

## ABSTRACT

Title of dissertation:      **ACTIVE DATA COLLECTION TECHNIQUES  
TO UNDERSTAND ONLINE SCAMMERS  
AND CYBERCRIMINALS**

Young Sam Park, Doctor of Philosophy, 2016

Dissertation directed by:  **Professor Elaine Shi  
Department of Computer Science**

Nigerian scam, also known as advance fee fraud or 419 scam, is a prevalent form of online fraudulent activity that causes financial loss to individuals and businesses. Nigerian scam has evolved from simple non-targeted email messages to more sophisticated scams targeted at users of classifieds, dating and other websites. Even though such scams are observed and reported by users frequently, the community’s understanding of Nigerian scams is limited since the scammers operate “underground”.

To better understand the underground Nigerian scam ecosystem and seek effective methods to deter Nigerian scam and cybercrime in general, we conduct a series of active and passive measurement studies. Relying upon the analysis and insight gained from the measurement studies, we make four contributions: (1) *we analyze the taxonomy of Nigerian scam and derive long-term trends in scams*; (2) *we provide an insight on Nigerian scam and cybercrime ecosystems and their underground operation*; (3) *we propose a payment intervention as a potential deterrent*

*to cybercrime operation in general and evaluate its effectiveness; and (4) we offer active and passive measurement tools and techniques that enable in-depth analysis of cybercrime ecosystems and deterrence on them.*

We first created and analyze a repository of more than two hundred thousand user-reported scam emails, stretching from 2006 to 2014, from four major scam reporting websites. We select ten most commonly observed scam categories and tag 2,000 scam emails randomly selected from our repository. Based upon the manually tagged dataset, we train a machine learning classifier and cluster all scam emails in the repository. From the clustering result, we find a strong and sustained upward trend for targeted scams and downward trend for non-targeted scams.

We then focus on two types of targeted scams: sales scams and rental scams targeted users on Craigslist. We built an automated scam data collection system and gathered large-scale sales scam emails. Using the system we posted honey-pot ads on Craigslist and conversed automatically with the scammers. Through the email conversation, the system obtained additional confirmation of likely scam activities and collected additional information such as IP addresses and shipping addresses. Our analysis revealed that around 10 groups were responsible for nearly half of the over 13,000 total scam attempts we received. These groups used IP addresses and shipping addresses in both Nigeria and the U.S. We also crawled rental ads on Craigslist, identified rental scam ads amongst the large number of benign ads and conversed with the potential scammers. Through in-depth analysis of the rental scams, we found seven major scam campaigns employing various operations and monetization methods. We also found that unlike sales scammers, most rental

scammers were in the U.S.

The large-scale scam data and in-depth analysis provide useful insights on how to design effective deterrence techniques against cybercrime in general. We study underground DDoS-for-hire services, also known as *booters*, and measure the effectiveness of undermining a payment system of DDoS Services. Our analysis shows that the payment intervention can have the desired effect of limiting cybercriminals' ability and increasing the risk of accepting payments.

ACTIVE DATA COLLECTION TECHNIQUES  
TO UNDERSTAND ONLINE SCAMMERS  
AND CYBERCRIMINALS

by

Young Sam Park

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2016

Advisory Committee:

Professor Elaine Shi, Chair/Advisor

Professor Damon McCoy, Co-Advisor

Dr. Markus Jakobsson

Professor Michelle Mazurek

Professor Hal Daume

Professor David Maimon

© Copyright by  
Young Sam Park  
2016

## Dedication

To my wife, Soyoung Hong  
and my children, Jisoo and Jiwon  
for all their unconditional love and support

## Acknowledgments

I am truly grateful to all those who helped me complete this thesis. If I forgot to mention you, I apologize, and please know that I really appreciate your help.

I first would like to thank my advisors, Dr. Elaine Shi and Dr. Damon McCoy, for their support throughout my whole graduate research. They inspired me to pursue an academic career doing research on cyber-security and cyber-criminal. I also would like to thank to Dr. Markus Jakobsson, who gave me advice on research and opportunities to co-work with his great company in cybersecurity area. I would like to thank to the committee members, Dr. Hal Daume, Dr. Davie Maimon and Dr. Michelle Mazurek for their time and service.

I am thankful to Jackie Jones, Mohammad Karami, Yali Li and Kai Wang for their great work and friendship. Without their help, this would not have been possible to complete this thesis.

Finally, I am grateful for my family, parents and friends for their true support throughout my life.

# Table of Contents

List of Tables	viii
List of Figures	x
1 Introduction	1
1.1 Outline	8
2 Trends In Nigerian Scams	9
2.1 Overview	9
2.2 Data collection	9
2.2.1 Scam reporting sites	10
2.2.2 FBI/IC3 annual report	10
2.3 Taxonomy of scam emails	11
2.3.1 Non-targeted scams	11
2.3.2 Targeted scams	14
2.3.3 Both non-targeted and targeted scams	16
2.4 Scam classification	17
2.5 Scam trends	19
2.5.1 Overall trends in scams	21
2.5.2 Targeted vs Non-targeted scams	22
2.5.3 Scams on the rise	23
2.5.4 Scams in the decline	25
2.5.5 Analysis	26
2.6 Conclusion	28
3 Sales Scams on Craigslist	29
3.1 Overview	29
3.2 Data Collection Methodology	30
3.2.1 Magnetic honeypot posts.	30
3.2.2 Automated communication with scammers.	31
3.2.3 IP address collection.	33
3.2.4 Ethics	33



3.3	Dataset	36
3.3.1	Overview and terminology.	36
3.3.2	Magnetic Honeypot Advertisements.	38
3.3.3	Collected emails and threads.	38
3.4	Analysis of scammers' IP addresses	40
3.4.1	IP geolocation.	40
3.4.2	IP blacklist.	42
3.5	Analysis of scammers' email accounts	44
3.5.1	Source, reply-to address discrepancy, and email account reuse.	44
3.5.2	Email service provider.	46
3.5.3	Sample bad email addresses.	47
3.6	Shipping Addresses and Phone Numbers	48
3.6.1	Shipping Addresses.	48
3.6.2	Phone Numbers.	49
3.7	Scam Patterns	49
3.7.1	When do scammers work?	50
3.7.2	How fast do scammers respond?	51
3.7.3	Do product category and price affect scammers' response rate?	52
3.8	Level of Automation	55
3.8.1	Signs of automation.	55
3.8.2	Signs of manual labor.	56
3.9	Classification	59
3.9.1	Conservative Classification Strategy.	59
3.9.2	Top 10 Groups.	60
3.9.3	More Aggressive Grouping Strategy and Findings.	64
3.9.4	Classification Summary.	65
3.10	Literature Review	65
3.11	Discussion	67
3.12	Conclusion	70
4	Rental scams on Craigslist	71
4.1	Overview	71
4.2	Data Sets	72
4.2.1	Rental Listing Crawling	72
4.2.2	Campaign Identification	73
4.2.3	Campaign Expansion Phase: Latitudinal	75
4.2.3.1	Human-generated scam signatures.	75
4.2.4	Campaign Expansion Phase: Longitudinal	75
4.2.4.1	Automated conversation engine.	76
4.2.4.2	Ethics.	78
4.2.5	Campaign Summaries	79
4.3	Analysis of Scam Campaigns	80
4.3.1	Credit Report Scams	80
4.3.1.1	Data collection.	81
4.3.1.2	Dataset sanity check.	81

4.3.1.3	In-depth analysis. . . . .	82
4.3.2	Clone Scam . . . . .	85
4.3.2.1	Data collection. . . . .	85
4.3.2.2	In-depth analysis of confirmed scams. . . . .	86
4.3.3	Realtor service scam . . . . .	88
4.3.3.1	Data Collection. . . . .	89
4.3.3.2	American Standard Online. . . . .	90
4.3.3.3	New Line Equity. . . . .	91
4.3.3.4	Search Rent To Own. . . . .	92
4.4	Flagged Ads Analysis . . . . .	93
4.5	Discussion . . . . .	95
4.5.1	Potential detection and conversation limitations. . . . .	95
4.5.2	Similarities and differences with previous study. . . . .	95
4.6	Related Works . . . . .	96
4.6.1	Advanced fee fraud. . . . .	96
4.6.2	Underground studies. . . . .	97
4.7	Conclusion . . . . .	98
5	Understanding and Deterring the Business of DDoS Services . . . . .	99
5.1	Overview . . . . .	99
5.2	Background . . . . .	100
5.2.1	Distributed Denial-of-Service . . . . .	100
5.2.2	Booter Services . . . . .	101
5.2.3	Ethical Framework . . . . .	102
5.3	Related Work . . . . .	105
5.4	Inside View of Booters . . . . .	107
5.4.1	Data Sets . . . . .	107
5.4.2	Subscribers . . . . .	109
5.4.3	Revenue . . . . .	110
5.4.4	Attacks . . . . .	111
5.5	Attack Infrastructure . . . . .	111
5.5.1	Data Set . . . . .	112
5.5.2	Frontend Servers . . . . .	113
5.5.3	Attack Servers . . . . .	114
5.5.4	Attack Techniques . . . . .	116
5.5.5	Amplifiers . . . . .	117
5.5.6	Amplifier Location . . . . .	117
5.5.7	Amplifiers Churn . . . . .	118
5.5.8	Amplification Factor . . . . .	119
5.6	Payment Intervention . . . . .	120
5.6.1	Payment Ecosystem . . . . .	120
5.6.2	Usage Pattern of PayPal Accounts . . . . .	122
5.6.3	Booters' Status . . . . .	125
5.6.4	Qualitative Assessments . . . . .	127
5.6.5	Booter Response . . . . .	128

5.7	Discussion and Future Work . . . . .	129
5.8	Conclusion . . . . .	132
A	Scam Examples . . . . .	133
A.1	Example fake payment scam emails . . . . .	133
A.2	Example rental scam ads . . . . .	137
A.3	Example rental scam emails . . . . .	138
A.4	Misc. . . . .	140
	Bibliography . . . . .	141

