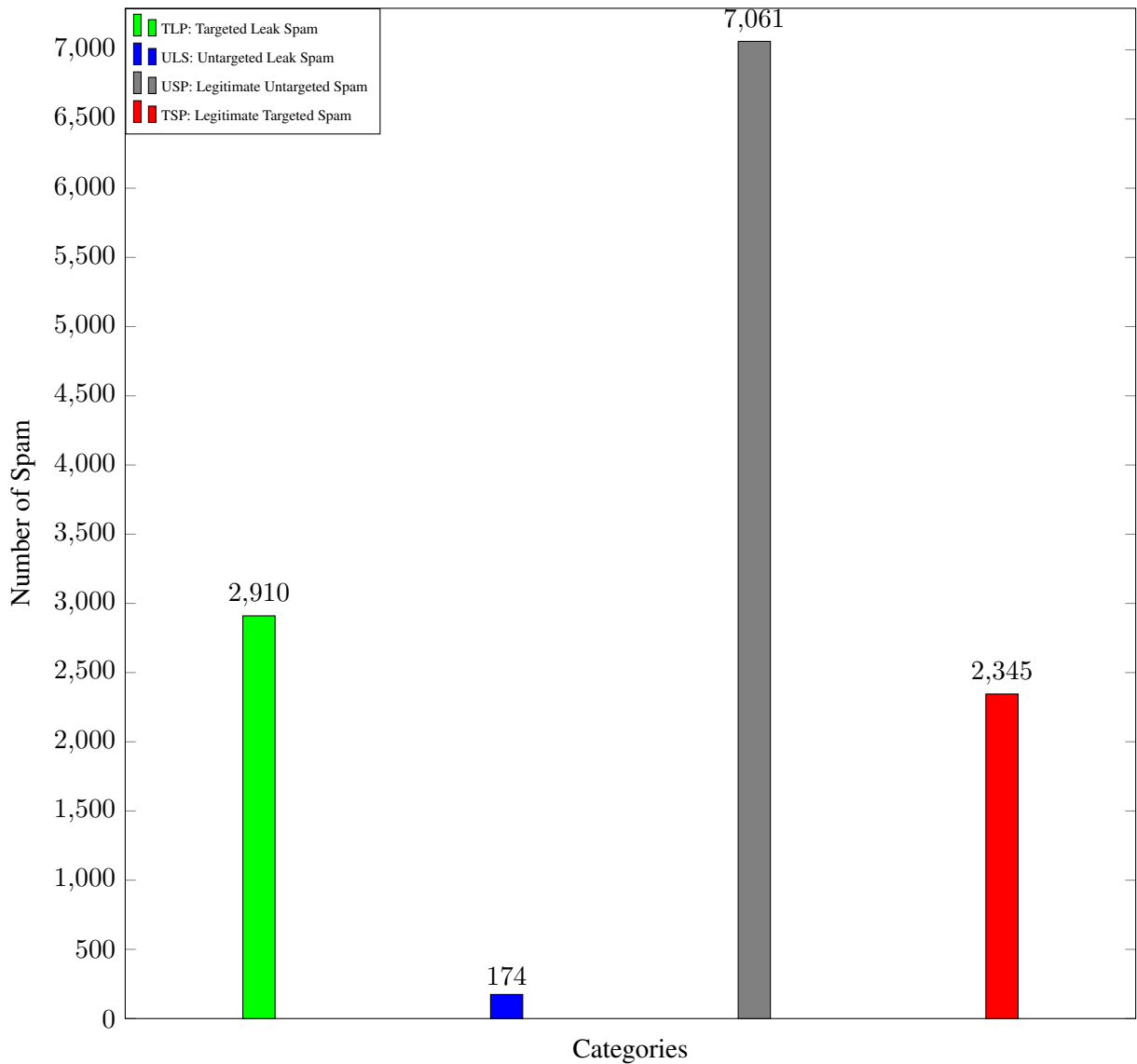


Figure 4.1: Spam Categories and Counts

## 4.2 Spammers

We identified the top 10 spam senders, which accounted for a significant proportion of the spam emails from our legitimate websites. As shown in Table 4.1, the top 10 spam senders from our legitimate websites are [childrenspplace.com](#) (1861 spams), [poshmark.com](#) (511 spams), [webmd.com](#) (403 spams), [bestbuy.com](#) (347 spams), [toyrus.ca](#) (296 spams), [lowes.com](#) (296 spams), [flipboard.com](#) (289 spams), [cafepress.com](#) (280 spams), [houzz.com](#) (256 spams), [datemyage.com](#) (239 spams), [linkedin.com](#) (239 spams), and [tasteofhome.com](#) (127 spams) respectively.

Given that spams are coming from domains that are otherwise perceived as legitimate, it may challenge traditional spam detection algorithms, which often rely on domain reputation and historical sender patterns. It suggests that spam filtering systems need to incorporate more sophisticated mechanisms,

like behavioral analysis, instead of just relying on sender reputation.

Table 4.1: Top 10 Legitimate Website Spammers and Their Spam Volume

Top Spammer	Spam Count	Percentage
childrensplace.com	1861	14.89%
poshmark.com	511	4.09%
webmd.com	403	3.23%
bestbuy.com	347	2.79%
toyus.com/lowes.com	296	2.36%
flipboard.com	289	2.31%
cafepress.com	280	2.24%
houzz.com	256	2.05%
datemyage.com/linkedin.com	239	1.91%
tasteofhome.com	127	1.01%

Furthermore, a comparison between targeted leak spams (23.2%) and untargeted leak spams (1.39%) revealed that users are significantly more likely to receive spam because of leak that contains personal information. This disparity can be attributed to spammers leveraging personalized content to capture the user’s attention more effectively. As spams containing personalized information appears more credible to recipients, reducing suspicion and increasing the likelihood of engagement. Detailed breakdown of all spammers is available in appendix A.

## 4.3 Timeline Analysis

### 4.3.1 Aggregate Monthly Spam

We aggregated the total volume of spam monthly from February 2024 to January 2025 as presented in Table 4.2. Spam volumes exhibited a growth trend starting in February, with peak activity recorded between March and May. A decline followed in June, July, and August, which can be attributed to our effort to unsubscribe from defaulting websites registered by 65 (43%)honeypots. However, a notable spike in spam activity was observed in September and November, with sharp drop in the months of October, December and January. The sharp rise seen in the month of September and November indicating an increase in spam volume from some legitimate websites such as seen with the case of childrensplace.com, poshmark.com, linkedin.com etc., and also new spam campaigns appearing, as seen with the case of tuango.ca etc. Figure 4.2 shows the growth curve.

The growth pattern of spams for the total spam received in Figure 4.2 can be described as seasonal and intervention-driven. The rise in spam volumes from February to the peak in March through May suggests a seasonal increase, possibly tied to heightened online activity during the early months of

the year (e.g., post-holiday promotions or targeted campaigns). The decline in June, July, and August correlates directly with our un-subscription effort. This demonstrates that proactive measures can significantly reduce spam influx and also that a second opting out from website implicit subscription was effective towards spam reduction.

We further determined the uncertainty in spam detection rates by calculating the mean spam rates of the aggregate spam volume for the 12-month period (99% confidence interval) for figure 4.2 (see appendix A for the mathematical formulas). The mean spam count monthly is 1024.83, the standard deviation (s) was 590.63, margin of error (ME) was 439.36, so that the 99% confidence interval (showing the lower and upper bound) were at 585.47 and 1,464.19 respectively.

Table 4.2: Monthly Spam Aggregation

<b>Month</b>	<b>Spam Count</b>
<b>Feb</b>	559
<b>Mar</b>	2038
<b>Apr</b>	2115
<b>May</b>	1579
<b>Jun</b>	989
<b>Jul</b>	877
<b>Aug</b>	422
<b>Sep</b>	1015
<b>Oct</b>	521
<b>Nov</b>	1336
<b>Dec</b>	478
<b>Jan</b>	369

### 4.3.2 Leak Spam Monthly Analysis

Also, we present the monthly growth pattern of spams received because of leak. These spam volume constitute 24.6% of the total spam dataset as shown in Figure 4.3. From these spam senders, there was no un-subscriptions done. From the figure 4.3, a non-linear and irregular growth pattern was seen, with alternating periods of growth and decline. There was a significant increase in the leak spam count from February (29) to April (353), showing an upward trend. The spam data fluctuation between sharp peaks and declines were seen between May and January. In May, the spams received dropped to 215 after April's peak at 353. June rebounded to 373, followed by a decrease to 321 in July. A sharp drop occurred in August (16), marking the lowest value in the dataset. A major surge is observed in September (519) and November (833) been the highest points in the dataset, indicating a significant anomaly or event. The number of spams dropped again in October (85) and a steady decline was seen in the months of

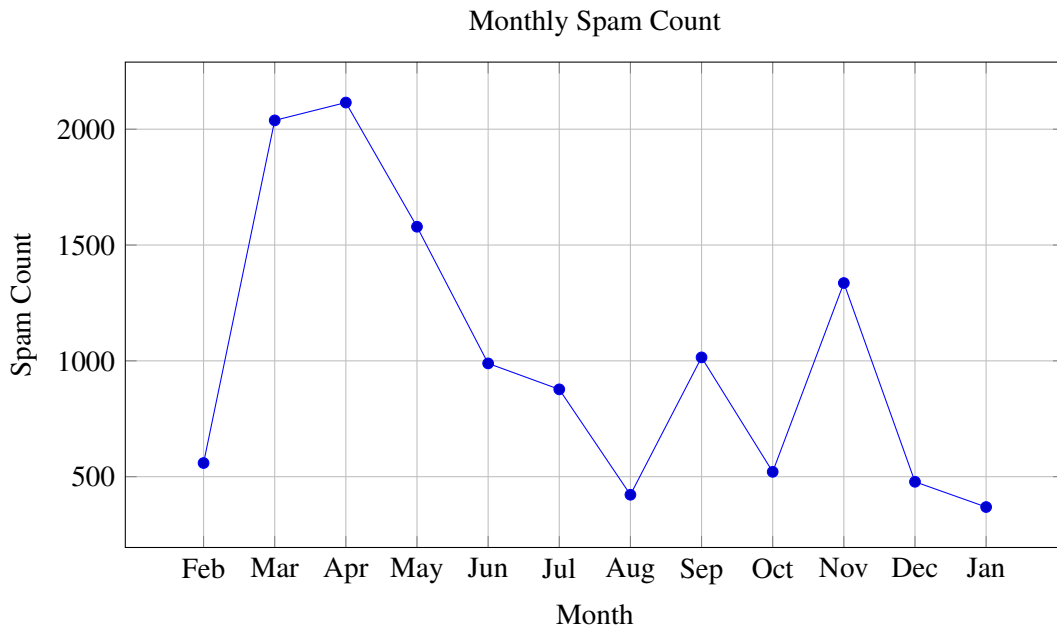


Figure 4.2: Monthly Spam Count aggregate over 12 Months

December and January.

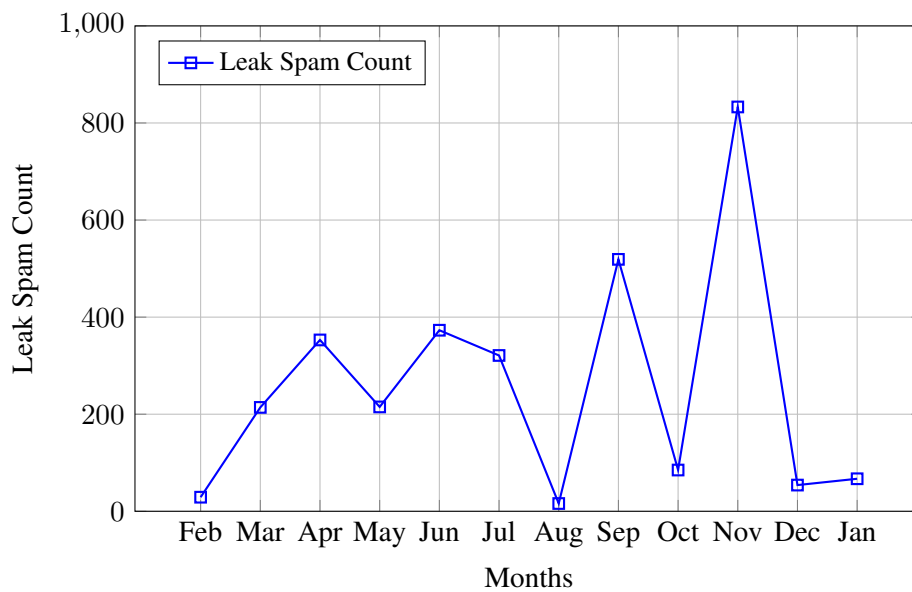


Figure 4.3: Growth Curve of Leak Spam Across Months

### 4.3.3 Spam Hourly Analysis

Furthermore, we analyzed the timeline information that was captured during the audit process. Figure 4.4 shows the hourly count of spams received over the 12-month period. The results indicated a peak of spam received between 1 - 2pm, 1-2am and 10 am (all recordings were done in the Eastern time zone). A decline is seen between 3am to 6am, as spam receiving commenced gradual rise between

7am to 9am. After 10am we see a decline that becomes steady between 11am and 12pm with sharp rise between 1 - 2pm and then a decline. The pattern in spam activity can be attributed to a combination of human behavior and automated spammer strategies.

Peaks at 1–2 PM and 10 AM may coincide with periods of high user activity, while peaks at 1–2 AM could reflect late-night spam campaigns. Given that Asia and North America consistently rank as the regions with the highest number of spam distributors globally [46, 47], the time difference likely influences the timing of spam receipt. Additionally, spammers may exploit reduced online monitoring during off-hours, as the spam received spans multiple senders. [48].

Declines between 3–6 am suggest periods when user engagement is minimal, and spammers may optimize their resources for more active times. This result revealed that high spam volumes are sent within the 1 - 2 am and pm hourly time frame., which gave us a high level summary of the spammer activity.

The hourly spam rate indicates that spammers often deploy automated bots to send emails at times of expected high engagement, leveraging analytics on user behavior. This aligns with the findings of *Levchenko et al.* [49]. For instance, medium.com sent spam messages at 11:10 AM, while biteable.com did so at 3:47 PM. The observed pattern of rises and declines suggests a deliberate timing strategy to maximize email engagement while evading spam filters that detect sudden surges. A detailed breakdown by spam category is provided in Appendix A.

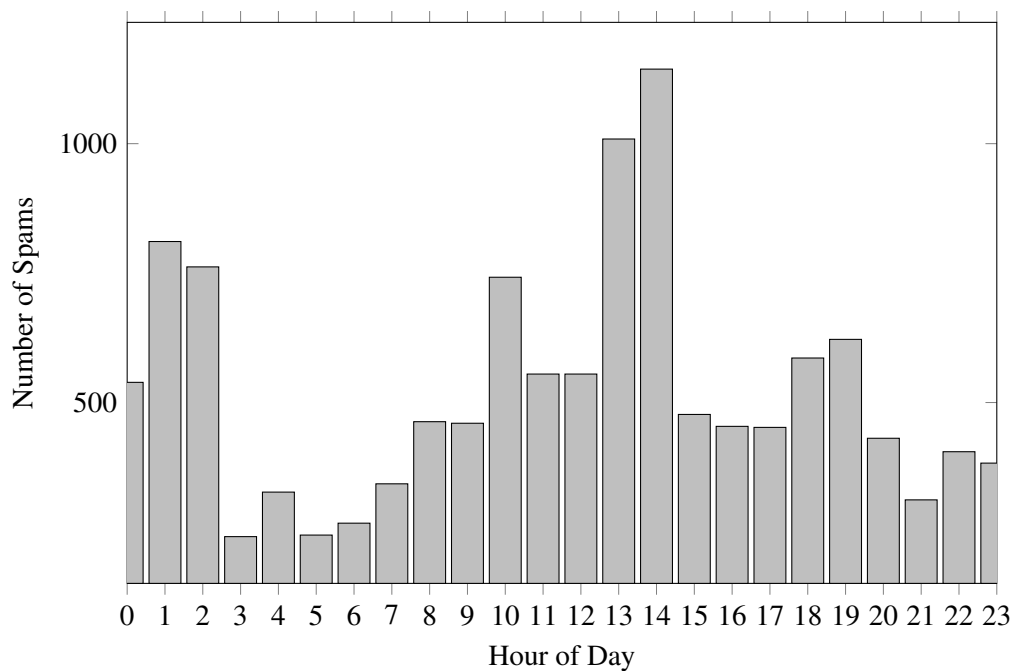


Figure 4.4: Hourly Spam Count Over 12-months

### 4.3.4 Spam Minutes- Interval Analysis

Furthermore, we performed 5 minute-interval based on count of the spams received to distinctly pinpoint the peak of spams receipt over 60 minutes, through our 12-month period. The chart is presented in Figure 4.5. The minutes-interval spam count revealed a distinct pattern of activity, with a significant peak in spam receipt during the 00–04 minutes of each hour, recording 1,748 spams.

Following the 0-4 minutes peak period of spam, the spam count dropped throughout the hour. A steady reductions in spam within the hour is most likely indicative that spammers tend to schedule their spams to be sent out at the beginning of the hours of their local time. In order to avoid being labeled as spams, they may slow down their spam sending process by sending smaller batches of spams every few minutes. The sharp spike of 145 spams seen between 30-34 minutes occurred since some part of the world has timezone offset of 30 minutes instead of 60 minutes (e.g., Newfoundland, Canada). Detailed spam volume by spam category is available in appendix A

These observations suggests that some spammers tend to schedule their campaign at a fixed time rather than a random time. The declining spam trends interspersed with sharp spikes could be associated with spammers targeting different geographic zones with varied schedules, possibly aligned with time zones or regional activity patterns.

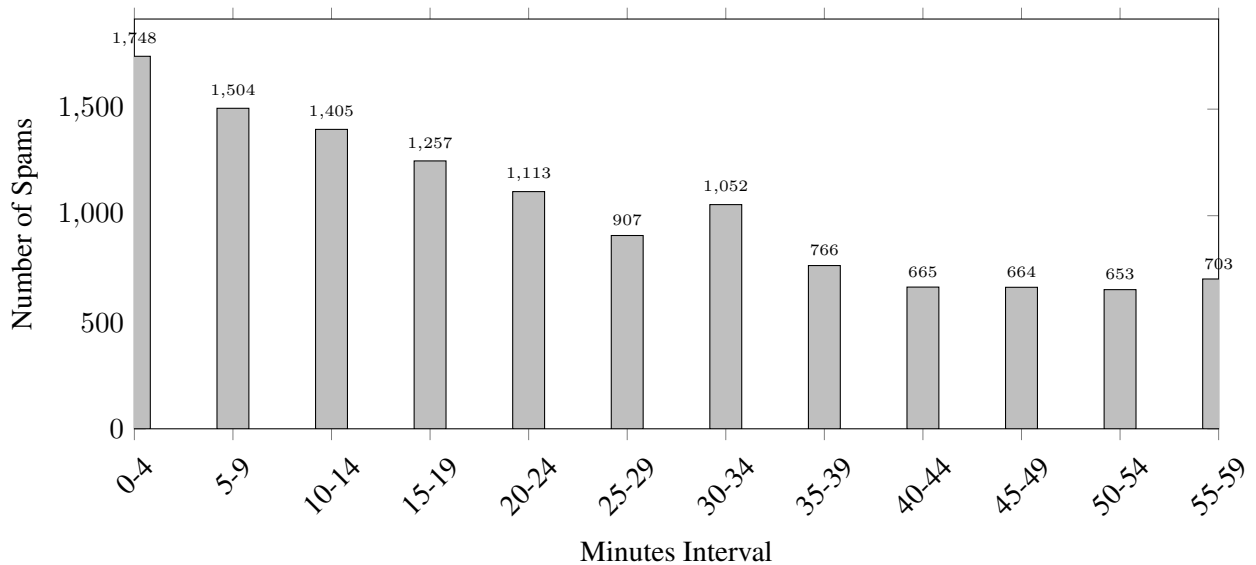


Figure 4.5: 5 Minute Interval Spam Capture

The sample mean of the 5-minute interval spam counts figure 4.5 is approximately 1036.42, standard deviation of 374.20 with a margin of error (ME) of approximately 278.27. The lower and upper bounds of the 99% confidence interval for figure 4.5 are 758.15 and 1314.68, respectively.

### 4.3.5 Day-of-the-week Spam Analysis

To further expand our spam filtering model recommendations, we present a day-of-the-week analysis of spam volume received over a 12-month period. Analyzing spam distribution by day of the week can provide insights into spam patterns, which may be useful for detecting anomalies, predicting spam trends, and refining security measures.

The total number of recorded dates was 3,146, as shown in Table 4.3. The data in Figure 4.6 reveals a distinct pattern, with a peak in spam volume on Thursdays (547), followed by a steady decline on Friday (489), Saturday (381), and Sunday (358). Conversely, spam volume rises again from Monday (422) through Wednesday (494), culminating in the Thursday peak.

Additionally, we observed that weekday spam volume was consistently higher than weekend spam volume. This suggests that spammers may be timing their campaigns to exploit user behavior, potentially aligning spam waves with business email activity, such as midweek corporate communications.

<b>Day of the Week</b>	<b>Count</b>
Thursday	547
Wednesday	494
Friday	489
Tuesday	455
Monday	422
Saturday	381
Sunday	358
<b>Total</b>	<b>3,146</b>

Table 4.3: Distribution of Spams by Day of the Week

These spam timeline findings highlight periods when spammers are most active. It also showed that a lot of spams are scheduled, not sent in real-time. These insights can be utilized to enhance spam detection and mitigation in several ways:

- The timeline data can be used to train models that predict peak spam hours, enabling the preemptive blocking of suspicious traffic during these periods.
- Automate tools that can be deployed to aggressively flag or quarantine spam during high-risk hours, minutes and day-of-the-week, reducing reliance on manual intervention.
- Investigate sender addresses, subject lines, and email content during peak periods, which can help identify recurring spam campaigns and patterns.

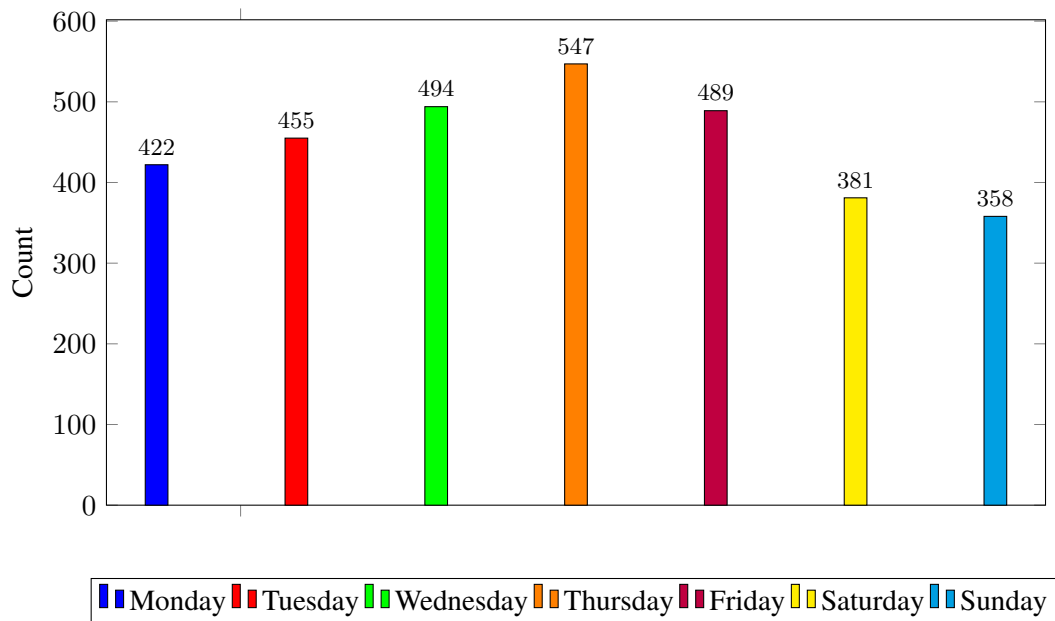


Figure 4.6: Distribution of Dates by Day of the Week

- During peak hours, minutes, and days, stricter spam filtering rules can be enforced by increasing sensitivity to key indicators such as suspicious keywords, unknown senders, or known spam sources, effectively reducing spam influx during high-traffic periods. Analyzing spam trends over time suggests that prevention strategies should be dynamic, addressing both peak and low periods. While proactive measures should be strengthened during surges, continuous monitoring and adaptive defenses are essential during drop periods to prevent resurgence. Additionally, it is important to recognize that email service providers may already filter out some spam through their own mechanisms. Therefore, a layered approach—integrating honeypot insights with existing spam filtering systems—offers a more robust defense against spam.

## 4.4 Nature of the Spam

To understand the nature of the spam received, We classified the audited spams using keywords and contexts in the body of the spam message. The classification description are as follows:

- (1) offers/promotions/Advertisement: spams seen to contain or connote keywords including offer, subscribe, advertise, best price, free/limited time, exclusive, congratulations, guaranteed, discount, price, urgent bonus, limited stock, news, product recommendation (courses, webinar promotion), product endorsement. (educational products/services, like software books, study materials).

- (2) scams: the keywords include urgent, claim, winner, account, verify, payment transfer, confidential, risk-free, limited time, fake romance, invitation to chatroom.
- (3) phishing: unsolicited survey, feedback, opinion, poll, survey, other phishing words are update info/account, password update, update change required, confirm security, suspended, action required, usual activity, fake scholarships offer, bogus academic journals (invitations to submit paper to fake journals/conferences),

In Figure 4.7, we present the classification of spam based on its nature. Our analysis revealed that 76% of the spam received by our honeypot consisted of advertisements and offers, followed by scams at 23.7% and phishing at 0.30%.

The high proportion of advert/offer spam (76%) suggests that the majority of spam is promotional in nature, aiming to entice recipients to engage with products, services, or websites. The significant presence of scam-related spam (23.7%) indicates a considerable risk of fraudulent activities, such as fake romance scams and chatroom invitations, commonly found on platforms like Amolatina.com, Travel-mates.com, Hotti.com, etc.

Additionally, the minimal presence of phishing emails (0.30%) suggests that phishing campaigns may be more targeted and less frequent, yet still pose a serious security threat due to their potential impact.

## 4.5 Demographic Analysis

The research experiment covered 64 age groups (13 to 75 years) and analyzed 370 websites, categorized into 12 communities: transport, education, accommodation, e-commerce, sports, news, recreation, health, jobs, forums, social networks, and family. Each community had a minimum of 8 websites, resulting in an uneven distribution of websites among communities. However, this uneven distribution did not impact honeypot activity, as each honeypot was consistently assigned unique five websites with only one overlap between any two profiles, irrespective of the community.

To address the challenge of unequal exposure - where communities with more websites had a higher likelihood of receiving spam - we ensured that our analysis was not biased toward communities with larger website counts. Additionally, we accounted for sampling bias, where a community with fewer websites might show a higher percentage of spam due to website characteristics (e.g., security practices, content type), rather than the age or gender of users.

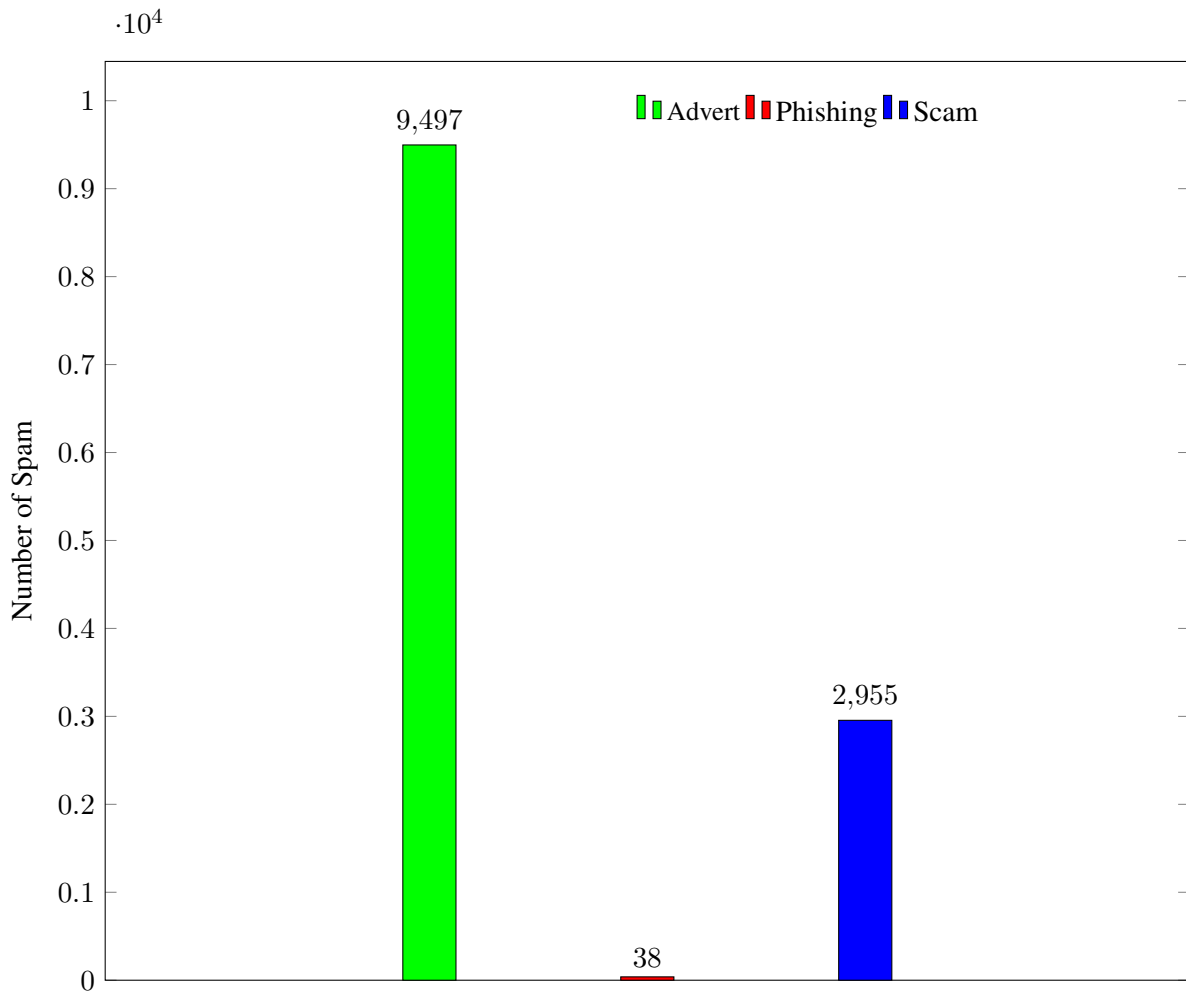


Figure 4.7: Distribution of Spam by Classification

To ensure a fair comparison, we analyzed spam rates per user rather than per community, providing a clearer reflection of which age and gender groups were more attractive to spammers, independent of website distribution. The honeypots were grouped by gender (male and female) and age (in 5-year intervals) to determine which demographic groups received the most spam, as shown in table 4.4.

Table 4.4 illustrates the total volume of spam received by gender, showing that men received the highest number of spam emails, with a total of over 9,352 emails, compared to women, who received 3,138. Further analysis, as shown in Figure 4.8, reveals that men in the 48-52 (M: 2,098, F: 62) and 53-57 (M: 2,696, F: 668) age groups received the highest volumes of spam. However, due to the limited sample size, it is important to note that these findings should not be generalized.

Based on the data we collected, the results suggest that middle-aged individuals, particularly men in their late 40s and 50s, are prime targets for spam. This could be attributed to their financial stability, which may make them attractive targets for spammers. Additionally, they may be less familiar with technology compared to younger generations, which could increase their vulnerability to spam-based

threats. For a detailed breakdown of spam counts per user, refer to Appendix A.

Younger users received relatively low spam counts. The 13-17 age group had the lowest spam rates (M: 46, F: 247). The spams were adverts from legitimate websites that they had accounts on. The low spam rate is likely because younger users are less likely to be targeted for financial fraud since they have fewer online accounts tied to commercial services and age requirements constraints for some websites. The 68-72 age group (M: 962, F: 235) and 73-75 (M: 453, F: 85) still showed significant spam activity, though lower than middle-aged users. Notably Spammers may target elderly users for scams like fake romance etc. as in the case of a honeypot - John Grenier, Jeremy Lake etc.

Each age-group - gender set considered in demographic analysis was made up of 5-males and 5-females that fall within each category, leading to minimum of 10 honeypots in each age-group- gender set. Although this set is considered small, these statistical findings hold significant relevance in multiple dimensions of cybersecurity research, user awareness. First, the observation that some honeypot accounts did not receive any spam while others received disproportionately high volumes suggests that spam campaigns are highly selective rather than random. This reinforces the argument that targeted data leaks, non-enforcement of privacy policy and forced consent practices, rather than mere user behavior, play a role in spam distribution. If all accounts were equally vulnerable, a more uniform spam distribution would be expected. Instead, the disparity in spam volume hints at underlying patterns in how spammers acquire and prioritize their targets.

Second, the demographic disparity in spam receipt—particularly the targeting of middle-aged men—has practical implications for cybersecurity education and fraud prevention. Financially stable individuals may be at greater risk of phishing attempts, investment scams, and other forms of financial fraud. The lower spam rates among younger users suggest that spammers optimize their efforts based on perceived financial viability rather than indiscriminately targeting all age groups. This insight could inform regulatory discussions on digital privacy protections, particularly concerning how data brokers and compromised databases may contribute to spam targeting.

The continued spam exposure of elderly individuals highlights ongoing risks related to romance scams and other fraudulent schemes, corroborating findings of [50, 51]. The presence of such patterns in honeypot data underscores the need for advanced spam detection models that integrate demographic risk factors rather than relying solely on traditional filtering techniques.

Table 4.4: Spam Counts by Age Group and Gender

Age Group	Female	Male
13-17	247	145
18-22	182	367
23-27	289	128
28-32	145	596
33-37	109	74
38-42	213	226
43-47	138	292
48-52	106	2098
53-57	668	2696
58-62	587	132
63-67	47	497
68-72	235	962
73-75	172	1139
Total	3138	9352

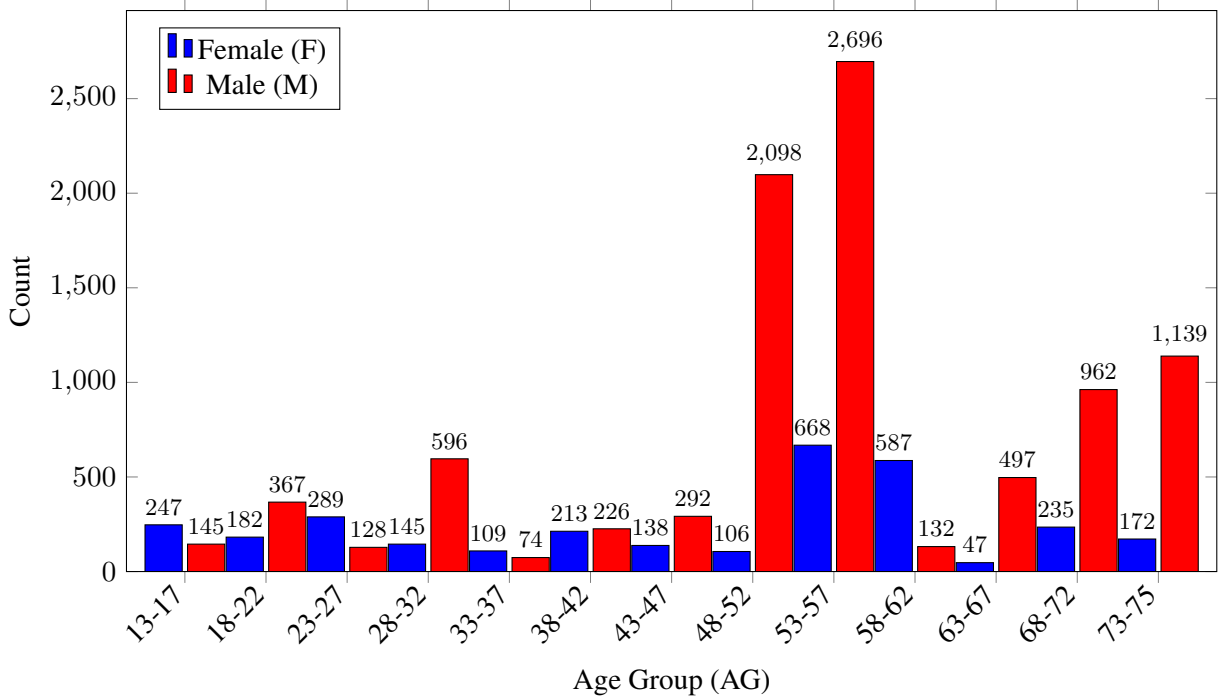


Figure 4.8: Demography of Spam

## 4.6 Addressing the Base Rate Fallacy in Our Study

In the context of spam detection and analysis, the base rate fallacy occurs when the probability of an event is misjudged due to ignoring the base rate (or prior probability) of that event. Specifically, when analyzing spam messages, there is a tendency to overestimate the likelihood of a message being spam based on features such as suspicious keywords or unknown senders, without considering the overall base rate of spam messages in the given dataset.