## List of Tables

2.1	Dataset summary . . . . .	10
2.2	<b>Scam email taxonomy.</b> 9 most prevalent scam categories and their estimated fractions based on manual tagging and Support Vector Machine (SVM) [1] classifier. . . . .	12
2.3	Most frequently reported scams in FBI/IC3 annual reports. . . . .	26
3.1	Pre-defined keywords and example sentence generated from the keywords. . . . .	32
3.2	Summary of experimental result . . . . .	36
3.3	Terminology . . . . .	37
3.4	IP addresses blacklisted by <i>Project Honey Pot</i> . . . . .	43
3.5	Top 20 IP addresses by the number of times observed. . . . .	43
3.6	Source and reply-to address discrepancy. . . . .	45
3.7	Number of source addresses mapped to a single reply-to addresses. . . . .	46
3.8	Distribution of email service providers. . . . .	48
3.9	Shipping addresses and phone numbers. . . . .	50
3.10	P-values of normal distribution tests . . . . .	53
3.11	P-values of pairwise KruskalWallis tests . . . . .	54
3.12	Analysis of scam automation. . . . .	58
3.13	Example Interarrival Times for Burst Traffic from a Single Email Address. . . . .	59
3.14	Top 10 Groups. . . . .	62
3.15	Top 10 Group Durations. . . . .	63
4.1	Dataset summary. . . . .	73
4.2	Example sentences used in first emails generated by the conversation engine. . . . .	78
4.3	Major rental scam campaigns. . . . .	80
4.4	Credit report scam campaigns. . . . .	83
4.5	Example inter-arrival time for burst email responses of CreditReport.Yahoo. . . . .	84
4.6	Top 3 clone scam groups. . . . .	88
4.7	Realtor service with advance fee campaigns. . . . .	90

4.8	Flagged ads categorization. . . . .	94
5.1	Average number of requests per second and average bandwidth consumed in kbps for each amplifier. . . . .	104
5.2	Summary of Asylum Stresser and Lizard Stesser leaked databases and scraped VDO reported data. . . . .	107
5.3	List of booter services we measured, the attack types offered, and the cost of the least expensive one-month subscription. . . . .	114
5.4	Spoofing enabled VPS services. . . . .	115
5.5	Number of total amplification servers and percentage of overlap with amplification servers used by other booters. . . . .	116
5.6	Top country locations and autonomous systems for amplifiers. . . . .	118
5.7	Number of PayPal accounts used by monitored booters before and after the intervention. . . . .	123

## List of Figures

2.1	Precision/recall and receiver operating characteristic (ROC) curve of SVM classifier. . . . .	20
2.2	Number of scam complaints reported to FBI/IC3. [2] . . . . .	22
2.3	Number and fraction of scam emails reported. . . . .	23
2.4	Targeted vs non-targeted scams. . . . .	24
2.5	Scams on the rise and in the declines. . . . .	27
3.1	Automated scam data collection using magnetic honeypot ads. . . . .	30
3.2	Example 419 scam thread . . . . .	34
3.3	Distribution of magnetic honeypot ads over ad posting time. . . . .	39
3.4	IP Geolocation of scammers. . . . .	41
3.5	Cumulative distribution of IP addresses over number of subnets. . . . .	42
3.6	Distribution of email account reuse count. . . . .	45
3.7	Number of email threads, source email addresses and reply-to email addresses. . . . .	46
3.8	Received time of scam responses. . . . .	52
3.9	Response time of scam emails. . . . .	53
3.10	Number of first scam responses per effective ad. . . . .	55
3.11	Scammer group by number of threads. . . . .	60
3.12	Emails Per Day - Top 10 Groups. . . . .	64
4.1	Time taken to flag scam clone ads. . . . .	94
5.1	Structure of booter services. . . . .	102
5.2	IP churn of amplifiers. . . . .	119
5.3	PayPal account usage over time. . . . .	124
5.4	Lifespans of PayPal accounts before and after the intervention. . . . .	126
5.5	Status of booters over time. . . . .	127
A.1	Recurring fake payment scam emails. . . . .	133
A.2	Sample emails with belligerent tones. . . . .	134
A.3	Sample broken subject and body lines. . . . .	134
A.4	Sample recurring emails bursts. . . . .	135

A.5	Sample emails indicating manual operation. . . . .	135
A.6	Sample emails invoking God. . . . .	135
A.7	Sample emails with CAPITALIZED text. . . . .	136
A.8	Sample themes in emails. . . . .	136
A.9	Example ad titles with sophisticated templates used by CreditRe- port_Yahoo campaign. . . . .	137
A.10	Example rental clone scam ad. . . . .	137
A.11	Example rental scam emails. . . . .	138
A.12	Example rent application template. . . . .	139
A.13	Example credit report scam email. . . . .	139
A.14	Craigslist phone verification scam. . . . .	140

## Chapter 1: Introduction

*Nigerian scams*, also commonly referred to as *Advance fee fraud* or *419 scams*<sup>1</sup>, is a prevalent form of online fraud that not only causes financial loss to individuals and businesses alike [3], but also can bring emotional or psychological damage to victim users [4]. An estimation of global losses to Nigerian scams in 2005 is more than 3 billion dollars [5]. Total financial loss from all referred cases of Internet scam in the U.S. was also increased from 17.8 million dollars in 2001 to 800 million dollars in 2014 [2]. While not all Nigerian scams nowadays originate from Nigeria, a significant proportion of them still do, and the name Nigerian scam is used to broadly refer to a large class of online fraud. This scam was originally mostly non-targeted and delivered via email spam. However, today there are more sophisticated targeted versions of this scam that are directed at users of classifieds, jobs and dating sites.

Despite the ubiquitous presence of online Nigerian scams, we currently lack a solid understanding of the online Nigerian scam ecosystem and the different techniques scammers use to deceive and profit off their victims. Many online websites, such as Craigslist, filter out scam postings to protect its legitimate users. For exam-

---

<sup>1</sup>We use all three terms interchangeably in this thesis.



ple, Craigslist has many safeguards in place to prevent scam postings, such as requiring phone number verification for a Craigslist account to prevent scammers from registering large numbers of Craigslist accounts and posting fraudulent advertisements, blocking suspicious IP addresses and accounts, and removing advertisements containing suspicious content. While most efforts to mitigate this problem focus on filtering suspicious posts or emails, this is only the visible part of a well honed set of scams and infrastructure established to extract money from their marks. An end-to-end understanding of a scam and its structural dependencies (message postings, email accounts, location of scammers, support companies, automated tools and payment methods) is often a crucial first step towards identifying potential weakness along the chain. In particular, this “understanding to deterrence” trajectory has resulted in suggesting weak points for disrupting other domain-specific threats, such as payment processing in the counterfeit software and pharmacy spam domain [6, 7].

To better understand the underground Nigerian scam ecosystem and seek effective methods to deter Nigerian scam and cybercrime in general, we conduct a series of active and passive measurement studies. Relying upon the analysis and insight gained from the measurement studies, we make four contributions: (1) *we analyze the taxonomy of Nigerian scam and derive long-term trends in scams*; (2) *we provide an insight on Nigerian scam and cybercrime ecosystems and their underground operation*; (3) *we propose a payment intervention as a potential deterrent to cybercrime operation in general and evaluate its effectiveness*; and (4) *we offer active and passive measurement tools and techniques that enable in-depth analysis of cybercrime ecosystems and deterrence on them*.

**Scam taxonomy and trends.** The first step toward better understanding of Nigerian scams is to create and analyze a large-scale Nigerian scam email dataset. We collected more than two hundred thousand user-reported scam emails over the recent nine years, stretching from 2006 to 2014 from four major scam reporting websites. Since there is no de facto standard of scam taxonomy, we built a list of major scam categories frequently observed and reported from users. We select ten most commonly observed scam categories along with seven additional subcategories, and analyze the underlying monetization methodology of each scam category.

In order to build a ground-truth dataset for scam email categorization, we manually tag 2,000 scam emails randomly selected from our repository. Based upon the manually tagged dataset, we train a machine learning classifier and then cluster all scam emails in the repository. From the clustering result, we find a strong and sustained upward trend for targeted scams and downward trend for non-targeted scams. We then look at individual trends for different types of scams, and the analysis of our scam repository show a dramatic rise in *Romance*, *Authority* and *BEC/Sales* scams while *Money transfer*, *Lottery* and *Phishing* scams are in decline. Our finding clearly suggests that scammers are Nigerian scammers are evolving from traditional spam-like scams to more sophisticated scams targeting at specific users.

**Understanding targeted sales scam.** [8] We focus on sales scam or fake payment scam<sup>2</sup>, a typical form of targeted scams, on Craigslist, one of the most popular online market websites whose monthly visitors are over 60 million in the U.S. alone<sup>3</sup>.

---

<sup>2</sup>We use these terms interchangeably in this thesis.

<sup>3</sup><http://www.craigslist.org/about/factsheet>

We present an in-depth measurement study of such scam activities. Through this measurement study, we aim to better *understand the underground economy of Nigerian scams*, and *seek effective intervention points*. In particular, we seek to address questions such as the following: “Where are scammers located?”, “How do scam factories operate?”, “How do scammers monetize their scam activities?” and “What features can we use to distinguish a scam email from a legitimate email?”

In order to better understand sales scams on Craigslist, we posted magnetic honeypot advertisements – designed to attract scammers but repel legitimate users. We received and replied to scam emails resulting from our advertisements, and analyzed the emails. For quantitative analysis of scams, we build an automated data collection system which posts advertisements, collects scam emails and interacts with scammers by sending out a response to the received scam emails. We also collect IP addresses of scammers to explicitly confirm geolocation of the scammers. We perform various analysis of the massively collected dataset to better understand how scammers work. We also cluster observed scammers into groups based on a few key factors such as email addresses, shipping address, phone number and email payload.

Our analysis reveals that these types of Nigerian scams are highly prevalent as our magnetic honeypot advertisements on average received 9.6 scam replies. The most enlightening result of our analysis is that *about 50% of the scam attempts observed can be linked back to the top 10 groups*. These groups are targeting advertisements spread over many classes of goods and geographic regions of Craigslist. In addition, our analysis reveals that many of the initial scam messages are automated

and arrive from a large number of email address that are quickly abandoned. However, most of these initial messages contain a different reply-to address to a smaller set of longer lived email accounts. We also find that 23% of the shipping addresses are located in the United States, although most of the IP addresses and shipping addresses are located in Nigeria. This indicates there are likely either accomplices or reshipping operations being used. Our analysis of the content of the messages shows certain occurrences of words such as, God, overseas military personnel, and capital letters that might be used to help filter these messages.

From this analysis we find several potential intervention points. Our analysis of the message content indicates that message filtering could be improved by looking for combinations of these pattern such as a reply-to address that does not match the sender's address, usage of these uncommon phrases, and identification and blacklisting of these more stable and long-lived secondary accounts. Also, shipping addresses might be the starting point for law enforcement investigations. Along these same lines the fact that only ten groups of scammers accounted for nearly half of the scams we received indicates that it might be possible to target and disrupt these groups, greatly reducing the prevalence of this scam.

**Understanding rental scam.** [9] We present a systematic empirical study of the targeted rental scams ecosystem as viewed through the lens of the Craigslist rental section. The initial challenge of this study is detecting rental scam posting amongst the large number of benign rental listings.

In this chapter, we conduct the first systematic empirical study of the online

rental scams ecosystem as viewed through the lens of the Craigslist rental section. Our in-depth analysis of these rental scam campaigns allows us to address questions geared at improving our understanding of the supporting infrastructure with the goal of exploring alternate points to undermine this ecosystem, such as: “What are the common underlying scams?”, “Where are these scammers located and what tools do they use?”, “How effective are current defences?”, “What payment methods do they use?”. We summarize our contributions and findings below.

By developing several effective detection techniques, we are able to identify several major rental scam campaigns on Craigslist. In addition, we extend Scam-baiter automated conversation engine to automatically contact suspected rental scammers, which enabled us to understand what support infrastructure they used and how they were monetizing their postings. In total we detected about 29K scam listings over the 20 cities we monitored, within a period of 141 days.

We find a diverse set of methods utilized for monetizing the rental scam campaigns we identified. These include attempts to trick people into paying for credit reports and “bait-and-switch” rental listings. When we explored the payment method used, five of the seven major scam campaigns identified used credit cards. Many campaigns also depended on businesses registered in the USA to collect payments. We also find that Craigslist’s filtering methods are currently removing less than half of the rental scam ads we detected.

Our results highlight the fact that scammers are highly customizing their monetization methods to the United States rental market. They also expose new scams and infrastructure that were not encountered in previous studies [10,11]. This differ-

ence highlights the need to understand a wider range of scam domain and suggests potential bottlenecks for many rental scam monetizing strategies at the regulatory and payment layers. For instance, United States regulatory agencies, such as the Federal Trade Commission (FTC) could investigate these companies and levy fines for their deceptive advertising practices. Another potential method of demonetizing these companies might be to alert credit card holder associations, such as Visa or MasterCard, to these merchants' deceptive billing and refund policies.

**Understanding and deterring DDoS service.** [12] We undertake a large scale active and passive measurement study of booter services to understand how they are structured both technologically and economically with the focus of isolating potential weaknesses. We explore booters from three different angles including analysis of leaked and scraped data, measurements of their attack infrastructure and a payment intervention.

Our analysis of leaked and scraped data from three booters—Asylum Stresser, Lizard Stresser and VDO <sup>4</sup>—demonstrates that these services have attracted over 6,000 paid subscribers that have launched over 600,000 attacks. We also find that the majority of booter customers prefer paying via PayPal and that Lizard Stresser, which only accepted Bitcoin, had a minuscule 2% sign-up to paid subscriber conversion rate compared to 15% for Asylum Stresser and 23% for VDO, which both accepted PayPal. By analyzing attack traffic directed at our own servers we are able

---

<sup>4</sup>We assign each booter service a unique three letter code based on their domain name to avoid unintentionally advertising their services. The two exceptions are Asylum Stresser, which ceased operation before our study and Lizard Stresser, which has already been highly publicized.

to characterize the set of amplifiers they use to direct large amounts of traffic at their victims. In order to measure the resilience of their payment infrastructure, we conduct a payment intervention in collaboration with PayPal and the FBI. Our evaluation of the effectiveness of this approach suggests that it is a promising method for reducing the subscriber base of booters.

## 1.1 Outline

The remainder of this dissertation is structured as follows. Chapter 2 presents a taxonomy of Nigerian scams and its trends over the last nine years. In chapter 3, we describe a quantitative measurement study of sales scams on Craigslist. We provide an insight into the underground ecosystem of Nigerian scam business on Craigslist, and discuss potential intervention points. In chapter 4, we present another empirical measurement study of rental scams on Craigslist. In chapter 5, we show a study of DDoS-for-hire service, another form of cybercrime, and present an experimental results of a large-scale payment intervention in collaboration with PayPal.

## Chapter 2: Trends In Nigerian Scams

### 2.1 Overview

This chapter focuses on taxonomy of scam emails collected from various sources and derive a long-term trend in scam emails. We first collect a large-scale scam emails and then analyze what kind of scams are there, what their structures are, how they are related. Based upon the taxonomy analysis, we build a machine learning classifier and cluster the scam emails into major scam categories. Then we further analyze and derive different trends from each scam category. Our analysis show a clear trend that spam-like *non-targeted scams* are decreasing while *targeted scams* with specific victims are getting more prevalent over the last 10 years. We discuss that our analysis can be used to provide the predictions of future scam developments based on historical trends and insight gained through the analysis.

### 2.2 Data collection

In order to analyze the trends in scams, we collect extensive scam emails and its statistics from several sources including scam reporting sites, spam email dataset and government report on internet crimes.



Source	URL	Start date	End date	# Scam emails
Anti Fraud International	antifraudintl.org	02/2007	12/2014	57,338
Scam Warners	scamwarners.com	07/2008	12/2014	54,352
Scamdex	scamdex.com	01/2006	12/2014	75,943
419baiter	419baiter.com	06/2008	02/2012	31,847
<b>Total</b>				<b>219,480</b>

Table 2.1: **Dataset summary.** Summary of scam email dataset between 2006 and 2014.

### 2.2.1 Scam reporting sites

The basis of our analysis of scam trends relies upon a large-scale scam email dataset. We first focus on scam reporting sites where many anonymous users post scam emails that they received. Amongst many scam reporting sites, we select 4 major ones with relatively large scam email database, more than 30K scam emails in each site. We crawl scam emails reported between 2006 and 2014, that is, more than 10K scam emails each year. Overall we collect about 220K scam emails over 4 scam reporting sites. Table 2.1 represents the summary of our scam email dataset.

### 2.2.2 FBI/IC3 annual report

Even though we are able to collect many number of scam emails, it is still unclear that what kind of scams are more threatening to people, that is, generating larger number of victims and financial losses. In this context, we reference *FBI/IC3 annual reports* [2] to see what kind of internet frauds are frequently reported to law enforcement and regulatory agencies, and how much financial losses they result in.

## 2.3 Taxonomy of scam emails

Since there is no de facto standard of scam categorization, we first need to build major scam categories frequently reported from users. There exist a few reports [2, 13] regarding a taxonomy of online fraud. National Fraud Authority [14] also releases annual reports on offline fraud activities. Also scam reporting sites have their own scam categorization rules. However we cannot apply existing categorization directly to our scam dataset primarily because they do not cover all scam categories we find from our dataset. Also, each report and scam reporting site use different scam categories. For example, *anti fraud international* has 20 scam categories while *419baiter* and *scammed.by* do not categorize scam emails at all. Hence we build our own scam categorization as described in Table 2.2. Through the rigid investigation of the collected scam emails and literature survey, we find 9 scam categories (14 if counting subcategories) commonly used or frequently reported to the scam reporting sites. We categorize scam emails based on 1) whom scammers are pretending to be and 2) how scammers try to persuade their victims. Table 2.2 summarize the list of scam categories and brief description of each scam category. We use our own scam categorization throughout this chapter.

### 2.3.1 Non-targeted scams

The topmost categorization rule is whether a scam emails is either *non-targeted*, *targeted* or *both*. Non-targeted scam is a traditional and typical form of email fraud where scammers do not set any designated victims: instead they send out as many

Type	Category	Subcategory	Manual tagging	SVM classification
Non-targeted	Authority	Bank	99 (4.95%)	18,680 (8.58%)
		Government, organization	78 (3.90%)	
		<i>Total</i>	<i>177 (8.85%)</i>	
	<b>Loan</b>		55 (2.75%)	6,428 (2.95%)
	<b>Lottery</b>		215 (10.75%)	24,132 (11.08%)
	Money transfer	Charity, dying person	105 (5.25%)	93,271 (42.83%)
		Business, commodity	185 (9.25%)	
		Next of kin	367 (18.35%)	
		Widow, orphan, refugee	118 (5.90%)	
		Conversation	47 (2.35%)	
		<i>Total</i>	<i>822 (41.10%)</i>	
<b>Phishing</b>		71 (3.55%)	9,368 (4.30%)	
Targeted	<b>Business</b>		54 (2.70%)	5,295 (2.43%)
	<b>Rental</b>		43 (2.15%)	7,228 (3.32%)
	<b>Romance</b>		203 (10.15%)	20,960 (9.63%)
Both	<b>Employment</b>		71 (3.55%)	10,191 (4.68%)
	<b>Sales</b>		11 (0.55%)	-
Etc.	<b>Others</b>		165 (8.25%)	22,192 (10.19%)
	<b>Invalid emails</b>		113 (5.65%)	
	<i>Total</i>		<i>278 (13.90%)</i>	
<b>Total</b>			2000	217,745

Table 2.2: **Scam email taxonomy.** 9 most prevalent scam categories and their estimated fractions based on manual tagging and Support Vector Machine (SVM) [1] classifier.

scam emails as possible to anonymous users, just like spam.