In our study, we made deliberate efforts to control for the base rate fallacy, ensuring that our spam detection process remained robust and unbiased. Our experimental design adhered to five key rules that guided interactions with email service providers and websites, the most significant being that our honeypots could not accept any subscription offers voluntarily. This controlled environment was critical in isolating spam messages and ensuring the accuracy of our results. Below are the specific measures taken to mitigate the base rate fallacy:

**Controlled Interaction with Websites:** Our honeypots were strictly restricted to interactions within the predefined list of websites. They were not allowed to click on any advertisements or promotions, which helped prevent the introduction of unintended variables that could skew spam generation based on user behavior or targeted marketing.

**Limiting Email Subscription Engagement:** The honeypots did not voluntarily accept any subscription offers from websites. This ensured that spam messages originated only from sites that were actively involved in our research, rather than from unsolicited or unrelated sources that could distort the spam patterns we were trying to measure.

By implementing these strict rules and ensuring a controlled environment, our study effectively mitigated the risk of base rate fallacy. Our holistic approach, which balanced individual message analysis with a broader look at spam distribution, has helped ensure that our conclusions are grounded in a comprehensive, unbiased view of spam activities, providing a reliable basis for understanding and improving spam detection mechanisms.

## 4.7 SMS Spam Analysis

The experiment was conducted using 20 SIM cards purchased from T-Mobile, USA. Over the course of the study, six of these SIM cards (30%) received spam messages, totaling 11 spam messages. Some spam message content suggested a connection to a previous owner of the SIM card. Additionally, some of the spam messages contained names that we believe may be linked to previous owners but also included keywords commonly associated with phishing schemes.

Upon categorizing the spam messages received via SMS, we identified the following distribution:

20% Subscription Spam (Advertisements & Offers) 10% Scam Messages 70% Phishing spams One of the key challenges observed was the lack of de-association from websites where phone numbers had been previously registered. Because telecoms providers do not cleanly de-associate phone numbers from user accounts before reissuing them and the spam messages did not contain our unique name variants,

we were unable to attribute any spam messages to a specific website. This issue was further highlighted when we attempted to create an account on UberEats with a reassigned number, only to find that the number was still linked to a previous user's account, rendering it unusable for new registration.

To solve the phone number to website association problem, we recommend a Nonce-Based De-Association - where websites that use phone numbers for authentication (e.g., login verification, OTP, and two-factor authentication) face a critical security risk when telecoms reassign numbers to new users. Since these numbers often remain linked to previous accounts, new owners may receive spam, phishing attempts, or unauthorized login prompts meant for the previous user. This problem arises because most websites do not automatically de-associate phone numbers from accounts after telecom reassignment.

A nonce-based verification system can mitigate this issue by ensuring phone number associations remain up to date. In this approach, a one-time-use, cryptographically unique nonce is generated during verification and tied to the telecom's verification process, preventing reuse or replay attacks. When a telecom reassigns a number, it triggers a de-association request, which is broadcasted to all websites previously linked to that number. Upon receiving this request, websites are required to unbind the number from the previous account and force re-verification for any new user.

This system improves security by ensuring that old accounts cannot be accessed using a reassigned number, thereby preventing unauthorized access, spam, and phishing risks. The implementation follows a structured four-step process:

- **User Registration:** Websites generate a nonce, which is cryptographically signed by the telecom, confirming number ownership.
- **Number Reassignment:** When a telecom reassigns a number, it broadcasts a de-association request.
- **Website Update:** Websites receiving the de-association request unlink the phone number from old accounts and require re-verification.
- **New User Registration:** The new owner registers without inheriting old account associations or receiving unwanted spam.

Additionally, a webhook system allows websites to subscribe to telecom de-association events, ensuring real-time updates. While the solution requires telecom cooperation and website compliance, its benefits in preventing spam, improving user privacy, and reducing account hijacking risks, make it a scalable and effective strategy for managing phone number reassignments securely.

## 4.8 Summary

This study found that untargeted spam (7061, 56.5%) from legitimate websites significantly outpaced targeted spam (2345, 18.8%), suggesting that websites with existing user accounts generate more untargeted spam. The analysis of top spam senders identified domains like `childrensplace.com` and `poshmark.com` as the largest contributors, challenging traditional spam detection systems that rely on domain reputation. The study also revealed that spam emails containing personal information were more likely to engage users, with targeted leak spam (23.2%) outnumbering untargeted leak spam (1.39%). Spam activity showed irregular growth, peaking in September and November, possibly due to specific events or campaigns. Hourly analysis indicated peaks in spam activity between 1–2pm, 1–2am, and 10am, aligned with user behavior and regional time zones. Monthly trends showed a spike in spam volume from March to May, followed by a decline during the summer, correlating with efforts to unsubscribe from certain websites. The study also included demographic analysis, revealing that men, particularly those in the 48-57 age range, received the most spam. Younger users (13-17) had the lowest spam rates due to fewer range of online accounts, while older users (68-75) also experienced significant spam, due to scams targeting elderly individuals. SMS spam analysis showed that 30% of SIM cards received spam, with phishing being the most prevalent type (70%). The study highlighted the lack of de-association between reassigned phone numbers and websites, recommending a nonce-based verification system to mitigate risks associated with reissued phone numbers, improving user privacy and reducing unauthorized access. The insights from these findings emphasize the need for more sophisticated spam filtering systems and proactive measures to prevent spam both at peak and drop periods, while acknowledging that some spam mails may have been filtered out by email service providers, who also are providing spam filtering.

## Chapter 5

# REGULATORY COMPLIANCE AND DATA LEAK FORENSICS

In this chapter, we provide a high-level summary of the Canadian Anti-Spam Law (CASL), given that this research was conducted in Canada. We reference CASL to evaluate the privacy policy statements of non-compliant websites against its implicit and explicit consent requirements.

Additionally, we explore the characteristics of data leaks, identifying suspect websites connected to the spam headers received by our honeypots. We assess the exposure impact, investigate potential malicious distribution or data harvesting tendencies associated with our legitimate websites, and present a comparative analysis of our methods against existing data leak detection and fingerprinting techniques. The content of the data leaks has been accepted for publication as a paper in CODASPY 2025 (see Appendix A).

### 5.1 Canadian Anti-Spam Law

The Canadian Anti-Spam Law (CASL) [52] is a regulation designed to manage the transmission of electronic messages with commercial intent. It applies to individuals and businesses sending Commercial Electronic Messages (CEMs) to or from Canadian electronic addresses.

A CEM includes communications like emails, text messages, and social media direct messages that promote participation in commercial activities. CASL's scope covers all Canadian electronic addresses, including email addresses, mobile phone numbers, and social media profiles, ensuring that any commercial message targeting Canadian users adheres to its regulations.

CASL governs three key scenarios. First, inbound messages to Canada, where businesses or individuals located outside Canada send marketing communications to Canadian recipients. For instance, a foreign software company promoting subscription upgrades to Canadian users must comply with CASL. Second, outbound messages from Canada, where Canadian entities target international recipients.

For example, a Canadian e-commerce website advertising sales to U.S. customers must adhere to CASL while also considering anti-spam laws in the recipient's jurisdiction. Finally, CASL applies to messages within Canada, such as a Canadian gym texting its members about promotional offers. This broad applicability ensures that CASL governs both domestic and international communications involving Canadian addresses.

Consent is a fundamental requirement under CASL, with senders needing either explicit or implied consent to contact recipients. Explicit consent involves a clear, affirmative agreement from the recipient, often obtained through opt-in forms or subscriptions. Implied consent applies in cases such as existing business relationships (e.g., a recent purchase) or inquiries initiated by the recipient. These messages must include a clear description of the message's purpose and identify any third parties involved in sending the message.

CEMs must include a functional and easy-to-use unsubscribe option that allows recipients to withdraw consent. Senders are required to process unsubscribe requests within 10 business days of receipt. CASL also includes exemptions for certain types of communications, such as personal or family messages, factual updates like product recalls, and responses to customer inquiries. These exemptions balance consumer protection with practical flexibility for specific scenarios. CEMs must clearly identify the sender, including their name, contact information, and business details.

Furthermore, certain activities are prohibited—using false or misleading information in the header, subject line, or body of a CEM, installing computer programs or software on a recipient's device without their explicit consent, and collecting or harvesting electronic addresses without permission (address harvesting). Violating CASL can lead to severe consequences, including fines of up to \$1 million for individuals and \$10 million for businesses per violation. Directors and officers can also face personal liability for non-compliance.

Regulatory oversight is provided by the Canadian Radio-television and Telecommunications Commission (CRTC) for compliance, the Competition Bureau for false or misleading claims, and the Office of the Privacy Commissioner (OPC) for privacy violations. Together, these agencies ensure that CASL's principles are upheld across all applicable communications.

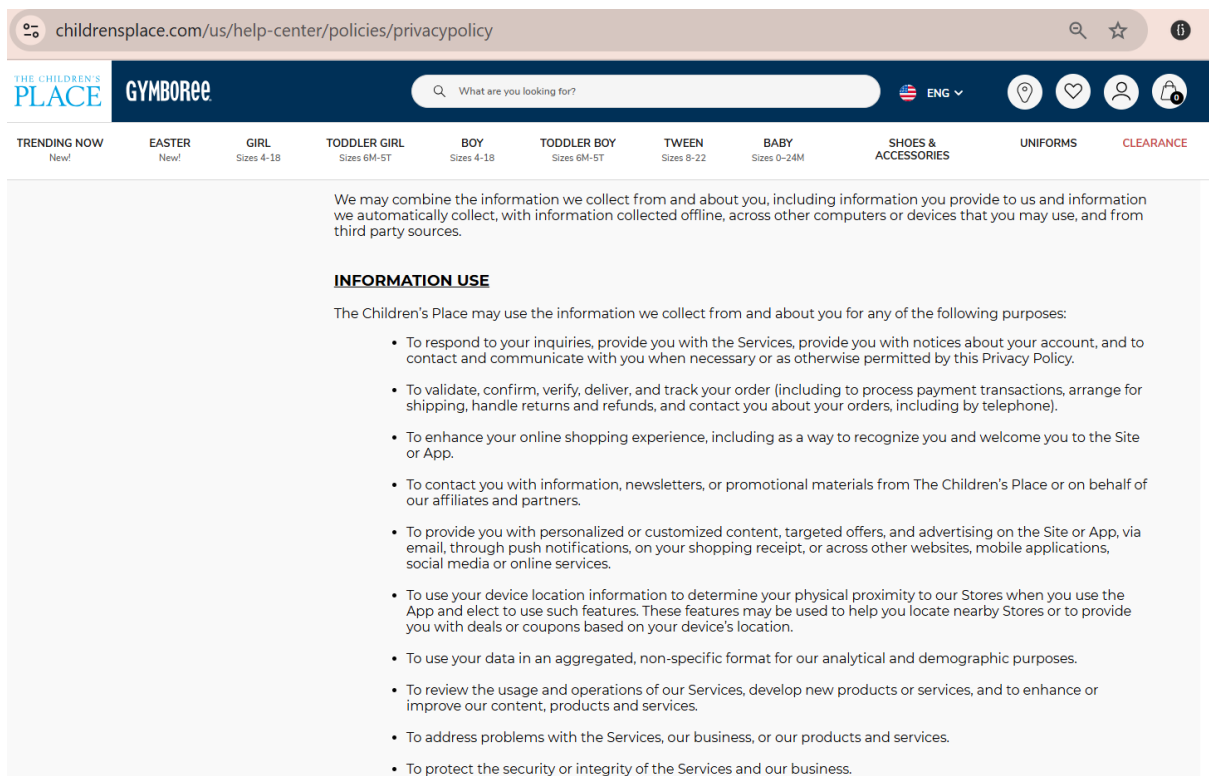


Figure 5.1: Exploitation of Implicit Consent

## 5.2 Consent Exploitation

**Findings on Website Non-Compliance with Canadian Anti-Spam Law:** Our findings revealed that websites associated with our honeypots accounted for 75.3% of the spam received, underscoring their role in forced or implied consent subscription practices. Notably, websites such as childrensplace.com, houzz.com, and toyrus.ca significantly contributed to the spam volume received by our honeypots. These websites exploit clauses in the Canadian Antispam Law (CASL) to justify their subscription practices.

**Exploitation of Implicit Consent Clauses:** Our analysis highlights the actions of non-compliant websites identified in our research experiment, examining the specific legal loopholes they exploit to shield themselves from liability. A review of privacy policies revealed that many defaulting websites rely on implicit consent clauses, implying user consent upon account sign-up. Some examples include rebag.com, coursera.com, jcpenny.com, lingoda.com, and zoosk.com, childrensplace.com etc. Figure 5.1 shows a sample privacy statement with implied status on how the data collected is used.

**Lack of Enforcement of Privacy Policies:** Another concerning trend observed is the lack of enforcement of website privacy statements. For instance, allrecipes.com, tripadvisor.com, ticketmaster.com, thestar.com, eharmony.com, thrivemarket.com etc., explicitly states that opt-in consent is required for subscription emails. However, our honeypots never subscribed at creation or during interactions, yet still received marketing emails - demonstrating a failure to enforce their own policies. Figure 5.2 show sample privacy statement with observed lack of policy enforcement.

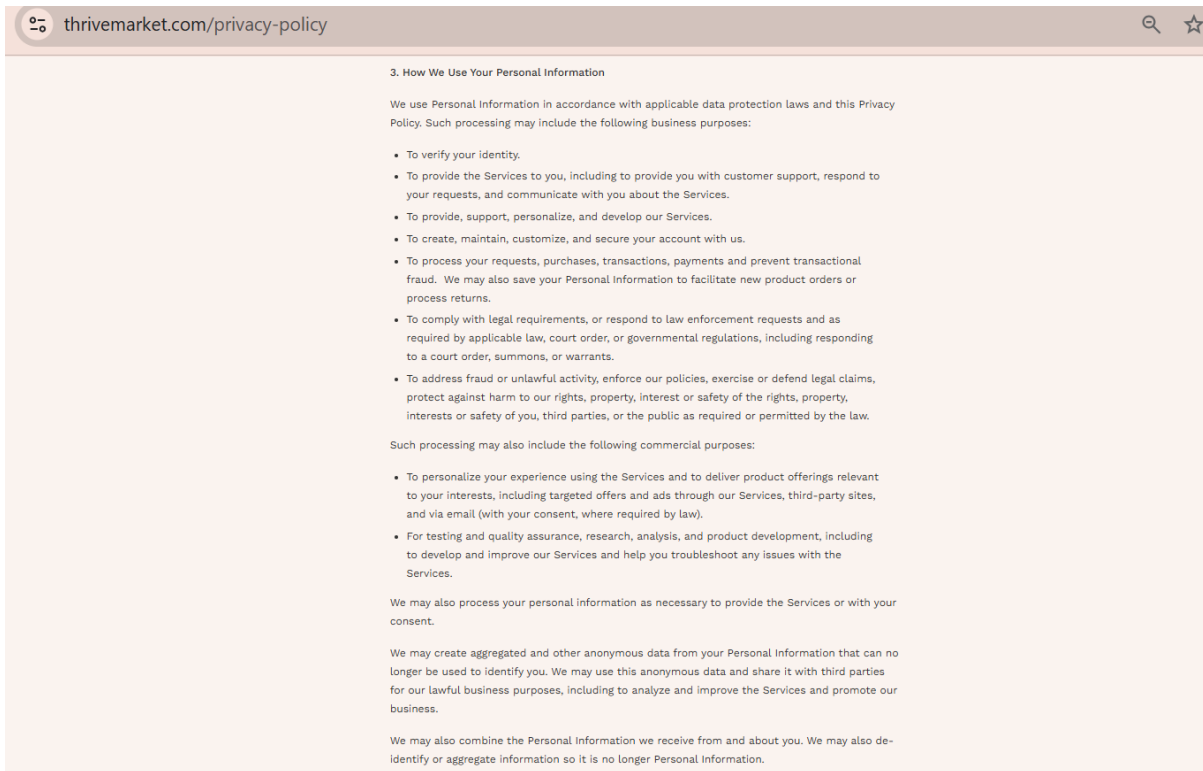


Figure 5.2: Lack of Privacy policy Enforcement

Also, some websites integrate explicit subscription consent into their privacy policies, effectively obliging users to accept marketing emails upon sign-up. Examples include koio.co and Frank and Oak etc. This practice leaves users with no genuine option to opt out of mailing lists, ultimately undermining user autonomy and informed consent. Figure 5.3 show sample privacy statement that integrate explicit subscription consent into their privacy policies.

**No Privacy Policy:** In addition to the identified non-compliance patterns, we encountered a notable case where a website had no privacy policy statement available. This raises concerns about transparency and compliance with data protection regulations, as users cannot review how their data is collected, stored, or shared.

Also, we observed an instance where one of the websites (takelessons.com) our honeypots visited ceased operations during the course of our experiment. This prevented further analysis of its practices,

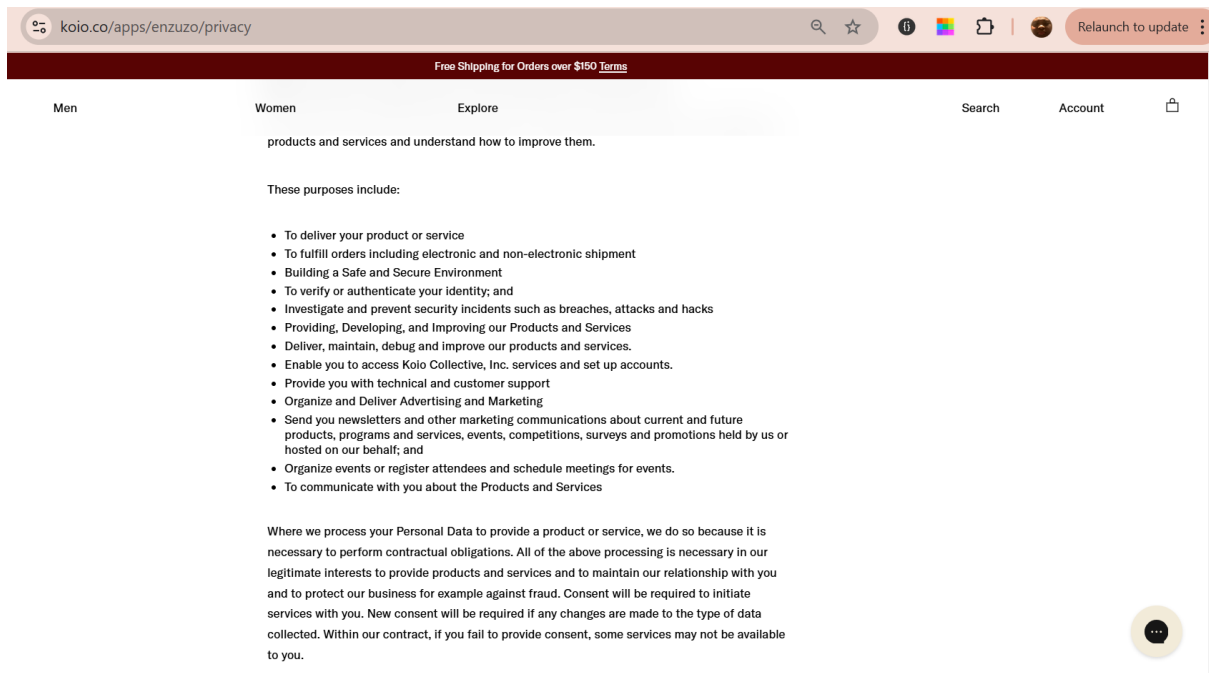


Figure 5.3: Sample explicit implied consent

leaving its prior data handling policies uncertain. For a comprehensive list of defaulting websites and their consent practices based on privacy policy reviews. (Refer to Appendix A for the repository url).

The lack of clear consent mechanisms, combined with forced subscription practices, seen with websites', the lack of privacy statement enforcement and third party share practices within the privacy policy term agreement, which is made compulsory for account holders to accept, suggests negligent data handling practices. Such negligence not only compromises user privacy, but also increases the likelihood of data exposure.

The result underscore the need for a review of the law to reduce spam receipt, without reinstating the currently suspended Private Right of Action. This would help enhance the effectiveness of the law in protecting users from spam while maintaining a balanced approach.

### 5.3 Our Recommendation

To reduce spams broadcast and improve privacy protection, several key recommendations should be implemented in the terms of agreement for online services:

**Limiting the Scope of Agreements to Core Services Only:** Terms of agreements should focus exclusively on the core services that users sign up for, ensuring that non-essential services (e.g., third-party marketing, data-sharing) are not bundled into the terms. Users should only agree to what is necessary for them to use the platform's primary features, and any additional services should require separate,

explicit consent. This approach helps reduce exposure to unwanted spam and minimizes unnecessary data collection.

**Users Should Be Allowed to Opt-Out of Third-Party Agreements:** Users should have the ability to easily opt-out of third-party services, such as marketing campaigns, data-sharing with external partners, or other non-essential features, without losing access to core functionalities. Opting out should not result in a loss of functionality in the main services provided by the website. This is crucial for maintaining user trust, as it allows them to retain control over their data while still accessing the services they originally intended to use. The ability to choose whether or not their data is shared or used for targeted ads should not impact their access to basic services like account management, purchases, or communication with customer support.

**Default Should Be Not Opt-In:** The default setting for any third-party or non-essential services should be "opt-out," meaning that users should not automatically be enrolled in data-sharing or marketing programs when they sign up for a service. Websites should refrain from pre-ticking boxes or using "opt-out" options that automatically include users in data-sharing or marketing initiatives. Instead, users should be required to actively opt-in to any additional services or data sharing beyond the core functionality of the website. This reduces the likelihood of users unintentionally agreeing to terms that they do not fully understand or consent to.

**Penalty for Lack of Enforcement of Privacy Statement Content:** As observed in the research, some websites failed to enforce the consent mechanisms outlined in their privacy policies, particularly when it came to marketing emails, despite their privacy statements explicitly stating otherwise. This discrepancy between stated privacy policies and actual practices highlights a significant gap in enforcing user privacy rights. To address this issue and serve as a deterrent, stricter penalties should be implemented for websites that fail to adhere to the privacy commitments they make. These penalties could include higher fines, restrictions on service access, or mandatory public disclosures of non-compliance.

The penalties should be structured to ensure that the cost of non-compliance outweighs any potential benefits from violating user privacy. Websites should be held accountable for failing to obtain explicit consent for marketing communications or improperly sharing user data with third parties. Furthermore, businesses that repeatedly violate privacy statements or engage in deceptive practices should face increasing penalties over time. These measures would not only protect users' rights but also encourage websites to take privacy seriously and actively enforce the terms of their privacy policies. This approach ensures that privacy policies are more than just words on a page but are enforced in a way that truly protects users from unwanted spam and potential data misuse.

By implementing these recommendations, websites can create a more user-friendly and privacy-conscious environment. Users will feel more in control of their data, and businesses can demonstrate a commitment to privacy and user rights, potentially leading to increased trust and customer loyalty.

## **5.4 Analysis Engine**

In our data leak forensics, we utilized the analysis engine to link spam senders to suspect legitimate websites, where our honeypots have accounts. The analysis engine employed both account-based and mapping-based analysis to identify the legitimate websites responsible for data leaks.

## **5.5 Account-based Analysis:**

For account-based analysis, spam leaks were identified based on the presence of unique honeypot name variants in the body of the messages, directly tying the spam to a specific associated website. Spammers created personalized emails in order to gain the attention of honeypots. Our usage of unique honeypot names in each account allowed us to accurately identify the honeypot account for each targeted spam, and then accurately identify the website who leaked the account information. Our account-based analysis followed these steps:

- (1) Identify targeted spam emails that were not from websites where our honeypots have accounts.
- (2) Extract the addressed names from the spam emails and identify their corresponding accounts. This step is possible since we used unique names for each account.
- (3) Return the websites associated with those accounts as data leak sources.

## **5.6 Mapping-based Analysis:**

Mapping-based analysis was used for untargeted spams received due to data leak. These spams originated from domains that our honeypots did not have accounts. However, since the spams were untargeted, it is not straight-forward which account they targeted, so account-based analysis did not apply. In this case, we used our mapping-based analysis to find the data leaker following the steps below:

- (1) Identify untargeted spam emails that were not from websites where our honeypots have accounts.

- (2) Group the spam emails by sender emails.
- (3) For each group, identify the affected honeypot profiles by the recipient emails, find the websites set they were mapped to, and look for intersections of those sets. Return the intersection websites as the leaking sources.

## 5.7 Leak Analysis Results:

As mentioned earlier, we recorded 2,910 targeted leak spams, 174 untargeted leak spams, 7,610 untargeted service provider spams, and 2,345 targeted service provider spams. The leak spams, both targeted and untargeted, were associated with 16 legitimate websites.

**Targeted leak spams:** During the the analysis period, 2,910 targeted leak spams (23.2% of the total spam volume) were identified. These spams were analyzed using the account-based analysis. Results show that those spams were the results of leaks from 7 legitimate websites (Table 5.1). More detailed information on the accounts and spammers can be found in the Appendix (Table .2).

Table 5.1: Leaking Source Websites and Leak Recipients

Leaker ID	Website	Leak Recipients
source1	cheryls.com	A
source2	datemyage.com	B, C, D, E, F, G
source3	garageclothing.com	H
source4	chinalove.com	B, I, J, K, L, D, F
source5	menshealth.com	N
source6	yemeksepeti.com	O
source7	simplyHired.com	P
source8	thestar.com	Q
source9	dating.com	I, J, K, L, D, E, R, S
source10	zoosk.com	T
source11	academicearth.org	U
source12	dominos.com	V
source13	blue mountain resort	W1, W2, W3, W4, W5, W6
source14	zdnnet.com	X
source15	istockphoto.com	Y
source16	lowes.com	M

medallia.com (A), dating.com (B), pinadate.com (C), eurodate.com (D), asian-date.com (E), getonce.com (F), anastasiadate.com (G), dynamistyle.com (H), datemyage.com (I), amolatina.com (J), amaldate.com (K), travelmates.com (L), decipherinc.com (M), hearstmags.com (N), hungryroot.com (O), indeed.com (P), torstar.ca (Q), zendate.com (R), hotti.com (S), zoosk.com (T), pitt.edu (U), factormeals.ca (V), emailmeform.com (W1), hmkw.de (W2), beuzpro.com (W3), consulting.de (W4), lesartsoseurs.org (W5), pothys.com (W6), redventures.com (X)

**Untargeted leak spams:** 174 untargeted leak spams (1.39% of the total spam volume) were recorded. We applied Mapping-based Analysis on this set of spams and successfully identified 9 leaking websites. They are yemeksepti.com, dominos.com, chinalove.com, datemyage.com, dating.com cheryls.com, thestar.com, menhealth.com, lowes.com,. Note that there is an overlap between this results and the results we obtained from account-based analysis, demonstrating different analysis methods can lead to the same findings.

## 5.8 Extent of Exposure

In this research, the extent of exposure of honeypot data within the spam ecosystem was evaluated based on the number of recipients of the leaked data, who sent targeted or untargeted spam messages to our honeypots. The email addresses of these recipients were crawled from the email messages. Our analysis, which linked the leak source to the recipient email addresses retrieved from each affected honeypot's account, revealed that honeypots that experienced data leaks received spam emails from 1 to 8 unique sender email addresses. The mapping from leaking websites to their corresponding spammers are presented in Table 5.1.

## 5.9 Discussion

It's worth noting that 1455 (50%) leak spams were received from dating.com, even though none of the honeypots had accounts interactions on this website. Also, 1320 leak spams (45%) were sent by datemyage.com (I), zendate.com (R), almaldate.com (K), travelmates.com (L), usadatzing.com (T), asian-date.com (E), anastasiadate.com (G), eurodate.com (D), and hotti.com (S). Of the remaining 140 spams (5%), they were sent by indeed.com (P) and hungryroots.com (O), factormeals.ca (V), emailmeform.com (W1), hmkw.de (W2), beuzpro.com (W3), consulting.de (W4), lesartoseurs.org (W5), pothys.com (W6), redventures.com (X) etc. Noting that our honeypots had no accounts on these websites.

In some cases, we examined third-party partner information associated with legitimate websites where our honeypots had registered accounts. Our focus was on isolated incidents where only a single honeypot received spam while its counterpart did not—an indication of a potential data leak (more details in the appendix A). This analysis aimed to evaluate whether third-party partners of these websites were involved in spam distribution. The following factors may explain these occurrences, based on patterns observed during the analysis:

- Spammers might have used data analytics to segment potential targets based on user profiles created from leaked information. If the honeypot’s unique name variant aligns with specific demographic or behavioral patterns, it could be specifically have been targeted.
- Websites may have inadvertently or intentionally exposed only a subset of user data. The unique name variant of the honeypot may have been part of a data set leaked due to lax security practices or a specific breach incident.
- The timing of the spammer’s access to leaked data could have aligned with specific honeypot entries and not others.
- Spammers may have selectively used portions of the acquired data to test response rates or effectiveness of their campaigns, which could leave the counterpart honeypot untouched.
- Some spammers might be focusing on geographic, cultural, or niche interests aligned with the variant names or domains.

### **5.9.1 Improper Use of Information**

The set of leaked spam received by our honeypot, which did not include the honeypot variant name or attributes in the spam body, was analyzed using the mapping analysis method. We reviewed the third-party partners and the business services offered by the legitimate websites associated with the affected honeypots. Using the principle of shared business models within our mapping-based analysis, we identified cases where brand affiliations or common ownership under the same parent entity were evident. This demonstrated the effectiveness of our approach, even with prior knowledge constraints.

While these cases may not strictly qualify as unintentional data leaks, they represent intentional misuse of user information from a privacy perspective. Our findings highlight non-compliance by third-party partners of certain websites. For example, Indeed.com disregarded our non-consent on SimplyHired.com, iStockPhoto.com, and GettyImages.com, leading to unauthorized email campaigns using honeypot data. Notably, the primary websites where our honeypots were registered—such as SimplyHired.com—maintained a strict no-mailing policy. However, their third-party partners violated these privacy safeguards, underscoring the risks associated with third-party data sharing and its contribution to spam proliferation.

The result revealed the following key findings:

- (1) some websites have brand partnerships with whom data was shared. These third party partners

did not uphold the no-mailing list subscription. We saw cases such as menhealth.com (source5) and hearstmag.com (N), thestar.com (source8) and torstar.com (Q), istockphoto.com (source15) and gettyimages.com (Y), zdnet.com (source14) and redventures.com (X).

- (2) Certain websites associated with targeted leak spams also sent untargeted spam emails. By analyzing these emails, we were able to trace the sources of the leaks using account-based analysis, as observed with eurodate.com, getonce.com, and datemyage.com

Notably, some honeypots received spam from as many as eight distinct spammer email addresses, which were traced back to legitimate website where the affected honeypots had registered accounts. This high number of distinct spammer email addresses was particularly observed with datemyage.com (source2) and dating.com (source9). Honeypots linked to chinalove.com (source4) also showed significant exposure, with seven distinct spammer email addresses identified. Honeypots linked to blue mountain restart (source13) received as much as six distinct spammer email addresses (see appendix [A](#) for more information).

## 5.10 Comparative Analysis Approach for Data Leak Detection

A comparative analysis evaluating the effectiveness of spam-based leak detection method against traditional data leak detection approaches was done. Instead of detecting spam itself, our approach uses spam as an indicator of data leaks, tracing leaks to their originating websites and assessing compliance violations. To benchmark its performance, we relied on the following key metrics:

**Leak Detection Rate (LDR):** This measures the percentage of verified leaks successfully identified using spam indicators.

**Source Attribution Accuracy:** This evaluates how accurately the model traces leaks back to their originating sources.

**Compliance Violation Rate:** This quantifies how many websites violate their privacy policies by leaking user data.

**Shared Business Model Leak Rate:** This Identifies suspected leaks where the spam leak source operates within the same business model as the legitimate website.

**Comparison with Other Data Leak Detection Methods** To evaluate our approach against existing detection techniques such as network-based intrusion detection systems (IDS), anomaly detection models, and forensic log analysis, we compared:

**Leak detection efficiency:** Traditional methods rely on network traffic analysis, system logs, or machine learning models to detect breaches. Our approach, however, leverages real-world spam as forensic evidence, which may detect leaks that other models overlook.

**Source attribution performance:** Unlike conventional methods that trace data ex-filtration through network logs, our method directly identifies the leak source based on spam characteristics (account-based) and recipient mapping.

**False positive rates:** Traditional methods may flag anomalous behaviors that do not necessarily indicate a data leak, while our method reduces false positives by relying on unique identifiers within spam content.

**Time to detection:** IDS-based models often detect leaks in real-time, whereas our approach detects leaks post-compromise (receipt of spam), but offers retrospective insights into compliance violations and trace to source of leaks.

**Regulatory compliance analysis:** While other approaches primarily focus on technical leak detection, our method provides an additional layer by assessing privacy policy violations.

Through this comparative study, we aim to determine whether spam-based leak detection provides complementary advantages to traditional data breach detection, particularly in regulatory compliance assessment and forensic analysis of data misuse.

Below are the formulas and calculated values for our spam-based data leak detection method.

## 1. Leak Detection Rate (LDR)

$$\text{LDR} = \left( \frac{\text{Total Traced Leaks}}{\text{Total Verified Leaks}} \right) \times 100 \quad (1)$$

$$\text{LDR} = \left( \frac{3084}{3084} \right) \times 100 = 100\%$$

## 2. Source Attribution Accuracy

$$\text{Source Attribution Accuracy} = \left( \frac{\text{Confirmed Source-Traced Leaks}}{\text{Total Verified Leaks}} \right) \times 100 \quad (2)$$

$$\text{Source Attribution Accuracy} = \left( \frac{3062}{3084} \right) \times 100 = 99.29\%$$

### 3. Compliance Violation Rate for spams due to leak

$$\text{Compliance Violation Rate} = \left( \frac{\text{Websites Involved in Leaks}}{\text{Total Websites Analyzed}} \right) \times 100 \quad (3)$$

$$\text{Compliance Violation Rate} = \left( \frac{16}{370} \right) \times 100 = 4.32\%$$

### 4. Shared Business Model Leak Rate

$$\text{Shared Business Model Leak Rate} = \left( \frac{\text{Leaks Linked by Business Model}}{\text{Total Verified Leaks}} \right) \times 100 \quad (4)$$

$$\text{Shared Business Model Leak Rate} = \left( \frac{22}{3084} \right) \times 100 = 0.71\%$$

The comparison table 5.2 highlights the strengths and limitations of different data leak detection approaches. Various methods exist for detecting data leaks, each with distinct strengths and limitations. The spam-based approach introduced in this study is compared against three conventional methods: Network Intrusion Detection Systems (IDS), Log Analysis, and Machine Learning-based detection. Table 5.2 summarizes their capabilities across the key performance metrics.

Table 5.2: Comparison of Different Leak Detection Approaches

Metric	Spam-Based	Network IDS	Log Analysis	Machine Learning
<b>Leak Detection Rate</b>	✓	✓	✓	✓
<b>Source Attribution Accuracy</b>	✓		✓	✓
<b>Compliance Violation Detection</b>	✓			
<b>Detection Speed</b>		✓		✓
<b>Shared Business Model Leak Detection</b>	✓			

**Leak Detection Rate** All four approaches demonstrate a strong ability to detect leaks. Spam-based detection achieves this by using spam emails as forensic evidence, while Network IDS relies on monitoring data ex-filtration attempts in real time. Log analysis and machine learning techniques also effectively detect leaks by analyzing system activity and behavioral anomalies.

**Source Attribution Accuracy** Spam-based detection and log analysis perform well in identifying leak sources because they provide concrete evidence linking honeypot identifiers to leak recipients. Machine learning models also exhibit high attribution accuracy by analyzing patterns, but Network IDS lacks direct attribution capabilities unless combined with forensic analysis.

**Compliance Violation Detection** Only the spam-based approach explicitly evaluates compliance

violations by examining whether websites adhere to their stated privacy policies. Traditional methods focus on breach detection rather than regulatory compliance. This makes the spam-based method particularly useful in privacy audits and legal assessments.

**Detection Speed** Network IDS and machine learning models provide real-time or near-instantaneous detection of potential leaks. However, spam-based and log analysis approaches rely on retrospective evidence, detecting leaks only after spam has been received or logs have been examined.

**Shared Business Model Leak Detection** The spam-based model uniquely identifies leaks attributed to shared business models, where user data may not be directly leaked but passed between affiliated entities. This capability is absent in Network IDS, log analysis, and machine learning approaches, which do not account for indirect data sharing.