**Authority scams.** In *authority scams*, scammers pretend to be employees at banks (*bank scams*), government agencies or international organizations (*government/organization scams*). The scammers abusively use the names of renowned organizations (e.g., Bank of Africa, FBI and IMF) to gain trust of victims. In bank

scams, scammers offers a pre-loaded ATM cards or charity fund to victims, and require a fee in advance. Similarly scammers in government/organization scams notify their victims that they are able to receive a charity fund and require a fee for the process. Sometimes this type of scams involve threatening (e.g., black money in victim's account) or malware propagation (e.g., attachment containing virus or worm). Authority scam emails mostly look like official emails from the organizations, and the whole processes explained in the emails also seem official work.

**Loan scams.** Scammers in *loan scams* offer loans to victims usually at an attractive interest rate. But the scammers ask an upfront fees for further loan service processing. This kind of scam emails do not look professional and especially the first emails usually contain couple of short sentences.

**Lottery scams.** *Lottery scams* bring an unexpected but happy news to a victim that his email address has been put into a lottery and wins the prize! Scammers usually require a fee in advance for transferring a huge amount of prize. This is one of most typical and prevalent form of un-targeted scams.

**Money transfer scams.** Scammers in *money transfer scams* usually have funds in African countries and want to transfer the funds to victims countries for several purposes. Scammers in *Charity/dying person scams* usually have inheritance of several million in US dollar and ask victims to help in moving the fund for charity business in victims' countries. *Business/commodity* scammers explain that they are looking for a business partner who will help them invest their money or sell their

commodity in victims' countries. *Next of kin scam* is a typical and one of the most prevalent form of fraud. Scammers in this scam categories are usually bankers or attorneys who have access to abandoned accounts of their client who have passed away without any wills or inheritors. They propose to put victim's name as a next of kin and inherit the money. *Widow/orphan/refugee* scammers are usually in unstable countries with internal wars or dictators. They have inheritance from their parents or husbands who have passed away recently, and want to transfer their money out of their countries for the safety. *Conversation* subcategory includes scam emails in the middle of conversation between scammer and victim, so it is hard to classify those emails into a specific subcategory. In all cases, victims are promised to get a certain percentage of the transferred funds in the return of helping the scammers, but the victims are also required to pay an upfront fee for the money transfer process.

### 2.3.2 Targeted scams

In targeted scams, scammers may have a certain context of their potential victims, e.g., the fact that the victim is looking for an apartment on Craigslist or is selling an old iPhone on eBay. Since the scammers are able to exploit this context of the victims, conversations in targeted scams are more natural and plausible.

**Business email compromise scams.** *Business email compromise (BEC) scams* generally target specific companies related with supplies from foreign businesses. In this scam category, scammers can be sellers who present product catalogs with attractive price tags, and sometimes they can be buyers who request product lists

from victim businesses. Since scammers are “foreign” businesses, payments are usually done via wire transfer or other electronic payments. Seller scammers prefer these payment methods since those are easy to perform but hard to reverse. Likewise buyer scammers also prefer those since it is relatively easy to fabricate “fake” payment notification to victims.

**Rental scams.** *Rental scams* may target users who post listing on classified advertisements websites to look for a rent, or also may post “fake” rental listing by themselves to lure the victims. A common methodology of those scammers is to attract victims with low-priced rents and then ask an upfront fee for the rent of first month and security deposit. The scammers often copy an actual rent listing and repost that with much lower rent price, and may ask the victims to look around the house first. But usually a victim is not allowed to enter the house since the home owner (scammer) is on travel with good purpose (e.g., mission trip to African countries). Hence it may be hard for the victim to figure out if the rent list is legit or not. We study more details of rental scams in Chapter 4.

**Romance scams.** *Romance scams* is little bit different from other types of scams in that scammers have to build a relationship with a victim for relatively long time. Once the scammer build a relationship with the victim successfully, then he may ask money for various reasons, e.g., to purchase a flight ticket. Since the initial phase of romance scam is just “normal” conversation, it is relatively hard to determine whether it is scam.

### 2.3.3 Both non-targeted and targeted scams

Certain types of scams can be categorized into both non-targeted and targeted scams.

**Employment scams.** *Employment scams* present in various forms. One typical form of employment scams starts with a decent job offer from a company located out of victim's home country. Then the victim is usually required to pay an upfront fee for documentation process, e.g., visa application. Employment scams can be both non-targeted and targeted: non-targeted employment scams are sent to unspecified email addresses just as in spams, and targeted employment scams start from job listings on classified advertisements websites.

**Phishing scams.** In general, the goal of *Phishing scams* is either to steal victim's private credentials (e.g., password or SSN) or to make victim install malware by spoofing famous companies that hold victim's money or account information (e.g., banks or PayPal). The key trick in this type of scam is links embedded inside emails which lead the victim to scammer's fabricated websites.

**Sales scams.** In *sales scams*, scammers can be either seller or buyer: seller scammer posts a fraudulent ad on classified advertisements websites, and buyer scammer responds to victim's legit ad and makes a fake payment, e.g., fake PayPal payment notification or bogus check. One typical example of sales scams is used car sales scam where a scammer posts a fraudulent ad selling a non-existing car on classified

advertisements websites. We investigate deep into sales scams in Chapter 3.

## 2.4 Scam classification

Since each scam reporting site uses different scam categories, we first try to cluster the scam email dataset we've collected from various sources based on a unified scam categories described in Chapter 2.3.

The first step of scam classification is to establish the ground truth for the classification. We manually inspected 2000 scam email samples selected randomly from our dataset and tagged them based on the scam categories listed in Table 2.2.<sup>1</sup> In some cases, it is not clear to classify a scam email into a scam category due to the similarity between two or more scam categories. For example, a scammer may pretend to be a banker and asks a help from a victim in transferring abandoned money of his client who passed away a few years ago. This scam email can be classified as *Authority - Bank* scam if we focus on the email sender or *Money transfer - Next of kin* scam if we focus on the content of the email. We set a major principle to resolve this kind of ambiguity— putting the highest priority on email contents. Although the example email was sent from a banker, the email is not pretended to be an official email from a bank and the email contents is more about Money transfer scam, we classified it as Money transfer scam. If an ambiguity is still not resolved, we then classify the scam emails into *Others* category. The result of manual tagging is also shown in Table 2.2. We find that about 41% of all scam email samples are *money transfer scams* that is known to be a traditional and typical form of email

---

<sup>1</sup>Note that the manual tagging was done by a single person in our research team.



scams. The second and third largest fractions are *lottery scams* and *romance scams* which account for about 11% and 10% of scam email samples relatively. Both scam categories are also well-known typical types of email scams.

We then implement SVM [15] classifier using Python *scikit-learn* library [16] and train the classifier using 2000 manually tagged email samples. We eliminate basic English stop words and other non-alphabetical characters. But we do not delete numbers and a special character “\$” since those are meaningful in scam emails since they represent a certain amount of money. Hence we replace numbers and a special character \$ with “number\_term” and “dollar\_term” respectively. We then extract TF-IDF [17] features from each sample and randomly select 80% of samples as *train set* and 20% as *test set*. Then our SVM classifier is trained using the train set and its performance is evaluated using the test set. We repeated the SVM classifier evaluation 10 times using different train and test sets, and the evaluation results are presented in terms of *precision/recall* and *ROC curve* as shown in Figure 2.1. For the ease of classification, we do not use subcategory and classify scam emails to category level only. We also merge business email compromise scams and sales scams due to the similarity of terms used in email contents. Although the email contents of BEC scams and sales scams are easily distinguished by human, SVM classifier with TF-IDF features failed in classifying those scam emails.

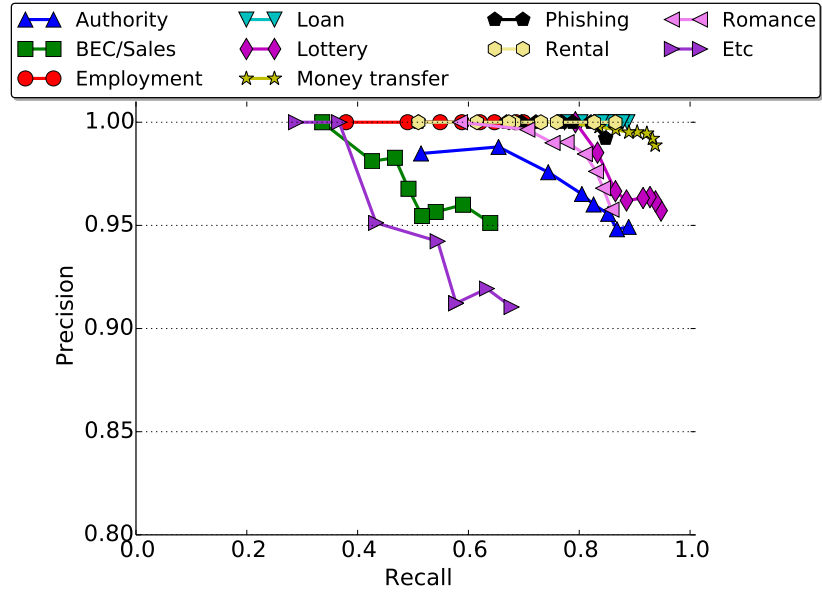
Figure 2.1 shows the performance of the SVM classifier in terms of precision/recall and ROC curve. In Figure 2.1(a), the SVM classifier shows higher than 90% precision with at least 60% recall for all scam categories. If *Etc* and *BEC* scam categories are excluded, the SVM classifier shows over 95% precision with at least

80% recall for the rest scam categories. According to the receiver operating characteristic curve in Figure 2.1(b), the SVM classifier shows over 90% true positive rate with lower than 5% false positive rate except Etc scam category. Since Etc scam category covers various scam types that account for small portions, and this scam category is not our interest. Although our SVM classifier may not show the minimal false positives and false negatives, it can be seen that the SVM classifier has reasonably good enough performance to show an overall trends in scams.