The spam-based approach provides a novel and effective method for leak detection, particularly in regulatory compliance assessments and source attribution. Network IDS and machine learning models excel in real-time detection but lack compliance violation evaluation. Log analysis offers strong source attribution but depends on extensive log retention and forensic expertise. The spam-based method is the only approach that explicitly detects shared business model leaks, providing unique insights into indirect data misuse.

## 5.11 Comparative Analysis Approach for Fingerprinting

In the domain of data leak detection, various techniques have been proposed to identify and trace the unauthorized exposure of sensitive data. Among these techniques, fingerprinting as used in our research has gained recognition for its unique advantages in detecting data leaks. This analysis compares our fingerprinting technique with other prominent methods, they include Watermarking, Fake Data Injection, and User Behavior Monitoring, as discussed in the works of *Avila et al.* [13] and *Gaikwad et al.* [14]. Table 5.3 shows the comparative analysis of our approach and the other fingerprinting methods.

The fingerprinting technique in our research, leverages spam emails received by honeypot accounts to trace data leaks, identify non-compliant websites, and attribute leaks to their sources. This method is based on unique name variants embedded in the spam content, which allows for efficient leak attribution and the identification of data exposure incidents.

Our fingerprinting technique offers a forensic approach, tracing leaks directly to the source based on the nature of the spam, which serves as an indicator of data exposure. Unlike traditional methods, it uses spam content, thus focusing on indirect indicators of data leaks that may not be directly detectable

by other techniques. our technique offers a high detection rate and is able to provide a reliable method for tracing and attributing leaks with great certainty.

The complexity of our technique is moderate. It requires embedding unique name variants within the dataset and detecting spam emails that match those variants. This method doesn't involve real-time or active monitoring of the system but instead waits for the leak to be detected through spam receipt. Although not overly complex in terms of real-time system monitoring, managing and analyzing the spam data can still require significant processing, especially when scaling. It's relatively simple in terms of real-time analysis, since it's only triggered by the arrival of the spam emails. However, the process of embedding name variants and tracing them requires setup and ongoing analysis of email data.

Watermarking [13, 14] is a method that embeds unique identifiers or invisible markers within datasets to trace leaked data back to its source. This method is commonly used in digital content protection but has been extended to the detection of data leaks. Watermarking provides clear traceability by embedding hidden identifiers in the data, allowing the detection of the source once the data is leaked. It also allows for very accurate leak detection if the watermark is preserved and is not easily detected or removed.

Watermarking can be detected or removed by sophisticated attackers, rendering it ineffective in some cases and it requires access to the leaked data, which may not be immediately available to the data owner or the regulator. Embedding unique identifiers in large datasets can be resource-intensive and impractical for large-scale systems. The complexity of the watermarking technique is high. Embedding unique identifiers in the dataset requires careful management and tracking of watermarks to ensure they are not accidentally altered or removed.

Table 5.3: Comparative Analysis of Data Leak Fingerprinting Techniques

Feature / Method	Our Fingerprinting	Watermarking	Fake Data Injection	User Behavior Monitoring
Real-Time Detection	✓	✗	✗	✓
Forensic Tool	✓	✓	✗	✗
Passive Detection	✓	✗	✗	✗
High Accuracy	✓	✓	✗	✗
Can Handle Large Datasets	✓	✗	✗	✗
Detection of Unauthorized Access	✗	✗	✓	✓
False Positive Rate	✓(low false positives due to unique name variants)	✓	✓	✓
Scalability	✓	✗	✗	✗
Complexity	Moderate	High	Moderate to High	High
Intrusiveness (Active Monitoring)	✓(active monitoring triggered by first spam)	✓	✓	✓
Data Privacy Concerns	✓(privacy policy violations and shared business model mapping)	✗	✗	✓

<b>Feature / Method</b>	<b>Your Fingerprinting</b>	<b>Watermarking</b>	<b>Fake Data Injection</b>	<b>User Behavior Monitoring</b>
Ability to Detect Indirect Leaks	✓	✗	✗	✗
Extent of Exposure	✓ (measured via spam headers and tracing leaks)	✓ (detected through watermark in data)	✗ (indirect measurement)	✗ (indirect measurement through behavior)

Furthermore, Watermarks also need to be persistent across data exchanges, making the implementation complex, especially with large datasets or sensitive data. Watermarking can be complex to implement, especially when dealing with un-trusted systems or adversaries who might attempt to detect and remove the watermark.

While both techniques offer high accuracy in leak attribution, our fingerprinting method provides a more indirect and dynamic approach by utilizing spam emails, which are often received in real-time. Watermarking, on the other hand, requires that the identifier is preserved in the data after the leak, which may not always be feasible or detectable immediately.

Fake data injection [13, 14] involves inserting false or decoy information into datasets to detect unauthorized data exposure. The fake data helps identify when a data leak has occurred, especially when unauthorized parties attempt to access or use the injected fake information. This method is effective at detecting data leaks in systems where the unauthorized access of data can be tracked based on the exposure of fake data. Fake data injection is relatively straightforward to implement and does not require complex tracking mechanisms. Although there is a risk of false positives, where legitimate use of data might be flagged as a leak due to the exposure of fake data. Furthermore, as the size of the dataset grows, managing fake data can become cumbersome, especially if a significant portion of the data is injected with fake information, making the complexity of the technique moderate to high. This complexity arises from ensuring the fake data is properly managed and the system can differentiate between real and fake data without errors. There is also a risk of creating false positives and missing the actual leaks. The complexity increases as the volume of fake data grows. Additionally, attackers might recognize the fake data and exploit it in ways that do not alert the system, leading to missed leaks.

While Fake data injection an effective tool for detecting data leaks, it often requires a proactive insertion of data, which could result in false positives. In contrast, fingerprinting provides an implicit and passive way of detecting leaks based on real-world behavior, such as spam receipt. Fingerprinting offers a more detailed and context-rich approach to detection, as it links spam data directly to specific leaked websites.

User behavior monitoring [13, 14] involves analyzing logs and detecting anomalies in user behavior to spot potential data leaks. This technique focuses on detecting unusual patterns, such as unexpected access to certain data or sudden spikes in data transmissions, that might indicate a leak. User behavior monitoring can offer real-time detection of suspicious activities, which is valuable for preventing further leaks. It enables proactive detection of abnormal behavior, such as unauthorized access to sensitive data. Anomaly detection methods can suffer from high false positive rates, where legitimate activities are

flagged as suspicious. Implementing user behavior monitoring can be complex, as it requires thorough baseline behavior analysis and continuous monitoring of large datasets. Continuous monitoring of user behavior can raise privacy concerns, especially if user activities are being tracked extensively. Our fingerprinting and user behavior monitoring rely on indirect indicators to detect data leaks, but user behavior monitoring offers real-time detection of suspicious activities, whereas fingerprinting relies on post-leak spam data. our Fingerprinting provides a more passive method, whereas user behavior monitoring requires active, ongoing tracking, which might not always be feasible or desirable in privacy-conscious environments. The complexity of the user behavior monitoring technique is high, since it requires constant monitoring of user actions, including log analysis and anomaly detection. The system needs to continuously process large amounts of data in real time, which can be complex, resource-intensive, and prone to high false positive rates. Implementing this method effectively requires sophisticated machine learning models or heuristics to distinguish between normal and anomalous behavior, and continuous monitoring of data.

In conclusion, while our fingerprinting techniques, provide a unique and effective method for detecting data leaks through spam emails, each of the other methods - Watermarking, Fake Data Injection, and User Behavior Monitoring - has its own advantages and limitations. Our approach stands out by offering a non-intrusive, forensic-level analysis, focusing on indirect indicators of leaks and making it particularly valuable for detecting leaks that result in spam-related activities. However, the scalability, real-time detection, and false-positive concerns of the other methods could limit their effectiveness in large-scale systems, whereas our fingerprinting offers a more sustainable and context-specific alternative.

## Chapter 6

# SUMMARY AND CONCLUSIONS

This chapter contains the summary of the experiment, findings, answers to our research question and hypothesis, limitation of the study and future work & recommendations.

### 6.1 Summary

This research introduced a spam-based data leak detection approach that leverages spam emails received by honeypot accounts to identify data leaks, identify non-compliant websites regarding their privacy policies, and attribute leaks to their legitimate website sources. Unlike traditional network-based or log analysis or other fingerprinting methods, this approach provides a forensic perspective by utilizing spam content, account-based analysis (through unique name variants), and mapping-based analysis.

The experiment was conducted over a 12-month period using 148 honeypots, producing 740 account, on 370 websites spread across 12 communities (transport, accommodation, education, sports, news, forum, jobs, health, family, recreation, and social networks). A total of 12,490 spam emails were received from 177 spam senders, including both service provider websites and leak-recipient websites. Of these, 9,497 were advertisement spams, 2,955 were scam spams, and 38 were phishing spams. Statistical analysis was performed to determine the timeline relevance of the spam received. The analysis revealed that most spams arrived within the 00-04 minutes of each hour, with a significant peak on Thursdays, suggesting that spammers align their activities with business email cycles and exploit user behavior.

Our analysis also showed that male individuals between the ages of 48 and 57 are more likely to be targeted by spammers. This may be attributed to their assumed financial stability and lower technical savvy compared to younger generations. Additionally, the relatively low number of spams received by individuals aged 13-17 could be due to their lower likelihood of being targeted for financial fraud, or

their fewer website interactions tied to commercial services. Interestingly, the male gender within the 62-75 age range also saw a significant number of spams, although at a lower level than the middle-aged demographic.

The study further revealed spam received via SMS, indicating a significant need to reduce such spam by de-associating phone numbers from websites through telecom providers. We proposed a nonce-based de-association approach, which would require telecoms to request the unbinding of phone numbers from websites upon line repossession. We believe this would significantly reduce spam while enhancing privacy.

The study also bench-marked the privacy policy statements of defaulting websites against the Canadian Anti-Spam Legislation (CASL). This analysis revealed issues such as the exploitation of implicit consent for mailing lists and third-party sharing, lack of enforcement of privacy policies, and the absence of privacy statements by a defaulting website. These findings underscore the need for a review of the current laws to reduce the spam sending practices.

A total of 3,084 spam emails were classified as leak-related, originating from 16 out of 370 analyzed websites that were confirmed to have leaked data. Using account-based analysis, 3,062 leaks were directly traced to their source through unique honeypot name variants embedded in the spam content and overlap with mapping analysis. The remaining 140 cases required mapping-based analysis, ultimately confirming 3,084 traced leaks with certainty. Additionally, some leaks were inferred through shared business models, where websites with similar operations exhibited suspicious data-sharing behavior.

From a compliance standpoint, 4.32% of the analyzed websites were found to have violated their privacy policies by leaking user data to third parties, despite explicitly stating otherwise. This highlights the limitations of self-regulation in data protection and emphasizes the need for automated compliance monitoring mechanisms.

A comparative analysis of data leak detection methods, including Network Intrusion Detection Systems (IDS), Log Analysis, Machine Learning-based detection, Watermarking, Fake Data Injection, and User Behavior Monitoring, revealed that spam-based detection using the fingerprinting technique is particularly effective for leak attribution and compliance auditing, areas where other methods fall short. While Network IDS and Machine Learning excel in real-time detection, they do not inherently assess compliance violations or track the origin of data leaks. Log Analysis is valuable for retrospective investigations but lacks the granularity of spam-based tracking. Watermarking and Fake Data Injection offer some benefits but depend on data persistence and cannot address indirect leaks or provide detailed leak attribution. User Behavior Monitoring helps detect anomalies but does not directly identify leak

sources and is prone to false positives. Our fingerprinting technique stands out by offering real-time detection, compliance monitoring, and the ability to trace indirect leaks, making it a more comprehensive and reliable solution for data leak detection and privacy enforcement.

Spam-based analysis is uniquely capable of detecting shared business model leaks, an emerging form of indirect data misuse.

## 6.2 How the study addresses the research Questions

This research provided valuable insights into the role of spam in detecting and tracing data leaks by addressing these questions:

### **Understanding the Link Between Data Leaks and Spam:**

Research Question: Can spam serve as a direct indicator of data leaks?

Answer: The study confirmed that spam can indeed serve as a direct indicator of data leaks. Through honeypots, account-based and mapping-based analysis, we traced a total of 3,084 leak-related spam emails to their respective data sources. The analysis of these emails demonstrated that spam, specifically received by users from compromised websites, could reliably identify and attribute data leaks.

### **Effectiveness of Honeypots in Leak Detection:**

Research Question: How can honeypots reveal patterns in the way spammers access and use leaked data?

Answer: Honeypots were integral in revealing how spammers target users based on leaked data. By tracking spam emails, we identified patterns of spamming behavior, such as time-based peaks in spam receipt and the specific demographic groups targeted. These patterns helped trace the leaks back to their originating websites, revealing the mechanism through which spammers accessed and distributed leaked information.

### **Analyzing Spam Ecosystems:**

Research Question: To what extent does leaked data spread through the spam ecosystem? How many entities use the same leaked data to send spam, and what does this reveal about the nature of spam networks?

Answer: The experiment demonstrated the widespread nature of leaked data across the spam ecosystem. It showed that several entities exploit the same leaked data to send spam, which indicates the multi-layered and collaborative nature of spam networks. This behavior highlights the scale of the issue and the potential for aggregated data misuse across various actors.

### **Ethics and Intent Behind Data Leaks:**

Research Question: Is spam an unintended consequence of negligent data practices, or do some websites intentionally leak user data?

Answer: The research suggests both negligence and intentional actions contribute to data leaks. Through privacy policy analysis and mapping of leak-related spam, we found that some websites violated their own privacy policies by leaking data without consent, while others demonstrated poor enforcement of privacy standards, potentially leading to intentional or avoidable data exposures.

### **Novelty of Spam as a Forensic Tool:**

Research Question: Can analyzing the nature and frequency of spam help detect and attribute intentional data leaks? How can spam forensics enhance current data breach detection and investigation methods?

Answer: Spam forensics proved to be a highly effective tool for detecting and attributing data leaks. The ability to analyze the nature, timing, and sources of spam emails allowed us to identify websites that were leaking data, something traditional detection methods like Network IDS or Log Analysis could not achieve. This approach enhances current data breach detection by providing a forensic mechanism that traces leaks back to their sources, even uncovering indirect leaks linked to shared business models.

### **Impact on Privacy and Security:**

Research Question: Can patterns in spam behavior inform new privacy safeguards or security protocols for websites handling user data?

Answer: The patterns observed in spam behavior revealed vulnerabilities in privacy practices (as seen with the case of the exploitation of the implicit clause of the CASL), indicating the need for improved privacy safeguards. Specifically, the research demonstrated the need for stronger enforcement of privacy policies, distinct separation of privacy statements from third party consent agreement (especially in cases where user data is shared with third parties) and mailing list consent agreement which are currently muddled up with privacy statements. This finding supports the argument for stricter privacy regulations and suggests that spam analysis can inform the development of more robust security practices to protect user data.

### **Quantifying the Reach of Leaked Data:**

Research Question: How can we measure the spread of leaked data based on the number of different senders in spam communications? What metrics can be used to assess the scale and impact of a leak through spam patterns?

Answer: The research used the number of spam senders and the distribution of spam over time as

key metrics to assess the scale and impact of data leaks. By analyzing the frequency and spread of spam from different sources, we quantified the extent to which leaked data was disseminated, providing a measurable way to evaluate the severity of data breaches and their impact on users.

Through these findings, the research successfully addressed its core motivation: to explore spam as a forensic tool for data leak detection, understand the mechanisms behind data leaks, and provide insights that could help inform future privacy policies and security practices.

### **6.3 Research Hypothesis:**

At the beginning of the research we postulated the following hypotheses to evaluate data handling practices of websites, consent management practices and factors influencing spam receipt:

**\*1. Users who engage in healthy Internet activities will still experience data leaks over time due to having accounts on websites that may not effectively enforce their privacy policies:**

Meeting the Hypothesis: Our research demonstrates that even users who engage in "healthy" internet activities—such as visiting and interacting solely with their assigned websites—are still at risk of data leaks. This is evident from your findings where spam messages continued to be received from websites that users were subscribed to, but that failed to properly enforce their privacy policies or used the implied clause in the anti-spam law. These websites leaked user data despite the user's intended safe online behavior. The analysis of spam messages confirmed that these leaks were not based on user actions but rather on the websites' failure to enforce privacy standards.

**\*2. There exist a potential link between spam receipt and websites associated with data leaks:**

Meeting the Hypothesis: This hypothesis is strongly supported by our findings. The spam messages we analyzed were traceable back to specific websites, confirming a link between spam receipt and websites with poor data privacy practices or direct data leaks. Our honeypot-based method successfully tracked spam messages and associated them with the originating websites, confirming that spam could be a key indicator of data leaks. In particular, 3,084 spam emails were classified as leak-related, originating from 16 out of 370 websites that were confirmed to have leaked data.

**\*3. The spammer's intent (distributive or harvesting) can be inferred from the nature of the received messages:**

Meeting the Hypothesis: The hypothesis has been met through our detailed analysis of the spam messages. We identified different types of spam (advertisements, scams, phishing) and analyzed their content to infer the intent behind each message. For example, phishing attempts were identified as

having a clear intent of harvesting personal information, while scam were more related to distributive spam. Our research showed that by analyzing the content and patterns of the spam messages, we could reliably infer whether the spammer's intent was to distribute or harvest user data.

**\*4. The receipt of spam by users is not directly correlated with their actual voluntary consented subscriptions at the time of account creation on websites:**

Meeting the Hypothesis: This hypothesis has been strongly validated in our research. Through the analysis of spam messages received by honeypot accounts, it became clear that the receipt of spam was not correlated with users' voluntary consented subscriptions. Many spam messages were received from websites to which users had never explicitly subscribed or had an account but not voluntarily subscribed to subscription spam as they were implicitly included in privacy policy statements, indicating that the users' actual consent was not the determining factor in whether they received spam. This supports the idea that spam receipt is often a result of data leaks or improper handling of user data by websites, not directly tied to the user's explicit consent when creating an account.

## **6.4 Conclusion**

This study demonstrates that spam emails can serve as valuable forensic evidence for detecting and attributing data leaks. By integrating account-based and mapping-based techniques, this approach effectively identified leaked data sources, tracked non-compliance, and exposed indirect leaks resulting from shared business models.

Overall, our findings, which span from the high volume of spam received from websites where our honeypots did not expressly consent to mailing list subscriptions, to tracing data leaks back to legitimate suspect websites, as well as observing the distributive patterns of certain websites and the timeline analysis of spam receipt at minute, hourly, day-of-the week and monthly intervals, strongly support the validity of our four hypotheses and research questions.