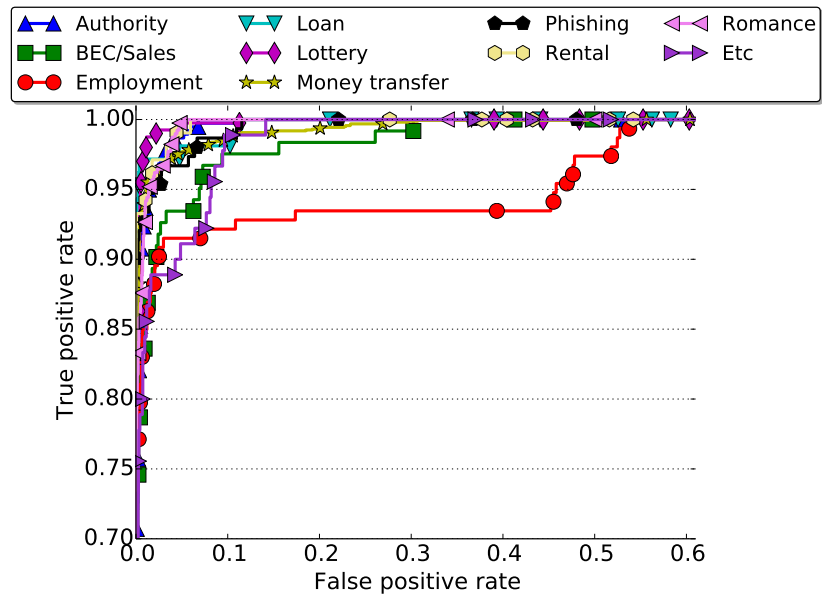
We then classify the rest scam email dataset using the SVM classifier already trained as described above. The overall classification result is presented in Table 2.2. *Money transfer* scam accounts for the largest portion of over scam emails in both results, 41% in manual tagging and 43% in SVM classification. Within Money transfer scam category, a typical Nigerian scam called *Next of kin* forms 18%, and *Widow, orphan, refugee* scams make up about 6% of all scam emails. The second largest scam category except Etc is *Lottery* scam in both results, accounting for about 11% in both manual tagging and SVM classification results. Besides those scam categories, the similarity in all scam categories between manual tagging and SVM classification results is clearly shown in Table 2.2. This observation strongly supports the preciseness and effectiveness of our SVM classifier.

## 2.5 Scam trends

In this chapter, we derive long-term trends in scams based upon the classification results described in Chapter 2.4 and FBI/IC3 annual reports. We focus on



(a) Precision/recall



(b) ROC curve

Figure 2.1: **Precision/recall and receiver operating characteristic (ROC) curve of SVM classifier.** Every scam category except “Etc” show over 95 % precision with over 60% recall and 90% false positive rate with under 5% false positive rate. Therefore it can be seen that the SVM classifier has good enough performance to show an overall trends in scam trends. All performance metrics were measured using 2000 samples divided randomly into 80% train set and 20% test set. Repeated 10 times.

the scam email dataset between the year of 2006 and 2014, and analyze the trends over the long term of nine years.

**Caveat.** Since our analysis relies upon the self-reported scam email dataset, the numbers and proportions of different scam categories may not represent the actual incidence of scams. Certain types of scam emails may be filtered out more frequently by email service providers and not delivered to users. In addition, people may not report specific type of scam emails since it is harder to recognize that those emails are scams. It is also reasonable to assume that self-reported scam email dataset may have a bias toward easily recognizable and observable scam emails. However, it is still meaningful to analyze the dataset since it would show the trends in scam emails that are actually perceived by people. We also tried to minimize possible biases by collecting scam emails from multiple sources.

### 2.5.1 Overall trends in scams

Annual number of scam complaints reported to FBI/IC3 is presented in Figure 2.2. In 2006, the number of complaints was about 200,000 and has increased up to 334,000 in 2009. Then the number of complaints has decreased slightly to 270,000 in 2014. Overall, the number of scam complaints reported to FBI/IC3 has increased by 33% between 2006 and 2014.

Figure 2.3 shows the overall trends in terms of fraction of scam emails reported in four scam reporting websites. Total number of scam emails reported has increased by 23% each year, from 9,700 in 2006 up to 29,500 in 2014. Although the number

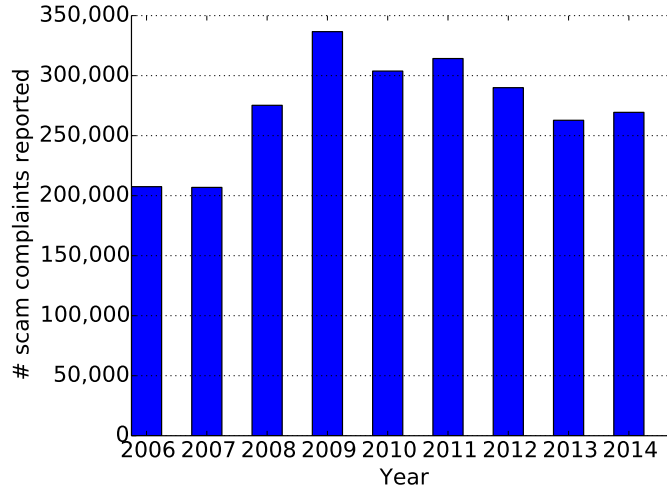


Figure 2.2: **Number of scam complaints reported to FBI/IC3.** [2] Number of scam complaints increases by 30% between 2006 and 2014.

of scam emails increased, we do not argue that changes in number of scam emails reported each year reflect increase or decrease of actual number of scam emails, since the activeness of each scam reporting website varies over time. Hence we analyze the changes in relative fractions of scam emails which would reflect trends in scam emails at certain degree. Later in this chapter, we focus on the fractions of each scam category and analyze the long-term trends in scam emails.

### 2.5.2 Targeted vs Non-targeted scams

We first analyze the overall trends in scams in terms of “targeted” vs “non-targeted” scams. We cluster scam emails in our dataset per Table 2.2 and derive a long-term trends as shown in Figure 2.4. We exclude scam categories belong to “Both targeted and non-targeted” scams to minimize confusions that may be resulted from the ambiguous nature of the mixed scam category. Our analysis clearly shows that non-targeted scams are decreased over the last 9 years while

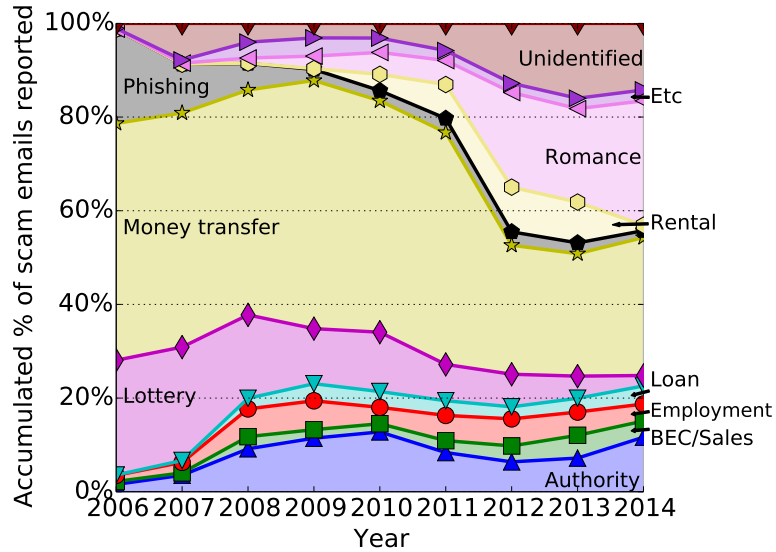


Figure 2.3: **Fraction of scam emails reported to scam reporting websites.**

targeted scams are in the opposite direction.

In 2006, non-targeted scams account for the majority of scam emails, about 96% of all scam emails while targeted scams has very limited percentage, about 0.7%. In 2014, on the other hand, the percentage of non-targeted scams decrease drastically down to 31%, while targeted scams show step increase up to 48%. This result implies that scammers' fraudulent methodology has improved, from simple spam-like scams (e.g., lottery and money transfer scams) to more personalized and plausible scams aided with context of specific victims (e.g., business email compromise, romance and sales scams).

### 2.5.3 Scams on the rise

We first focus on which scam categories are getting increased over time. We inspect the overall trends in scams shown in Figure 2.3 and find out three scam

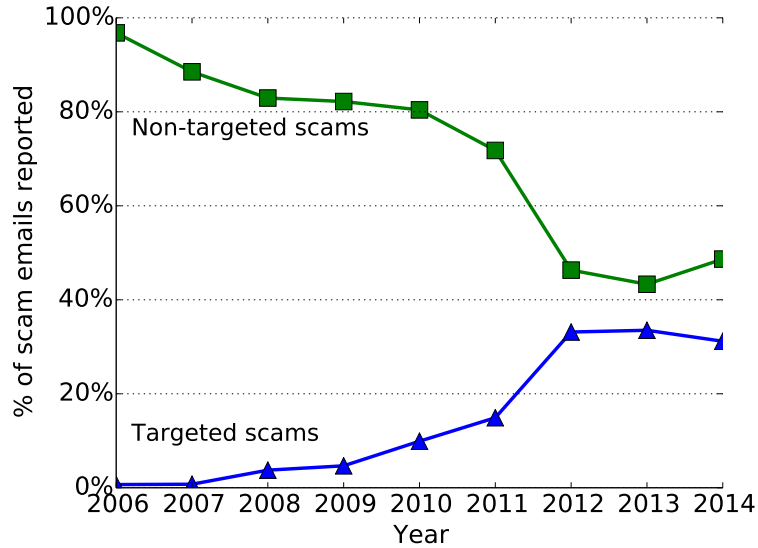


Figure 2.4: **Targeted vs non-targeted scams.** Targeted scams increases from 0.7% in 2006 up to 31% in 2014, but non-targeted scams decreases from 96% in 2006 down to 48% in 2014.

categories that show explicit increase over the last 9 years: *Authority*, *BEC/sales* and *romance* scams. We then compare our findings to FBI/IC3 annual reports [2] to cross-check and justify our observation. Figure 2.5(a) and Table 2.3 show our analysis of scams categories on the rise and corresponding results from FBI/IC3 annual reports, respectively.