Our findings also showed how honeypots was/can be used to reveal patterns in the spammer's timeline activities. This insight can be useful in training spam filtering models or implementing stricter spam rules at certain timelines for spam detection.

The high detection rate of 100% and the ability to confirm leak sources with 3,084 traced cases validate the reliability of this method. However, real-time detection remains a challenge, as this approach depends on receiving spam emails, unlike traditional IDS-based detection methods. Additionally, while

spam-based detection is effective for compliance violation assessment, it does not prevent leaks proactively but rather serves as an investigative tool.

In addition, the experiment expounded on how spammers access and utilize leaked data. As well as the extent to which leaked data spreads. Furthermore the number of entities leveraging the same leaked data to send spam within the spam ecosystem was seen as with the case of some websites in the family community (e.g. datemyage.com) The experiment, also showed privacy and consent practices of legitimate websites that require urgent reforms to the existing privacy safeguards for websites that manage user data.

The findings reinforce the urgent need for stronger privacy enforcement and highlight how spam-based leak detection can serve as an early warning system for data leaks, aiding regulatory bodies and security professionals in protecting user data. This research provides a novel forensic approach to data leak detection that not only traces leaks but also identifies privacy policy violations. The high accuracy and effectiveness of our spam-based method make it a valuable tool for regulatory compliance auditing.

## **6.5 Limitation of the Work**

One potential limitation of our technique is its reliance on spam emails as a data source for detecting data leaks. Since the system depends on spam messages, real-time leak detection may be compromised as there could be delays in identifying leaks that do not directly result in spam messages. As such, data leaks that do not lead to spam may not be detected immediately, which makes the technique less suitable for real-time detection compared to methods that provide instant alerts.

Additionally, false positives could arise from the mapping-based analysis, as indirect associations between spam sources and leaked data might mistakenly identify leaks that are not actually related. This can be especially problematic when spam messages share characteristics with legitimate communications, making it harder to distinguish between actual leaks and unrelated activities. Despite this challenge, we believe this limitation can be addressed by implementing continuous training of an automated mapping model that can learn from the data over time. By incorporating machine learning techniques, the method can adapt to new patterns, reducing false positives and improving leak detection accuracy for the mapping based analysis. Given the high rate of certainty from the account-based (fingerprinting) approach, this continuous learning process would help enhance the system's reliability and mitigate delays in detecting potential leaks.

Lastly, spam filtering mechanisms may impact the effectiveness of this technique by reducing the

volume of spam messages available for analysis. This reduction in data could limit the sample size, making it harder to identify patterns and reduce the overall effectiveness of the technique in certain environments where spam filtering is aggressive.

## **6.6 Future Work & Recommendations**

To better enhance the method, we look to

- integrate deep learning techniques to develop an unsupervised spam data leak forensic tool, which can be further extended to create a more robust spam filtering model with 0% false positives in our mapping based analysis.
- Enhancing real-time detection by integrating spam-based analysis with machine learning models
- Developing a regulatory compliance monitoring system based on spam-driven forensic analysis.
- Expanding honeypot diversity to analyze a broader range of industries to include financial institutions (banks), government platforms, etc., to capture a wider range of privacy practices and data leakage risks.
- Expanding regulatory applications to monitor large-scale privacy compliance following other standard privacy regulator bodies such as Europe (GDPR compliance) or Asia.

# Appendix A

## Appendix

### A.1 Definition of terms:

**Legitimate website:** These are websites in our website pool accessed in the research and form part of our website list.

**Leak recipient website:** Websites where our honeypot did not create accounts or interact with these website.

**SWU = USP:** These codes are equivalent. The code represents the untargeted spams received from legitimate websites within our website pool.

**SWT = TSP:** These codes are equivalent. The code represents the targeted spams received from legitimate websites within our website pool.

**UTL = USL:** These codes are equivalent. These codes represents the untargeted spams received from leak recipient websites.

**TLP:** These codes represents the targeted spams received from leak recipient websites.

### A.2 Publication

The details of the paper for this work is accepted for publication in:

- Oghenerukevwe Oyinloye and Carol Fung. "Evaluating website Data Leaks through Spam Collection on Honeypots" (2025) The 15th ACM Conference on Data and Application Security and Privacy (CODASPY) ACM, 2025.

### A.3 Research Data

The list of websites used in this research with their associated ID can be found at our repository:

<https://github.com/Spams-from-Honeypots/Honeypot-Spam-Collection.git>

Honeypot website assignment can be accessed at: <https://github.com/Spams-from-Honeypots/Honeypot-Spam-Collection.git>

To access our research raw email-spam dataset, visit our repository at: <https://github.com/Spams-from-Honeypots/Honeypot-Spam-Collection.git>

To access our research raw SMS-spam dataset, visit our repository at: <https://github.com/Spams-from-Honeypots/Honeypot-Spam-Collection.git>

List of defaulting websites with their associated Consent Practice Exploitation, visit our repository at:

<https://github.com/Spams-from-Honeypots/Honeypot-Spam-Collection.git>

#### Honeypot - Website Assignment:

Table A.1: Honeypot - Website Assignment Table

HoneypotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
1	ellen	adams	1	2	3	4	5
2	david	allen	1	6	7	8	9
3	doreen	baker	2	6	10	11	12
4	michele	payne	3	7	10	13	14
5	mary	Nohis	4	8	11	13	15
6	james	bennett	5	9	12	14	15
7	penny	chans	16	17	18	19	20
8	george	chens	16	21	22	23	24
9	michelle	cloutier	17	21	25	26	27
10	pierre	collins	18	22	25	28	29
11	wanda	cormier	19	23	26	28	30
12	steven	cote	20	24	27	29	30

HoneyPotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
13	jason	scotia	31	32	33	34	35
14	monique	lesile	31	36	37	38	39
15	julie	dubie	32	36	40	41	42
16	ernest	ferguson	33	37	40	43	44
17	sandra	fraser	34	38	41	43	45
18	fernand	gaboury	35	39	42	44	45
19	carol	newman	46	47	48	49	50
20	donald	Jean-Brillon	46	51	52	53	54
21	christine	girad	47	51	55	56	57
22	stan	occonor	48	52	55	58	59
23	sarah	hamilton	49	53	56	58	60
24	Terry-Pie	harris	50	54	57	59	60
25	gilles	huppe	61	62	63	64	65
26	marc	jackson	61	66	67	68	69
27	elizabeth	Groulx	62	66	70	71	72
28	rick	Verdun	63	67	70	73	74
29	margaret	Champlain	64	68	71	73	75
30	raymond	Bishop-Powder	65	69	72	74	75
31	wendy	Banting	76	77	78	79	80
32	dianne	legare	76	81	82	83	84
33	Ryan	francois	77	81	85	86	87
34	alain	kennedy	78	82	85	88	89
35	melissa	namur	79	83	86	88	90
36	jack	savane	80	84	87	89	90
37	denise	lambert	91	92	93	94	95
38	ronald	Lavioe	91	96	87	98	99
39	fred	Leclerc	92	96	100	101	102
40	lorne	charlene	93	97	100	103	104
41	cheryl	Levesque	94	98	101	103	105
42	marcel	cavendish	95	99	102	104	105

HoneyPotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
43	louise	moore	106	107	108	109	110
44	manon	morrison	106	111	112	113	114
45	lori	murray	107	111	115	116	117
46	bill	nadeau	108	112	115	118	119
47	andrea	maurice	109	113	116	118	120
48	barry	patricia	110	114	117	119	120
49	tammy	cameron	121	122	123	124	125
50	venessa	desrosiers	121	126	127	128	129
51	yves	leblanc	122	126	130	131	132
52	dale	cooper	123	127	130	133	134
53	craig	marshall	124	128	131	133	135
54	janice	martin	125	129	132	134	135
55	betty	millier	136	137	138	139	140
56	louis	murphy	136	141	142	143	144
57	danielle	poirier	137	141	145	146	147
58	megan	beaudry	138	142	145	148	149
59	neil	reynolds	139	143	146	148	150
60	richard	jonathan	140	144	147	149	150
61	sylvie	alexis	151	152	153	154	155
62	maurice	pelletier	151	156	157	158	159
63	rose	robinson	152	156	160	161	162
64	shawn	rogen	153	157	160	163	164
65	joyce	Rusell	154	158	161	163	165
66	kenneth	Sartre	155	159	162	164	165
67	glen	shaws	166	167	168	169	170
68	victor	mederios	166	171	172	173	174
69	tanya	smith	167	171	175	176	177
70	StJean	christopher	168	172	175	178	179
71	ashley	thibault	169	173	176	178	180
72	erin	thomas	170	174	177	179	180

HoneyPotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
73	trevor	turner	181	182	183	184	185
74	jackie	brousier	181	186	187	188	189
75	colleen	waler	182	186	190	191	192
76	bassey	haimona	183	187	190	193	194
77	rachel	wilson	184	188		193	195
78	jamie	russell	185	189	192	194	195
79	dorothy	green	196	197	198	199	200
80	dean	morgan	196	201	202	203	204
81	carole	fister	197	201	205	206	207
82	colin	coleman	198	202	205	208	209
83	tara	kimberly	199	203	206	208	210
84	darren	bilodeau	200	204	207	209	210
85	jerry	kelly	211	212	213	214	215
86	damisi	aldrix	211	216	217	218	219
87	pauline	steele	212	216	220	221	222
88	philip	warren	213	217	220	223	224
89	todd	carrol	214	218	221	223	225
90	ross	sharpe	215	219	222	224	225
91	connie	mayer	226	227	228	229	230
92	thersea	moreau	226	231	232	233	253
93	jacqueline	bellemare	227	231	235	236	237
94	jung	tanjung	228	232	235	238	239
95	grace	vallee	229	233	236	238	240
96	danny	bertrand	230	234	237	239	240
97	yvonne	pattison	241	242	243	244	245
98	marlene	baxter	241	246	247	248	249
99	cynthia	rioux	242	246	250	251	252
100	bryan	carrier	243	247	250	253	254
101	dana	lamoureux	244	248	251	234	255
102	vincent	lauzon	245	249	252	254	255

HoneyPotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
103	claire	irvine	256	257	258	259	260
104	kediat	phang	256	261	262	263	264
105	sylvia	rondeau	257	261	265	266	267
106	gaetan	sidhu	258	262	265	268	269
107	kerry	dickson	259	263	266	268	270
108	aaron	boily	260	264	267	269	270
109	jose	neufeld	271	272	273	274	275
110	samantha	berg	271	292	277	278	279
111	kyle	martinez	272	276	280	281	282
112	sinyrat	surbakti	273	277	280	283	284
113	bruno	norman	274	278	281	283	285
114	evelyn	labonte	275	279	282	284	285
115	sonia	clement	286	287	288	289	290
116	jeremy	lake	286	291	292	293	294
117	tyler	johns	287	291	295	296	297
118	bernie	berry	288	276	295	298	299
119	alison	laurendeau	289	294	296	298	300
120	june	dussault	290	293	297	299	300
121	oyinbolade	ahumere	301	302	303	304	305
122	langa	ndaba	301	306	307	308	309
123	kuban	manik	302	306	310	311	312
124	pinem	phiong	303	307	310	313	314
125	sitepu	thung	304	308	311	313	315
126	theron	mabunda	305	309	312	314	315
127	karen	anderson	316	317	318	319	320
128	robert	amstrong	316	321	322	323	324
129	jeffrey	beaulieu	317	321	325	326	327
130	benjamin	jacob	318	324	325	328	329
131	linda	bergeron	319	323	326	328	330
132	mark	bernier	320	324	327	329	330

HoneyPotID	first_name	last_name	w_1	w_2	w_3	w_4	w_5
133	susan	bouchard	331	332	333	334	335
134	kevin	Boulevard	331	336	337	338	339
135	diane	couture	332	336	340	341	342
136	frank	davis	333	338	340	343	344
137	claudio	fortin	334	337	341	343	345
138	alexander	fournier	335	362	342	344	345
139	heather	gagne	346	366	348	349	350
140	emily	henderson	346	351	352	353	354
141	francis	stephen	347	351	355	356	357
142	doug	gauthier	361	352	355	358	369
143	barbara	grant	349	353	356	358	360
144	john	grenier	350	354	357	359	360
145	laura	harvey	361	339	363	364	365
146	bruce	Angrinon	348	365	366	367	368
147	anne	Vertu	362	347	370	366	359
148	mike	Deshawn	363	367	370	368	369

## A.4 Mathematical Formula for Confidence Intervals

(1) The mean spam count is given by  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ,

where  $n = 12$  represents the total number of months:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\bar{X} = \frac{1}{12} (X)$$

(2) the standard deviation:

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

$$s = 590.63$$

(3) Margin of Error(ME) at 99% confidence level, the Z-critical is 2.576

$$ME = Z \times \frac{s}{\sqrt{n}}$$

Where:

- $ME$  is the margin of error
- $Z$  is the Z-score corresponding to the desired confidence level
- $s$  is the standard deviation of the population
- $n$  is the sample size

(4) Confidence Interval (CI)

$$CI = \bar{X} \pm Z_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

Where:

- $\bar{X}$  is the sample mean
- $Z_{\alpha/2}$  is the Z-critical value corresponding to the desired confidence level
- $s$  is the population standard deviation (or sample standard deviation if population standard deviation is unknown)
- $n$  is the sample size

## **A.5 Data Leak Incident Report:**

Table .2: Summary of Spam Incidents and Leaks

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account- Mapping- Based	Based
source1	cheryl's food	maurice, pelletier	medallia.com (A)	Mar. 10, unique name variant	Phish	✓	
source2	datemyage.com	colleen, waler, jamie, russell	dating.com (B)	Apr. 7, Apr. 21, overlapping web-site + unique name variant	Scam	✓	✓
source2	datemyage.com	colleen, waler	PinaDate.com (C)	Apr. 25, unique name variant	Scam	✓	
source2	datemyage.com	colleen, waler	EuroDate.com (D)	Apr. 16, unique name variant	Scam	✓	✓
source2	datemyage.com	colleen, waler, jamie, russell	AsianDate.com (E)	May 5, May 16, overlapping web-site + unique name variant	Scam	✓	
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac-counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account-Based	Mapping-Based
source2	datemyage.com	colleen, waler, jamie, russell	Once notifications@getonce.com (F)	Apr. 3, Apr. 18, overlapping website + unique name variant	Scam	✓	✓
source2	datemyage.com	colleen, waler	Anastasiadate.com (G)	Apr. 27, unique name variant	Scam	✓	
source3	garageclothing.com	ross sharpe	Dynamitestyle.com (H)	Jun. 12, brand association	Scam		✓
source4	www.chinalove.com	jeremy, lake, june, dussault	Dating.com (B)	Mar. 21, Mar. 31, overlapping website + unique name variant	Scam	✓	✓
source4	www.chinalove.com	jeremy, lake, june, dussault	Amolatina.com (J)	Mar. 28, Mar. 31, overlapping website + unique name variant	Scam	✓	✓
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account- Mapping- Based	Based
source4	www.chinalove.com	jeremy, lake, june, dussault	datemyage.com (I)	Mar. 21, Mar. 31, overlapping website + unique name variant	Scam	✓	✓
source4	www.chinalove.com	june, dussault	getonce.com (F)	Mar. 31, over- lapping website + unique name vari- ant	Scam	✓	✓
source4	www.chinalove.com	jeremy, lake, june, dussault	AmalDate.com (K)	Apr. 2, Apr. 6	Scam	✓	✓
source4	www.chinalove.com	jeremy, lake, june, dussault	Your TravelMates.com (L)	Apr. 3, Apr. 6, overlapping web- site + unique name variant	Scam	✓	✓
source4	www.chinalove.com	jeremy, lake, june, dussault	EuroDate Team (D)	Apr. 9, Apr. 10	Scam	✓	
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account Based	Mapping- Based
source5	www.menshealth.com	fermand, gaboury	menshealth@eml.hearstmags.com (N)	Mar 22, brand as- sociation	Advert		✓
source6	yemeksepeti.com	christine girad	Hungryroot hotlogif- mom.hungryroot.com (O)	Feb. 18, shared ser- vice model	Advert		✓
source7	simplyHired	wanda, cormier	Indeed donotre- ply@indeed.com (P)	Mar. 3, business type	Advert		✓
source8	thestar.com	les, huppe	Toronto Star nore- ply@torstar.ca (Q)	Mar. 8, brand asso- ciation	Advert		✓
source9	www.dating.com	john, grenier, fran- cis, stephen	Amolatina.com (J)	Mar. 13 and Apr. 6 respectively, over- lapping website + unique name vari- ant	Scam	✓	✓
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account- Mapping- Based	Based
source9	www.dating.com	john, grenier, fran- cis, stephen	AmalDate.com (K)	Mar. 21, Nov. 3 respectively, over- lapping website + unique name vari- ant	Scam	✓	✓
source9	www.dating.com	john, grenier, fran- cis, stephen	Your TravelMates.com (L)	Mar. 6, Nov. 5 re- spectively	Scam	✓	✓
source9	www.dating.com	john, grenier, fran- cis, stephen	EuroDate.com (D)	Mar. 16, Apr. 3 respectively, over- lapping website + unique name vari- ant	Scam	✓	✓
source9	www.dating.com	francis, stephen	AsianDate.com (E)	Apr. 26, unique name variant	Scam	✓	✓
source9	www.dating.com	john, grenier	Hotti.com (S)	Oct. 10, unique name variant	Scam	✓	
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account-Mapping- Based	Based
source9	www.dating.com	john, grenier	DateMyAge.com (I)	Feb. 26	Scam	✓	✓
source9	www.dating.com	francis, stephen	ZenDate.com (R)	Nov. 3, unique name variant	Scam	✓	✓
source10	zoosk.com	dale, cooper	laura@usadatingz.com (T)	Oct. 10, unique name variant	Scam	✓	
source11	academicearth.org	glen, shaws	pitt.edu (U)	Feb. 25, unique name variant	Scam	✓	
source12	dominos	alexander, fournier	news@n.factor-meals.ca (V)	Nov. 18, shared business model	Scam		✓
source13	blue mountain resort	alexander, fournier	no-reply@emailmeform.com (W1)	Aug. 18, shared business model	Phish		✓
source13	blue mountain resort	alexander, fournier	studienberatung-koeln@hmkw (W2).de	Aug. 25, shared business model	Phish		✓
Continued on next page							