Authority scam shows gradual increase from about 2% in 2006 up to about 13% in 2010 and 12% in 2014. Although it is hard to match our analysis to the reports due to different scam categorizations, we are able to find from FBI/IC3 reports that *FBI scam* is included as one of the most frequently reported scams in recent years, from 2009 to 2014. Even though FBI scam in FBI/IC3 reports does not cover all kinds of authority scams, it still partially support our analysis.

We also observe a similar finding for romance scams. Romance scam shows

rapid increase from less than 1% in 2006 up to 20% in 2012. According to FBI/IC3 reports, romance scam also has been prevalent from 2011 and included in the most frequently reported scams.

BEC/sales scams also increased over nine years, from 0.6% in 2006 to 2.5% in 2014. Although neither BEC or sales scams was identified as one of the most frequently reported scams in FBI/IC3 annual reports, BEC was considered an emerging scam category recently. According to 2014 FBI/IC3 annual report, BEC scam was reported in 2010 for the first time, and has evolved into more sophisticated and various forms from 2013. Total financial loss resulted from BEC scam in 2014 was estimated to be \$226 million dollars.

#### 2.5.4 Scams in the decline

From the analysis of scam trends, we are also able to find three scam categories that are in decline: *Lottery*, *money transfer* and *phishing* scams. Figure 2.5(b) shows the fraction of those three scam categories each year. Similarly, we try to find out corresponding scam categories from FBI/IC3 reports and list relevant results in Table 2.3.

Money transfer scam is the most frequently observed scam in our dataset, but we are not able to find exact matches in FBI/IC3 reports. We find from FBI/IC3 reports the statistics of *investment* and *Nigerian letter scams*, which are directly related to money transfer scams. Combining both investment and Nigerian letter scams, they are reported to be included in the most frequently reported scams



Category	Years								
	2006	2007	2008	2009	2010	2011	2012	2013	2014
FBI scams	X	X	X	✓	✓	✓	✓	✓	✓
Romance scams	X	X	X	X	X	✓	✓	✓	✓
Real estate scams	X	X	X	X	X	X	✓	✓	✓
Identity theft	✓	✓	✓	✓	✓	X	X	X	X
Investment & Nigerian letter scams	✓	✓	✓	X	X	X	X	X	X

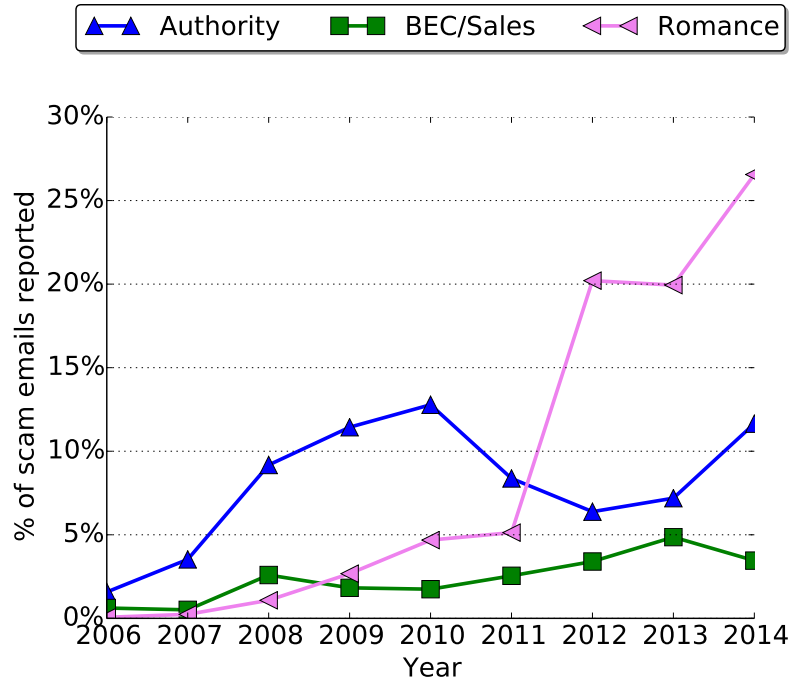
Table 2.3: **Most frequently reported scams in FBI/IC3 annual reports [2]**. FBI and romance scams are more frequently reported in recent years. On the other hand, identity theft, investment and Nigerian letter scams are less frequently reported in recent years.

until 2008 and are not included afterwards. Hence it is reasonable to argue that the statistics of FBI/IC3 reports support our observation that money transfer and phishing scams are getting less frequent in recent years.

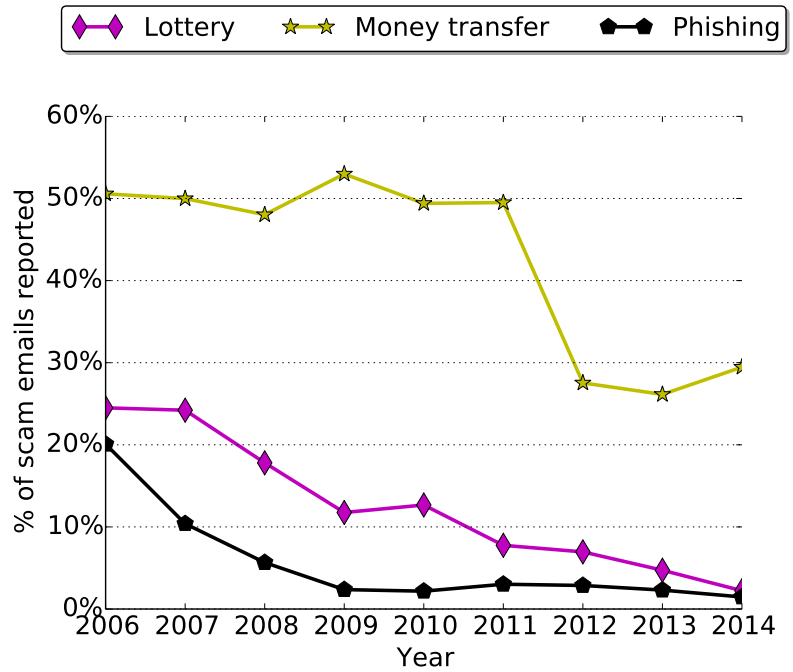
Phishing and Lottery scams, another typical non-targeted scams, are also decreased from 20% and 24% in 2006 to 3% for both in 2014. In FBI/IC3 reports, we are able to find statistics of *identity theft* which is a close match to phishing scam. As shown in Table 2.3, identity theft is one of the most frequently reported scams until 2010 and is not included afterwards.

### 2.5.5 Analysis

Through the analysis of trends in scams based on our scam email dataset and cross-checking FBI/IC3 reports, we find strong evidence suggesting that “traditional, un-targeted” scams such as lottery, money transfer and phishing scams are in steep decline. On the other hand, more targeted and credible scams such as



(a) Scams on the rise.



(b) Scams in decline.

Figure 2.5: **Scams on the rise and in the declines.** Fraction of scam categories on the rise and in decline.

BEC/sales, romance, rental and authority scams are getting increased. Our study on sales scams [3](#) reveals that scammers are organized in a few large groups and they use semi automated tools to deal with many number of potential victims. Therefore it is reasonable to assume that scammers are moving from spam-like non-targeted scams towards more sophisticated targeted scams.

## 2.6 Conclusion

In this chapter we presented a quantitative measurement study of Nigerian scam emails. We built a large-scale scam email dataset and created a list of scam categories commonly observed and reported by many users online. Then we clustered the scam emails per the scam categories using a machine learning classifier built based upon the manually tagged dataset. Our analysis of trends in scams strongly implies that targeted scams are getting more prevalent while non-targeted scams are in decline.

## Chapter 3: Sales Scams on Craigslist

### 3.1 Overview

Chapter 2 reveals that targeted scams are getting more prevalent while non-targeted scams are decreasing. In this context, this chapter and next chapter focus on two typical forms of targeted scams, *sales scam* and *rental scam* respectively. In this chapter, we present the measurement and analysis results of sales scams on Craigslist. Through this study, we aim to better understand the underground economy of Nigerian scams, and seek effective intervention points. In particular, we seek to address questions such as the followings: “Where are scammers located?”, “How do scam factories operate?” and “What features can we use to distinguish a scam email from a legitimate email?”.

In order to better understand Nigerian scams on Craigslist, we built an automated scam data collection system, *Scambaiter*. Using Scambaiter system, we posted honeypot ads to attract scammers. Then we received and replied to scam emails resulted from our honeypot ads. We perform various analyses of the massively collected dataset to better understand how scammers work. We also cluster observed scammers into groups based on a few key factors such as email addresses, shipping address, phone number and email payload.

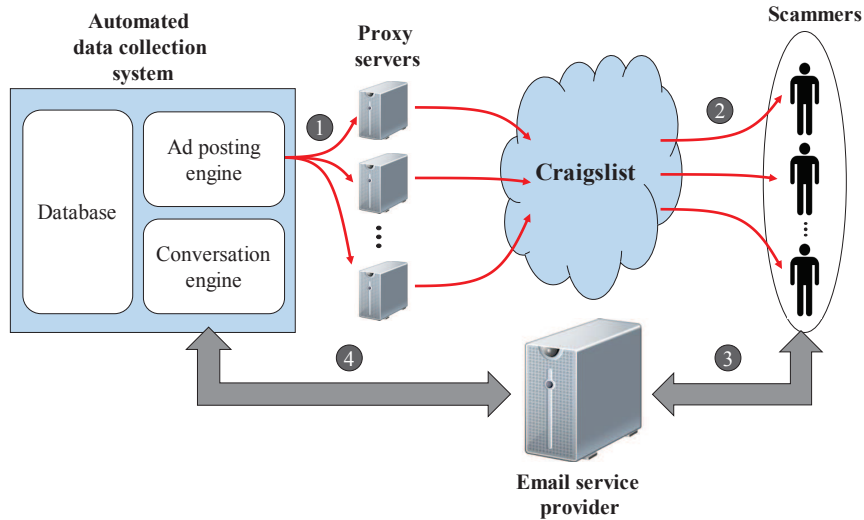


Figure 3.1: **Automated scam data collection using magnetic honeypot ads.**  
**(1, 2):** The system posts “magnetic honeypot” ads which would attract scammers only;  
**(3):** the scammers send scam emails in response to the magnetic honeypot ads;  
**(4):** the system automatically engages in email conversations with scammers.  
**Fraud attempt:** The conversation eventually leads to a fraud attempt, where the scammer sends a fake PayPal notification or fake check, and urges the victim to send the goods to the scammer-indicated mailing address.

## 3.2 Data Collection Methodology

We have built an automated data collection system that collects scam data on Craigslist as illustrated in Figure 3.1. Our data collection methodology is explained below.

### 3.2.1 Magnetic honeypot posts.

Our data collection focused on selling a variety of goods on Craigslist.

Our idea is to create *magnetic honeypot* advertisements that would *selectively attract scammers but not legitimate users*. To do this, we post unattractive advertisements, e.g., selling a used iPad at a price higher than new. More specifically, we

choose goods among a list of popular items on Amazon to make sure that the goods we are selling can be easily bought from Amazon or anywhere else. The selling price is set to be a little bit higher than the price of new product found on Amazon. Any sensible real user would conceivably not reply to such posts. However, scammers would — they might be using bots to crawl Craigslist or automate the response process, or might not carefully check the contents of each post due to lack of labor.

### 3.2.2 Automated communication with scammers.

We have built an automated conversation engine that performs linguistic analysis of incoming emails from scammers, and automatically engages in multiple rounds of communication with scammers. The engine periodically checks inboxes of email accounts used for Craigslist accounts and reads in all unread emails. Then it classifies the emails to identify valid scam emails. Our automated engine replies to a subset of the scam emails we receive — specifically, emails with a subject line that replies directly to the subject of our post. Henceforth, our automated engine exchanges multiple rounds of emails with the scammer, leading to the fraud attempt, e.g., fake PayPal notifications or fake checks.

The conversation engine is designed to generate 192 kinds of reply emails based on 13 pre-defined keywords. (e.g., sale, price and condition) observed frequently during the preliminary experiment conducted manually. The pre-defined keywords are divided into four groups and the conversation engine generates a corresponding sentence per each keyword group. Hence, the conversation engine creates up to four

Keywords	Example sentence
sale	Yes, it's still on sale.
available	Yes, the product is still available.
still	I still have it for sale.
price	The price is < <i>price</i> >.
firm	The firm price is < <i>price</i> >.
final	The final price is < <i>price</i> >.
condition, shape, detail	The condition is almost perfect since it was not used frequently.
why	I'm selling this just because it's not in use.
reason	I'm trying to sell it since I got a similar one correctly.
paypal	Sounds great. My paypal account is < <i>account</i> >. Please let me know when the payment is done.
check	I'm sorry but could you pay the money via paypal please?

Table 3.1: **Pre-defined keywords and example sentence generated from the keywords.** Corresponding sentences are changed periodically during the experiment. Keywords “paypal” and “check” are checked only for second and later replies.

sentences excluding greeting and enclosure, and if it failed in finding any keyword from a scammer’s email, it does not send out a reply email. Beside, specific keywords such as “paypal” and “check” are checked only for generating second or later replies, since usually those terms are not used in scammers’ first response. List of keywords and example sentences are listed in Table 3.1.

We also inspected 200 randomly selected scammer emails to check if we received emails from legitimated users. We expected that any legitimate users would mention unreasonably high price of the honeypot ads, or at least they would try to negotiate the price. Through the manual inspection, we found no emails explicitly sent from legitimate users. In many cases, however, the contents of emails are too short (e.g., only one sentence), so it is not clear to identify if an email is from scammer or not. Those emails were considered scam emails based on the nature of our

honeypot ads.

The most common type of fraud we observe is a fake PayPal notification stating that funds have arrived at the victim’s PayPal account, followed by requests for the victim to send the product to the scammer’s mailing address. A typical example of email conversation is posted in Figure 3.2, and more examples are posted in Appendix A.1.

### 3.2.3 IP address collection.

The IP address of an email sender provides insightful information, such as scammers’ geolocation. However, collecting IP addresses from email headers is not always feasible when the emails are relayed by the site (e.g., Craigslist), or if the webmail provider does not include source IP address in email headers (e.g., Gmail). To collect IP addresses of scammers, our automated conversation engine embeds an external image link into emails generated in response to a received scam email. Since the embedded link leads to a web server under our control, we can collect IP addresses of anyone who accesses image files we’ve embedded. The embedded link is unique to the corresponding advertisement so that we can later analyze the collected IP addresses based on factors such as city, product category and price.

### 3.2.4 Ethics

Since our experiment ultimately deals with human subjects, we put several controls in place to manage any harm to the participants. In addition, we went



iPhone 5 64GB (WashingtonDC)

[from: cathy caraballo <cathycaraballo93@gmail.com>]

Still available for sell??

*[Our response]*

Yes, the product is still available. Please let me know if you need more information.

[from: cathy caraballo <cathycaraballo93@gmail.com>]

Thanks for getting back to me [words omitted] I will give you \$680 for the item in order to out bid other buyer and \$60 for shipping via a register mail down to my Son, kindly get back to me with your PayPal email account so I can proceed now with your payment and if you don't have an account with PayPal, its pretty easy, safe and secured to open one. Just log on to WWW.PayPal.com [words omitted]

Thanks and God Bless.

*[Our response]*

Sounds great. My paypal account is sarkadejan@gmail.com. Thanks!

[from: cathy caraballo <cathycaraballo93@gmail.com>]

Hello Friend, just want you to know that your payment has been made paypal just mailed me now so check your inbox or spam and your money has been deducted from my account pending to your account.. [words omitted] tracking number and scanned receipt for verify and Here is the Shipping Details below [address omitted]

**Fake Paypal notification:**

[from: service@paypal  
<verifedtrackingshipp@mail2consultant.com>]

Dear Sarkadejan@gmail.com, You've received an instant payment of \$770.00 USD from Cathy Caraballo93, [words and images omitted]

Figure 3.2: **Example 419 scam thread.** The first scam response usually has one or couple of simple sentences showing scammer's interest in goods posted by the victim. The second scam response contains a fraud attempt through fake PayPal or bogus check. The scammer's offer is usually attractive since their offer price is higher than then victim's list price. Finally, the third and later scam responses urge the victim to send the goods to the designated mailing address.

through the process of getting our experiment approved by our institution’s human subjects review process. During the experiment, we collected scam emails by posting honey pot advertisements which may attract responses from legitimate users as well as scammers. Even though our honey pot advertisements are designed to be “unattractive” such that legitimate users would not be interested in replying, it is still possible that our experiment might receive responses from legitimate users that send an actual payment to buy a product that we have posted on Craigslist. In order to prevent this unintentional “victimization”, we consistently check if there were any actual payments made by legitimate Craigslist users. If a payment was made by a legitimate user, the victim would be provided with pertinent information about our experiment, and the item would be shipped to them or the refund procedure would be initiated immediately. In addition, any messages from this user would be purged from our collected data. Note that fortunately, we found no payment made by any legitimate users during the entire experiment.

Another issue concerns how we use the collected data that might contain private information about scammers. Throughout the experiment, we gathered messages that contain information such as shipping addresses and phone numbers which could potentially be used to identify scammers. We limit the use of raw data to email addresses, IPs, and text from messages that will not clearly identify the actual identity of the scammer. All other information is only included in aggregate to avoid revealing the identity of any scammers.

Finally, we adhered to Craigslist’s terms of use regarding posting advertise-

<b>Overview</b>	Duration of experiment	97 days (4/15/2013 - 7/19/2013)
	Cities/areas	20
	Product categories	4
<b>Honeypot ads</b>	Total number of ads	1,376
	Effective ads	629
	Flagged ads	747
<b>Emails</b>	Emails received	19,204
	Emails sent	9,902
<b>Email threads</b>	First scammer responses received	<b>13,215</b>
	First replies sent	8,048
	Second scam-related response received	<b>1,626</b>
	Fake PayPal payment emails (not threads)	<b>751</b>
	Bogus check payment threads	<b>885</b>

Table 3.2: **Summary of experimental result**

ments <sup>1</sup>. Specifically, each of our accounts only posted in a single location and were restricted to a posting rate of once every 48 hours.

### 3.3 Dataset

Table 3.2 presents a summary of the dataset we collected using the methodology described in Section 3.2. More details of each part of the table are explained below.

#### 3.3.1 Overview and terminology.

Our 419 scam data collection spans a duration of roughly *three months*, from 4/15/2013 to 7/19/2013. We selected 20 locations including 10 large and 10 small cities/areas from a list provided by Craigslist. The large cities include San Francisco,

<sup>1</sup><http://www.craigslist.org/about/terms.of.use>

Seattle, New York, Boston, LA, San Diego, Portland, Washington DC, Chicago and Denver and small cities/areas include Twin Tiers, Cumberland Valley, Meadville, Susanville, Siskiyou, Hanford-Corcoran, Santa Maria, Winchester, Southwest and Eastern Colorado.