Continued from previous page

Source SpamID	Suspect Websites	Honeypot Ac- counts Affected	Who was it leaked to	Estimated Date of Leak	Class of Leak	Account-Mapping- Based	Based
source13	blue mountain resort	alexander, fournier	vote4kellysemrad- gmail.com@beuzpro.com (W3)	Sep. 18, shared business model	Phish		✓
source13	blue mountain resort	alexander, fournier	anton@consulting.de (W4)	Oct. 13, shared business model	Phish		✓
source13	blue mountain resort	alexander, fournier	diffusion@lesartoseurs.org (W5)	Oct. 21, shared business model	Phish		✓
source13	blue mountain resort	alexander, fournier	bookings1@pothys.com (W6)	Nov. 1, shared business model	Phish		✓
source14	zdnnet.com	jacqueline, belle- mare	mtsupport.redventures.com (X)	Apr. 9, shared busi- ness model	Phish		✓
source15	istockphoto.com	jeffrey, beaulieu	gettyimages.com (Y)	Jan. 7, shared busi- ness model	Advert		✓
source16	http://www.lowes.com	marlene, baxter	CorporateSearchLower.com invite@decipherinc.com (M)	Jun. 18, close busi- ness model survey content	Phish		✓

## .1 SpammerID by Spammer Email Address:

Table .3: Spam Sender Information: TLP

<b>TLP Spam ID</b>	<b>Sender</b>
TLP000001	Pitt Health Informatics Online healthinformaticsonline@shrs.pitt.edu
TLP000002	notifications@dating.com
TLP000003	PinaDate Team notifications@pinadate.com
TLP000004	Longs Drugs From: cvs@express.medallia.com
TLP000005	DateMyAge Team From: notifications@datemyage.com
TLP000006	AmoLatina Team From: notifications@amolatina.com
TLP000007	AmalDate Team From: notifications@amaldate.com
TLP000008	YourTravelMates Team From: notifications@yourtravelmates.com
TLP000009	EuroDate Team From: notifications@eurodate.com
TLP000010	Hungryroot hello@from.hungryroot.com
TLP000011	AsianDate Team notifications@asiandate.com
TLP000012	Once notifications@getonce.com
TLP000013	AnastasiaDate Team notifications@anastasiadate.com
TLP000014	Indeed donotreply@indeed.com
TLP000015	laura@usadatingz.com
TLP000016	ZenDate Team notifications@zendate.com
TLP000017	Hotti Team From: notifications@hotti.com
TLP000018	messaging-service@post.xero.com

Table .4: Spam Sender Information: (UTL=ULS)

<b>UTL Spam ID</b>	<b>Sender</b>
UTL000001	Support zdnet@mtsupport.redventures.com
UTL000002	Dynamite From: style@emails.dynamitestyle.com
UTL000003	CorporateResearch@Lowes.com From: invite@decipherinc.com
UTL000004	notifications@datemyage.com
UTL000005	notifications@getonce.com
UTL000006	menshealth@eml.hearstmags.com

<b>UTL Spam ID</b>	<b>Sender</b>
UTL000007	notifications@amolatina.com
UTL000008	Toronto Star noreply@torstar.ca
UTL000009	Hungryroot hello@from.hungryroot.com
UTL000010	birdsandblooms@email.birdsandblooms.com
UTL000011	notifications@dating.com
UTL000012	Instagram security@mail.instagram.com
UTL000013	notifications@eurodate.com
UTL000014	Post CA From: no-reply@emailmeform.com
UTL000015	studienberatung-koeln@hmkw.de
UTL000016A	vote4kellysemrad-gmail.com@beuzpro.com
UTL000016B	anton@comsulting.de/ agof@comsulting.de
UTL000016C	diffusion@lesartsoseurs.org
UTL000016D	salahdliy183@gmail.com
UTL000016E	bookings1@pothys.com
UTL000017	notifications@m.factormeals.ca
UTL000018	notifications@yourtravelmates.com
UTL000019	notifications@zendate.com
UTL000020	notifications@amaldate.com
UTL000021	info@engage.gettyimages.com

Table .5: Spam Sender Information (SWT-TSP)

<b>Spam ID</b>	<b>Sender</b>
SWT00001	Filip from ADONIS Team news@boc-group.com
SWT00002	GameStop Canada - EDGE donotreply@edge.gamestop.ca
SWT00003	lmccarty@edx.org
SWT00004	The Moz Team team@moz.com From: team@moz.com
SWT00005	noreply@uber.com
SWT00006	mailer@connect.match.com
SWT00007	ZenDate Team From: notifications@zendate.com
SWT00008	Spencer from Podia From: create@podia.com

<b>Spam ID</b>	<b>Sender</b>
SWT00009	Cheesecake Rewards From: no-reply@rewards.thecheesecakefactory.com
SWT00010	Dating Team From: notifications@dating.com
SWT00011	Animaker tona@animaker.es
SWT00012	The meez team newsletter@getmeez.com
SWT00013	HyreCar partner@hyrekar.com
SWT00014	American Airlines AAdvantage Program loyalty@loyalty.ms.aa.com
SWT00015	contact@olivercabell.com
SWT00016	mail@messaging.zoosk.com
SWT00017	newsletter@topuniversities.com
SWT00018	Medium noreply@medium.com
SWT00019	LinkedIn jobs-listings@linkedin.com
SWT00020	Indeed no-reply@indeed.com
SWT00021	Salt Spring Coffee orders@saltspringcoffee.com
SWT00022	Lingoda students@hello.lingoda.com
SWT00023	Alibaba.com alibaba@service.alibaba.com
SWT00024	Microsoft Microsoft@notificationmail.microsoft.com
SWT00025	eharmony info@offers.eharmony.ca
SWT00026	DateMyAge Team notifications@datemyage.com
SWT00027	Poshmark Info info@poshmark.com
SWT00028	paulr@flixier.com
SWT00029	recommendations@discover.pinterest.com
SWT00030	info@vitasave.ca
SWT00031	bestbuy@email.bestbuy.com
SWT00032	hello@duolingo.com
SWT00033	premium@academia-mail.com
SWT00035	Ridesharing info@ridesharing.com
SWT00036	Narcity hello@mail.narcity.com
SWT00037	The Flickr Team flickrteam@arrow.flickr.com
SWT00038	My Place Rewards thechildrensplace@notification.childrensplace.com
SWT00039	SmartSweets hello@smartsweets.com
SWT00040	JCPenney Rewards JCPenney@email.jcpenney.com

<b>Spam ID</b>	<b>Sender</b>
SWT00041	Andie Swim hello@andieswim.com
SWT00042	CareerBuilder From: no-reply@sites.careerbuilder.com
SWT00043	Anna Ford From: anna@m.bookclubs.com
SWT00044	SENREVE hello@senreve.com
SWT00045	BabyCenter babycenter@newsletters.babycenter.com
SWT00046	teachable@news.teachable.com
SWT00047	email@email.matchesfashion.com
SWT00048	hello@udacity.com
SWT00049	mail@e.jackjones.com
SWT00050	Robert Half no.reply@email.roberthalf.com
SWT00051	support@shopdiva.ca
SWT00052	noreply@glassdoor.com
SWT00053	hometeam@epicure.com
SWT00054	SmokersPersonals notify@datingvipnotifications.com
SWT00055	Candy Funhouse CA noreply.canada@candyfunhouse.com

Table .6: Spam Sender Information:(SWU=USP)

<b>Spam ID</b>	<b>Sender</b>
SWU00001	Frank And Oak do-not-reply@send.frankandoak.com
SWU00002	info@meetup.com
SWU00003	news@sfmtc.edx.org
SWU00004	Julianna at Thrive Market From: hello@thrivemarket.com
SWU00005	Match From: mailer@connect.match.com
SWU00006	Uber From: noreply@uber.com
SWU00007	Garage From: style@email.garageclothing.com
SWU00008	Toys R Us Canada From: toysrus@emails.toysrus.ca
SWU00009	MyLowe's Rewards From: lowes@e.lowes.com
SWU00010	The Houzz Pro Team From: messages@houzz.com
SWU00011	Titan Computers From: sales@titancomputers.com
SWU00012	HelloFresh From: news@newsletter.hellofresh.ca

<b>Spam ID</b>	<b>Sender</b>
SWU00013	X (formerly Twitter) From: info@twitter.com
SWU00014	AOL Mail Team From: aolmailteam@comms.aol.net
SWU00015	DateMyAge Team From: notifications@datemyage.com
SWU00016	ZenDate Team From: notifications@zendate.com
SWU00017	Cheesecake Rewards From: no-reply@rewards.thecheesecakefactory.com
SWU00018	Dating Team From: notifications@dating.com
SWU00019	The RealReal From: email@e.therealreal.com
SWU00020	Moores From: no-reply@mooresclothing.com
SWU00021	Arkadium From: community@arkadium.com
SWU00022	Tona from Animaker From: tona@animaker.es
SWU00023	Kijiji Autos From: email@shop.kijijiautos.ca
SWU00024	Grant Goldberg From: support@musicbed.com
SWU00025	globe@globeandmail.com
SWU00026	Team Biteable From: support@biteable.com
SWU00027	Tripadvisor From: awards@mp1.tripadvisor.com
SWU00028	The meez Podcast From: newsletter@getmeez.com
SWU00029	Square From: noreply@messaging.squareup.com
SWU00030	Twitch From: no-reply@twitch.tv
SWU00031	Dribbble From: no-reply@n.dribbble.com
SWU00032	Candy Funhouse CA From: noreply.canada@candyfunhouse.com
SWU00033	MakeMyTrip From: noreply@zen-makemytrip.com
SWU00034	Lufa Farms From: info@lufa.com
SWU00035	editorialstaff@flipboard.com
SWU00036	HyreCar From: help@hyrekar.com
SWU00037	hello@koio.co
SWU00038	Support@email.yidio.com
SWU00039	contact@olivercabell.com
SWU00040	mail@info.gumtree.com
SWU00041	no-reply@indeed.com
SWU00042	mail@messaging.zoosk.com
SWU00043	memberservices@mail.christianmingle.com

<b>Spam ID</b>	<b>Sender</b>
SWU00044	delish@newsletter.delish.com
SWU00045	no_reply@notification.moomoo.com
SWU00046	news@info.misfitsmarket.com
SWU00047	hello@e.animoto.com
SWU00048	newsletter@topuniversities.com
SWU00049	health@messages.webmd.com
SWU00050	contact@rebag.com
SWU00051	Men's Health newsletter@menshealth.com
SWU00052	The Khan Academy team hello@khanacademy.org
SWU00053	no-reply@email.latinamericancupid.com
SWU00055	First Up — The Star newsletters@thestar.ca
SWU00056	Instacart no-reply@customers.instacartemail.com
SWU00057	Alibaba.com member@notice.alibaba.com
SWU00058	Lingoda students@hello.lingoda.com
SWU00059	Wix.com wix-team@emails.wix.com
SWU00060	Target targetnews@em.target.com
SWU00061	thechildrensplace@promo.childrensplace.com
SWU00062	info@twobears.ca
SWU00063	Yarnspirations yarnspirations@email.yarnspirations.com
SWU00064	CafePress cafepress@mail.cafepress.com
SWU00065	Poshmark Info info@poshmark.com
SWU00066	microsoft@notificationmail.microsoft.com
SWU00067	GoodRx no-reply@contact.goodrx.com
SWU00068	Robert Half no.reply@email.roberthalf.com
SWU00069	ArtPal hello@ArtPal.com
SWU00070	Artblr.com info.artblr.com@email.cyberimpact.com
SWU00071	no-reply@animal-id.net
SWU00072	email@email.matchesfashion.com
SWU00073	contact@coffeemeetsbagel.com
SWU00074	xumo@email.xumo.com
SWU00075	promote@shopperplus.com

<b>Spam ID</b>	<b>Sender</b>
SWU00076	support@shopdiva.ca
SWU00077	recommendations@discover.pinterest.com
SWU00078	donotreply@smokersingles.com
SWU00079	bestbuy@email.bestbuy.com
SWU00080	info@vitasave.ca
SWU00081	hello@duolingo.com
SWU00082	info@join.netflix.com
SWU00083	Animoto hello@e.animoto.com
SWU00084	Tuango - Montreal South Shore mail@email.tuango.ca
SWU00085	Chris from INDOCHINO reply@sfmtmail.indochino.com
SWU00086	Tiny Arduino Drone Instructables no-reply@mail.newsletter.instructables.com
SWU00087	The Flickr Team flickrteam@arrow.flickr.com
SWU00088	Deezer mail@music.deezer.com
SWU00089	ShopRite shoprite@e.shoprite.com
SWU00091	Teachable teachable@hello.teachable.com
SWU00092	Rakuten Canada memberservices@newsletter.rakuten.ca
SWU00093	BabyCenter babycenter@newsletters.babycenter.com
SWU00094	community@em.takelessons.com
SWU00095	no-reply@email.latinamericancupid.com
SWU00096	Wayfair editor@members.wayfair.com
SWU00097	Teleflora Rewards loyalty@loyalty.teleflora.com
SWU00098	Allrecipes The Scoop email@email.allrecipes.com
SWU00099	Lyft no-reply@marketing.lyftmail.com
SWU00100	Medium hello@medium.com
SWU00102	22Fresh deals@22fresh.com
SWU00103	Tubi updates@watch.tubitv.com
SWU00104	SmartSweets hello@smartsweets.com
SWU00105	JCPenney JCPenney@email.jcpenney.com
SWU00106	Andie Swim hello@andieswim.com
SWU00107	Bookclubs help@m.bookclubs.com
SWU00108	SmokersPersonals notify@datingvipnotifications.com

<b>Spam ID</b>	<b>Sender</b>
SWU00109	The Montreal Gazette marketing-e@postmedia.com
SWU00110	thebeardstruggle@thebeardstruggle.com
SWU00111	Recipe of the Day from Taste of Home tasteofhome@email.tasteofhome.com
SWU00112	POF From: donotrespond@pof.com
SWU00113	Kohl's From: kohls@s.kohls.com
SWU00114	Narcity From: hello@mail.narcity.com
SWU00115	same as 31
SWU00116	Clipchamp From: noreply@info.clipchamp.com
SWU00117	Bulk Barn From: bulkbarn@mail.bulkbarn.com
SWU00118	SENREVE From: hello@senreve.com
SWU00119	Ticketmaster newsletter@email.ticketmaster.com
SWU00120	LiveHealth Online startlivehealthonline@em.startlivehealthonline.com
SWU00121	monster@emails.monster.com
SWU00122	hudsonsbay@thebay.hudsonsbay.com
SWU00123	marriottbonvoy@email-marriott.com
SWU00124	hello@kaaibags.com
SWU00125	Playful Minds hello@playfulmindstoys.ca
SWU00126	info@glassdoor.com
SWU00127	coursera@m.learn.coursera.org
SWU00128	mail@e.jackjones.com
SWU00129	hometeam@epicure.com
SWU00130	noreply@summersalt.com
SWU00131	ae-news-interest02@mail.aliexpress.com
SWU00132	dairyqueen@mail.dq.com
SWU00133	customerservice@jrtoycompany.ca

## **.2 Spam Count by ID**

Table .7: Leak Data Table Count

<b>spammer_ID</b>	<b>num_spam</b>	<b>spammer_ID</b>	<b>num_spam</b>
TLP000001	22	UTL000001	1
TLP000002	1427	UTL000002	1
TLP000003	147	UTL000003	2
TLP000004	20	UTL000004	17
TLP000005	424	UTL000005	21
TLP000006	257	UTL000006	6
TLP000007	146	UTL000007	11
TLP000008	112	UTL000008	44
TLP000009	252	UTL000009	7
TLP000010	1	UTL000010	2
TLP000011	52	UTL000011	28
TLP000012	17	UTL000012	2
TLP000013	4	UTL000013	11
TLP000014	3	UTL000014	2
TLP000015	1	UTL000015	1
TLP000016	13	UTL000016	6
TLP000017	10	UTL000017	3
TLP000018	1	UTL000018	3
		UTL000019	3
		UTL000020	2
		UTL000021	1

Table .9: Service Provider Spam Data: TSP and USP

<b>TSP (SWT)</b>		<b>USP (SWU)</b>	
<b>spammer_ID</b>	<b>num_spam</b>	<b>spammer_ID</b>	<b>num_spam</b>
SWT00001	34	SWU00001	5
SWT00002	2	SWU00002	39
SWT00003	3	SWU00003	25
SWT00004	6	SWU00004	70
SWT00005	19	SWU00005	58
SWT00006	7	SWU00006	57
SWT00007	10	SWU00007	57
SWT00008	45	SWU00008	296
SWT00009	22	SWU00009	296
SWT00010	117	SWU00010	256
SWT00011	1	SWU00011	8

TSP (SWT)		USP (SWU)	
spammer_ID	num_spam	spammer_ID	num_spam
SWT00012	9	SWU00012	49
SWT00013	5	SWU00013	6
SWT00014	9	SWU00014	3
SWT00015	1	SWU00015	28
SWT00016	58	SWU00016	11
SWT00017	10	SWU00017	1
SWT00018	105	SWU00018	30
SWT00019	239	SWU00019	49
SWT00020	3	SWU00020	3
SWT00021	4	SWU00021	10
SWT00022	15	SWU00022	1
SWT00023	12	SWU00023	16
SWT00024	2	SWU00024	12
SWT00025	35	SWU00025	18
SWT00026	211	SWU00026	13
SWT00027	263	SWU00027	4
SWT00028	16	SWU00028	34
SWT00029	32	SWU00029	2
SWT00030	1	SWU00030	4
SWT00031	13	SWU00031	81
SWT00032	5	SWU00032	114
SWT00033	25	SWU00033	5
SWT00034	1	SWU00034	2
SWT00035	2	SWU00035	289
SWT00036	4	SWU00036	2
SWT00037	65	SWU00037	19
SWT00038	776	SWU00038	29
SWT00039	6	SWU00039	49
SWT00040	1	SWU00040	9
SWT00041	6	SWU00041	9

<b>TSP (SWT)</b>		<b>USP (SWU)</b>	
<b>spammer_ID</b>	<b>num_spam</b>	<b>spammer_ID</b>	<b>num_spam</b>
SWT00042	1	SWU00042	65
SWT00043	1	SWU00043	8
SWT00044	3	SWU00044	113
SWT00045	36	SWU00045	125
SWT00046	7	SWU00046	43
SWT00047	4	SWU00047	42
SWT00048	1	SWU00048	1
SWT00049	39	SWU00049	403
SWT00050	9	SWU00050	24
SWT00051	1	SWU00051	132
SWT00052	22	SWU00052	14
SWT00053	4	SWU00053	62
SWT00054	13	SWU00055	41
SWT00055	3	SWU00056	1
		SWU00057	18
		SWU00058	3
		SWU00059	38
		SWU00060	95
		SWU00061	1085
		SWU00062	17
		SWU00063	8
		SWU00064	280
		SWU00065	248
		SWU00066	1
		SWU00067	9
		SWU00068	3
		SWU00069	9
		SWU00070	5
		SWU00071	11
		SWU00072	89

TSP (SWT)		USP (SWU)	
spammer_ID	num_spam	spammer_ID	num_spam
		SWU00073	12
		SWU00074	5
		SWU00075	16
		SWU00076	56
		SWU00077	3
		SWU00078	4
		SWU00079	334
		SWU00080	54
		SWU00081	13
		SWU00082	5
		SWU00083	9
		SWU00084	57
		SWU00085	18
		SWU00086	4
		SWU00087	13
		SWU00088	1
		SWU00089	1
		SWU00090	29
		SWU00091	107
		SWU00092	41
		SWU00093	30
		SWU00094	2
		SWU00095	41
		SWU00096	63
		SWU00097	1
		SWU00098	30
		SWU00099	2
		SWU00100	5
		SWU00102	4
		SWU00103	21

TSP (SWT)		USP (SWU)	
spammer_ID	num_spam	spammer_ID	num_spam
		SWU00104	39
		SWU00105	116
		SWU00106	177
		SWU00107	68
		SWU00108	18
		SWU00109	72
		SWU00110	56
		SWU00111	127
		SWU00112	19
		SWU00113	95
		SWU00114	2
		SWU00115	43
		SWU00116	1
		SWU00117	13
		SWU00118	27
		SWU00119	10
		SWU00120	13
		SWU00121	21
		SWU00122	25
		SWU00123	1
		SWU00124	24
		SWU00125	3
		SWU00126	14
		SWU00127	69
		SWU00128	121
		SWU00129	54
		SWU00130	4
		SWU00131	126
		SWU00132	1

TSP (SWT)		USP (SWU)	
spammer_ID	num_spam	spammer_ID	num_spam
		SWU00133	1

### .3 Demography & Honeypot Spam Count

Table .10: Spam Demographic Information

Num_spam per user	Gender	Age
59	F	13
	M	13
152	F	14
12	M	14
24	F	15
34	M	15
	F	16
	M	16
12	F	17
	M	17
130	F	18
165	M	18
25	M	19
10	F	19
	F	20
96	M	20
	F	21
51	M	21
42	F	22
30	M	22
74	F	23
	M	23
67	F	24

<b>Num_spam per user</b>	<b>Gender</b>	<b>Age</b>
	M	24
86	F	25
29	M	25
	F	26
45	M	26
59	F	27
54	M	27
65	F	28
10	M	28
363	M	29
39	F	29
27	F	30
206	M	30
6	F	31
17	M	31
	M	32
8	F	32
	F	33
13	M	33
5	F	34
31	M	34
29	F	35
29	M	35
22	F	36
1	M	36
53	F	37
	M	37
113	F	38
87	M	38
23	M	39
56	F	39

<b>Num_spam per user</b>	<b>Gender</b>	<b>Age</b>
35	F	40
84	M	40
4	F	41
	M	41
5	F	42
	M	42
5	F	43
76	M	43
30	F	44
	M	44
	F	45
25	M	45
27	F	46
	M	46
76	F	47
186	M	47
	F	48
88	M	48
5	M	49
	F	49
993	M	50
22	F	50
	F	51
645	M	51
40	F	52
372	M	52
119	F	53
38	M	53
	F	54
102	M	54
512	M	55

<b>Num_spam per user</b>	<b>Gender</b>	<b>Age</b>
2	F	55
436	F	56
56	M	56
1988	M	57
111	F	57
14	M	58
	F	58
	F	59
	M	59
146	F	60
82	M	60
3	F	61
27	M	61
	F	62
	M	62
9	F	63
	M	63
	F	64
	M	64
36	F	65
156	M	65
	F	66
340	M	66
1	M	67
2	F	67
117	M	68
	F	68
21	M	69
159	F	69
51	F	70
490	M	70

<b>Num_spam per user</b>	<b>Gender</b>	<b>Age</b>
25	M	71
13	F	71
12	F	72
309	M	72
1	F	73
12	M	73
300	M	74
	F	74
84	F	75
141	M	75
	F	13
99	M	13
3	F	27
	M	27
	F	34
	M	34
	F	39
32	M	39
44	F	51
	M	51
	F	45
5	M	45
42	F	58
8	M	58
343	F	62
1	M	62
87	F	74
668	M	74
53	F	60
	M	60
	F	74

<b>Num_spam per user</b>	<b>Gender</b>	<b>Age</b>
18	M	74

#### .4 Timeline Counts

Table .11: UTL Table: Hourly Spam Receipt

<b>Hour</b>	<b>am/pm</b>	<b>a.m/p.m</b>	<b>UTL Total</b>
0	1		1
1	8	6	14
2	2	1	3
3	0	0	0
4	1	2	3
5		2	2
6			0
7	10	1	11
8	27		27
9	2	3	5
10	10	7	17
11	8	6	14
12	7	11	18
13	2	5	7
14	9	13	22
15	2	3	5
16	6	5	11
17	9	2	11
18	8	1	9
19	11	4	15
20	3	3	6
21	1	2	3
22	1	1	2
23	1		1

Table .12: TLP Hourly Spam Count

<b>Hour</b>	<b>am/pm</b>	<b>a.m/p.m</b>	<b>TLP Total</b>
0	61	44	105
1	102	96	198
2	113	96	209
3	58	55	113
4	59	38	97
5	70	43	113
6	66	39	105
7	65	33	98
8	67	39	106
9	62	44	106
10	67	51	118
11	53	42	95
12	54	52	106
13	168	108	276
14	129	110	239
15	60	58	118
16	62	60	122
17	76	62	138
18	71	51	122
19	100	175	275
20	85	52	137
21	62	54	116
22	55	44	99
23	103	57	160

Table .13: SWU Hourly Spam Count

<b>Hour</b>	<b>am/pm</b>	<b>a.m/p.m</b>	<b>SWU Total</b>
0	318	6	324
1	358	133	491
2	441	20	461
3	103	36	139
4	143	99	242
5	63	59	122
6	108	69	177
7	82	92	174
8	160	174	334
9	145	204	349
10	364	191	555
11	258	125	383
12	239	149	388
13	545	136	681
14	604	281	885
15	205	103	308
16	188	123	311
17	223	70	293
18	218	223	441
19	253	39	292
20	104	153	257
21	103	98	201
22	263	39	302
23	227	36	263

Table .14: SWT Hourly Spam Count

<b>Hour</b>	<b>am/pm</b>	<b>a.m/p.m</b>	<b>SWT Total</b>
0	204	0	204
1	240	36	276

<b>Hour</b>	<b>am/pm</b>	<b>a.m/p.m</b>	<b>SWT Total</b>
2	225	2	227
3	62	5	67
4	45	4	49
5	61	4	65
6	20	9	29
7	92	9	101
8	28	12	40
9	17	17	34
10	89	24	113
11	78	33	111
12	82	14	96
13	236	27	263
14	160	41	201
15	133	41	174
16	70	34	104
17	60	13	73
18	119	6	125
19	115	16	131
20	92	24	116
21	45	6	51
22	98	1	99
23	59	5	64

# REFERENCES

- [1] K. D. Michael Nieves and V. Y. Pillitteri, “An introduction to information security,” 2017, special Publication 800-12 Revision 1, p. 85.
- [2] C. Griffiths. (2001, Dec) the latest 2023 phishing statistics? [Online]. Available: [aag-it.com/the-latest-phishing-statistics](https://aag-it.com/the-latest-phishing-statistics)
- [3] Kaspersky, “Spam and phishing report in 2023,” <https://securelist.com/spam-phishing-report-2023/112015/>, Kaspersky, Tech. Rep., 2023.
- [4] T. Micro. (2021) Types of phishing. [Online]. Available: [https://www.trendmicro.com/ense/what\\_is\\_phishing/types\\_phishing.html](https://www.trendmicro.com/ense/what_is_phishing/types_phishing.html)
- [5] P. Network. (2023, Jun) What is a data leak? [Online]. Available: <https://www.paloaltonetworks.com/cyberpedia/data-leak>
- [6] S. Henry, “CrowdStrike’s work with the democratic national committee: Setting the record straight,” <https://www.crowdstrike.com/en-us/blog/bears-midst-intrusion-democratic-national-committee/>, CrowdStrike, Tech. Rep., June 2020, blog Report.
- [7] L. H. Newman. (2023, January) What twitter’s 200 million-user email leak actually means. Date accessed not recorded. [Online]. Available: <https://www.wired.com/story/twitter-leak-200-million-user-email-addresses/>
- [8] E. M. T. G. K. Scarfone, “Guide to protecting the confidentiality of personally identifiable information (pii),” National Institute of Standards and Technology (NIST), Tech. Rep., 2020, special Publication 800-122 Revision p. 8.
- [9] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, “Phishing attacks: A recent comprehensive study and a new anatomy,” *Frontiers in Computer Science*, vol. 3, p. 563060, 2021.

- [10] C. Casey. (April 29, 2021, Dec) The dirty dozen: the 12 most costly phishing attack examples. [Online]. Available: <https://www.thesslstore.com/blog/the-dirty-dozen-the-12-most-costly-phishing-attack-examples/>
- [11] A. T. Tunggal. (Dec) What is data leak? stop giving cybercriminals free access. [Online]. Available: [https://www.upguard.com/blog/data-leak",year="2023](https://www.upguard.com/blog/data-leak)
- [12] R. Mansoor, N. D. Jayasinghe, and M. M. A. Muslam, "A comprehensive review on email spam classification using machine learning algorithms," in *2021 International Conference on Information Networking (ICOIN)*. IEEE, 2021, pp. 327–332.
- [13] R. Ávila, R. Khoury, R. Khoury, and F. Petrillo, "Use of security logs for data leak detection: a systematic literature review," *Security and communication networks*, vol. 2021, no. 1, p. 6615899, 2021.
- [14] N. D. Gaikwad and D. Bhosle, "Review on data leakage detection & data prevention techniques," *International Journal of Science and Research (IJSR)*, 2017.
- [15] F. El Mendili, M. Fattah, N. Berros, and et al., "Enhancing detection of malicious profiles and spam tweets with an automated honeypot framework powered by deep learning," *International Journal of Information Security*, vol. 23, pp. 1359–1388, 2024. [Online]. Available: <https://doi.org/10.1007/s10207-023-00796-7>
- [16] A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti, and M. Alazab, "A comprehensive survey for intelligent spam email detection," *Ieee Access*, vol. 7, pp. 168 261–168 295, 2019.
- [17] T. Burdon, "The role of online marketplaces in enhancing consumer protection," *OECD*, 2021.
- [18] Office of the Privacy Commissioner of Canada. Pipedata requirements in brief. [Online]. Available: [https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda\\_brief/](https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/)
- [19] National Institute of Standards and Technology (NIST), "Nist privacy framework: A tool for improving privacy through enterprise risk management," National Institute of Standards and Technology, Tech. Rep., 2023, accessed: 2023-12-13. [Online]. Available: <https://www.nist.gov/privacy-framework#:~:text=The%20NIST%20Privacy%20Framework%20is,services%20while%20protecting%20individuals'%20privacy>

- [20] N. I. of Standards and T. (NIST). Personally identifiable information (pii). [Online]. Available: [https://csrc.nist.gov/glossary/term/personally\\_identifiable\\_information](https://csrc.nist.gov/glossary/term/personally_identifiable_information)
- [21] R. Verma, V. Gautam, C. P. Yadav, I. Gupta, and A. K. Singh, “A survey on data leakage detection and prevention,” 2020. [Online]. Available: <https://ssrn.com/abstract=3603736>
- [22] R. D. J. M. C. S. D. I. T. D. I. . R. J. J. Herrera Montano I., García Aranda J J., “Survey of techniques on data leakage protection and methods to address the insider threat,” vol. 25, no. 6,, 2022, pp. 4289–4302.
- [23] S. Alneyadi, E. Sithiraseenan, and V. Muthukkumarasamy, “Detecting data semantic: a data leakage prevention approach,” in *2015 IEEE Trustcom/BigDataSE/ISPA*, vol. 1. IEEE, 2015, pp. 910–917.
- [24] K. V. Samarthrao and V. M. Rohokale, “Enhancement of email spam detection using improved deep learning algorithms for cyber security,” *Journal of Computer Security*, vol. 30, no. 2, pp. 231–264, 2022.
- [25] S. Srinivasan, V. Ravi, V. Sowmya, M. Krichen, D. B. Nouredine, S. Anivilla, and K. Soman, “Deep convolutional neural network based image spam classification,” in *2020 6th conference on data science and machine learning applications (CDMA)*. IEEE, 2020, pp. 112–117.
- [26] H. T. Renuka, D. Karthika, M. R. Chakkaravarthi, and P. L. Surya, “Spam classification based on supervised learning using machine learning techniques,” in *2011 International Conference on Process Automation, Control and Computing*, 2011, pp. 1–7.
- [27] M. Erdélyi, A. Garzó, and A. A. Benczúr, “Web spam classification: a few features worth more,” in *Proceedings of the 2011 Joint WICOW/AIRWeb Workshop on Web Quality*, ser. WebQuality ’11. New York, NY, USA: Association for Computing Machinery, 2011, p. 27–34. [Online]. Available: <https://doi.org/10.1145/1964114.1964121>
- [28] N. F. Shah, e. P. K. Kumar, Pramod”, and M. W. Y. M. B. Sahoo, Manmath Narayanand Murugappan, “A comparative analysis of various spam classifications,” in *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*. Singapore: Springer Singapore, 2018, pp. 265–271.
- [29] S. K. Nayak and A. C. Ojha, “Data leakage detection and prevention: Review and research directions,” *Machine Learning and Information Processing: Proceedings of ICMLIP 2019*, pp. 203–212, 2020.

- [30] A. A. Naeem, “Honeypots: Concepts, approaches and challenges,” 2021. [Online]. Available: <https://hal.science/hal-03324407/document>
- [31] R. Alharthi, A. Alhothali, and K. Moria, “Detecting and characterizing arab spammers campaigns in twitter,” *Procedia Computer Science*, vol. 163, pp. 248–256, 2019.
- [32] K. Lee, J. Caverlee, and S. Webb, “The social honeypot project: protecting online communities from spammers,” in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 1139–1140.
- [33] Y. Zhang, H. Zhang, X. Yuan, and N.-F. Tzeng, “Pseudo-honeypot: Toward efficient and scalable spam sniffer,” in *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2019, pp. 435–446.
- [34] A. Bhowmick and e. A. D. S. S. K. Hazarika, Shyamanta M.”, “E-mail spam filtering: A review of techniques and trends,” in *Advances in Electronics, Communication and Computing*. Singapore: Springer Singapore, 2018, pp. 583–590.
- [35] Y. Shapira, B. Shapira, and A. Shabtai, “Content-based data leakage detection using extended fingerprinting,” *arXiv preprint arXiv:1302.2028*, 2013.
- [36] C. Bhandari and S. N. Kini, “A survey paper on data leak detection using semi honest provider framework,” *International Journal of Science and Research*, pp. 2319–7064, 2015.
- [37] P. Zilberman, S. Dolev, G. Katz, Y. Elovici, and A. Shabtai, “Analyzing group communication for preventing data leakage via email,” in *Proceedings of 2011 IEEE international conference on intelligence and security informatics*. IEEE, 2011, pp. 37–41.
- [38] I.-H. Hann, K.-L. Hui, Y.-L. Lai, S.-Y. T. Lee, and I. P. Png, “Who gets spammed?” *Communications of the ACM*, vol. 49, no. 10, pp. 83–87, 2006.
- [39] R. S. Miani, D. Oliveira, K. J. B. Park, and B. B. Zarpelao, “An empirical study of factors affecting the rate of spam,” in *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC, 2018, pp. 239–252.
- [40] A. Almaatouq, E. Shmueli, M. Nouh, A. Alabdulkareem, V. K. Singh, M. Alsaleh, A. Alarifi, A. Alfaris, and A. S. Pentland, “If it looks like a spammer and behaves like a spammer, it must be a spammer: analysis and detection of microblogging spam accounts,” *International Journal of Information Security*, vol. 15, pp. 475–491, 2016.

- [41] E. Ezpeleta, U. Zurutuza, and J. M. G. Hidalgo, “A study of the personalization of spam content using facebook public information,” *Logic Journal of the IGPL*, vol. 25, no. 1, pp. 30–41, 2017.
- [42] K. Jamal, M. Maier, and S. Sunder, “Privacy in e-commerce: Development of reporting standards, disclosure, and assurance services in an unregulated market,” *Journal of Accounting Research*, vol. 41, no. 2, pp. 285–309, 2003.
- [43] M. T. Ribeiro, P. H. C. Guerra, L. Vilela, A. Veloso, D. Guedes, W. Meira Jr, M. H. Chaves, K. Steding-Jessen, and C. Hoepers, “Spam detection using web page content: a new battleground,” in *Proceedings of the 8th annual collaboration, electronic messaging, anti-abuse and spam conference*, 2011, pp. 83–91.
- [44] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, “Detecting and characterizing social spam campaigns,” in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement (IMC’10)*. ACM, 2010, pp. 35–47.
- [45] T-mobile. [Online]. Available: <https://https://www.t-mobile.com/>
- [46] Kaspersky. (2022) Kaspersky spam and phishing report 2022. Accessed December 6, 2024. [Online]. Available: <https://securelist.com/spam-and-phishing-in-2022/>
- [47] A. VPN. (2023) 44
- [48] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage, “Spamalytics: An empirical analysis of spam marketing conversion,” in *Proceedings of the 15th ACM Conference on Computer and Communications Security (CCS)*. ACM, 2008, pp. 3–14.
- [49] K. Levchenko, A. Pitsillidis, N. Chachra, B. Enright, N. Feamster, and G. M. Voelker, “Click trajectories: End-to-end analysis of the spam value chain,” in *IEEE Symposium on Security and Privacy*. IEEE, 2011, pp. 431–446.
- [50] National Initiative for the Care of the Elderly (NICE), “Internet scams targeting older adults,” 2021, accessed: 2025-03-24. [Online]. Available: <https://www.nicenet.ca/articles/internet-scams-targeting-seniors>
- [51] T. Knutson, “Romance scams surged for seniors during pandemic,” *Forbes*, September 2021, accessed: 2025-03-24. [Online]. Available: <https://www.forbes.com/sites/tedknutson/2021/09/23/romance-scams-surged-for-seniors-during-pandemic/>

[52] G. of Canada, “Canada’s anti-spam legislation (casl),” 2014, accessed: 2024-02-21. [Online].  
Available: <https://crtc.gc.ca/eng/internet/anti.htm>