We selected four product categories including *cell phone*, *computer*, *jewelry* and *auto parts*, which are used by many Craigslist users and therefore, many advertisements are posted daily as usual. As mentioned in Section 3.2.4, we posted our ads at very low rates, so that they account for only an unnoticeable fraction of the total traffic volume in each city on Craigslist. Specifically, we posted at most one advertisement per category per city every 48 hours, which makes at most 80 advertisements per 48 hours in total. The price of products used in the experiments ranged from \$80 to \$7,000.

Table 3.3 shows the terminology that we use to refer to honeypot ads and received emails throughout this thesis.

<b>Effective ads</b>	Magnetic honeypot ads that are not flagged by Craigslist until the expiration (1 week)
<b>Email thread</b>	Several emails in the same conversation
<b>First response received</b>	The first email sent by the scammer to us after seeing our Craigslist ads.
<b>First reply sent</b>	Our response to the first response received.
<b>Second response received, second reply sent</b>	The scammer's response to our first reply; our reply to that in turn

Table 3.3: **Terminology**

### 3.3.2 Magnetic Honeypot Advertisements.

During the experiment, we posted 1,376 magnetic honeypot advertisements over 20 large and small cities in the U.S. Among the whole advertisements posted, 747 advertisements were flagged by Craigslist, leaving 629 *effective* advertisements. 42 emails accounts (Craigslist accounts) were used during the experiment. We designed our system to post magnetic honeypot advertisements evenly distributed over posting time and product category to minimize possible biases in the collected dataset. Figure 3.3 illustrates distribution of effective advertisements over time of day. In this figure, the slight unevenness in distribution (over different times of the day and product category) partly stems from Craigslist’s flagging policy.

The average number of effective ads posted per each hour is 26.2. (min/max/median/standard dev = 20/38/26/4) The average number of effective ads posted per product category is 157.3. (min/max/median/standard dev = 144/187/149/20) It is believed that the degree of variation observed in both distribution would not cause any significant bias in the collected dataset.

### 3.3.3 Collected emails and threads.

The total number of emails received during the experiment was 19,204 and the number of emails sent is 9,902. Several emails in the same conversation are together referred to as a thread.

Among the total of 19,204 emails received in our data collection 15,270 were first responses. Among these first responses, our system determined that 13,215 rep-

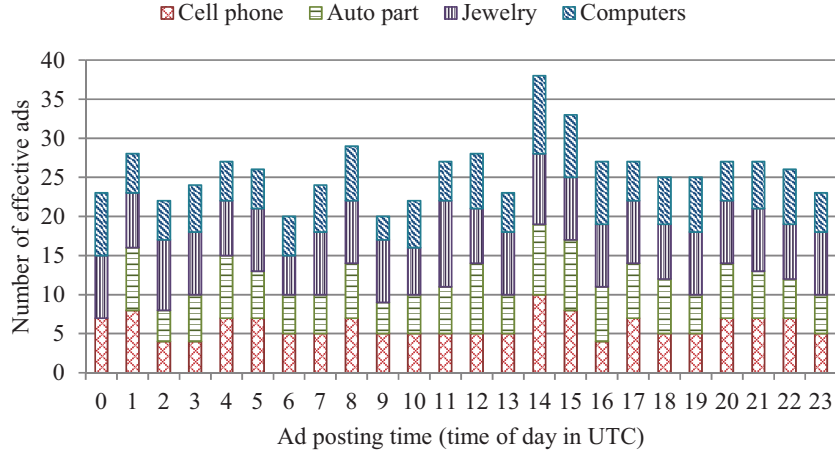


Figure 3.3: **Distribution of magnetic honeypot ads over ad posting time.** The ad posting engine posts magnetic honeypot ads every 48 hours or more in each city and category.

resented scam-related activities based on subject, pre-defined keywords and sender email addresses, whereas the remaining include spams and fake PayPal payment emails and emails delivered from email service providers. As a result, our system attracted 9.6 scam trials (first scam responses) per ad. We are not arguing in this thesis that we attracted all kinds of scammers since we might have missed some scammers who were too smart to take our honeypot ads. Even if our scam dataset is somewhat biased, it would be still meaningful to collect and analyze a large-scale scammer information to get a fundamental understanding of their operation.

From the 13,215 scam-related first responses, our automated data collection engine sent 8,048 first replies. As mentioned in Section 3.2, presently we only send replies to emails that directly reply to our posts. There are 9,008 out of 13,215 first responses reply directly to our posts — by including the subject line that we used for our ads. We did not send first replies for 960 out of 9,008 valid first responses

since our system failed in finding out pre-defined keywords necessary to generate corresponding replies.

For 1,626 of the threads, we received a second response from the scammer. Finally, we received 751 fake PayPal payment notifications emails and 885 bogus check fraud attempts. Note that we received multiple fake PayPal payment emails for some threads, and it was not always possible to tie a PayPal notification back to an email conversation thread, since for most fake PayPal notifications the source email address is different from those used in the email conversation.

### 3.4 Analysis of scammers' IP addresses

As described in 3.2, we collected IP addresses of scammers by embedding an external link to a product image. We gathered IP addresses from web logs of the server that hosts product image files.

#### 3.4.1 IP geolocation.

In the experiment, we observed 965 IP addresses over 22 countries. The total number of accesses to the image hosting server from those IP addresses were 7,759, and each IP address was observed 8 times on average. (min/max/median/standard dev = 1/182/3/16.1) 16 out of 965 IP addresses were web crawlers such as *MSN search engines* and excluded in our analysis. Figure 3.4 illustrates the IP geolocation of scammers who have accessed the embedded image links more than once. The source [hostip.info](http://www.hostip.info)<sup>2</sup> was referenced to retrieve a geolocation information of each IP

---

<sup>2</sup><http://www.hostip.info>

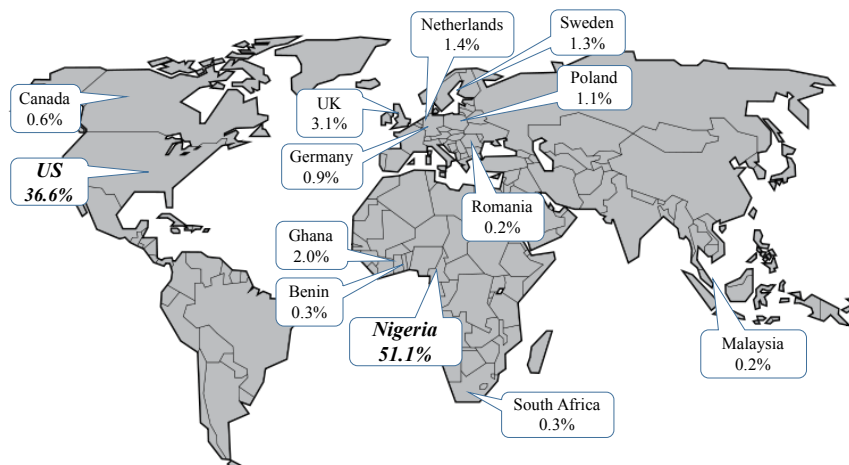


Figure 3.4: **IP Geolocation of scammers.** For 949 IP addresses observed, 51.1% are from Nigeria and 36.6% are from the U.S.

address.

Scammers' IP addresses were observed from all over the world but most of them were located in Nigeria and the U.S. In particular, 51.1% of collected IP addresses were from Nigeria and 36.6% were from the U.S. Note that this figure is plotted based on the number of unique IP addresses observed. It is also possible that some scammers could be using proxies, so the IP geo-location does not reflect their true location.

In Figure 3.5, the distribution of IP addresses over number of C class subnets (255.255.255.0/24) and B class subnets (255.255.0.0/16) is illustrated. We observed 406 class C subnets in total, and 40 of them take about half of whole IP addresses. Also, 10 out of 242 class B subnets account for about half of whole IP address. The result shows that small portion of subnets take major number of IP addresses observed, and it might imply the possibility of small number of scam factories dominating whole scam business.



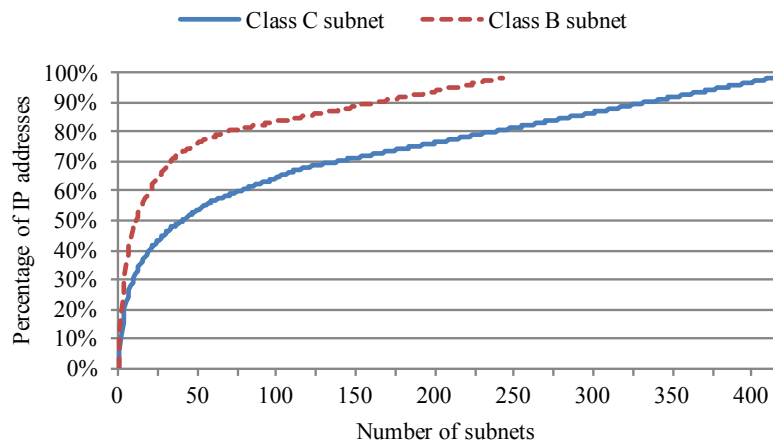


Figure 3.5: **Cumulative distribution of IP addresses over number of subnets.** The total number of class C subnets is 406 and class B subnet is 242. Half of IP addresses observed belong to 40 class C subnets or 10 class B subnets.

### 3.4.2 IP blacklist.

We cross-checked the collected IP addresses with a publicly available blacklist, *Project Honey Pot* [18] containing IP addresses of user-reported spam and scam generators. The result is outlined in Table 3.4. In particular, Project Honey Pot contains blacklisted IP addresses which were confirmed to be malicious; and graylisted addresses, which were detected but have not been confirmed to be malicious. From the IP addresses we collected, 41.1% are blacklisted by Project Honey Pot. 14.2% of IP addresses are in graylisted but not blacklisted.

About 44.7% out of the 949 IP addresses we found have not been blacklisted or graylisted by Project Honey Pot. Therefore, one contribution of our measurement study is to provide data to supplement existing IP blacklists.

Table 3.5 shows the top 20 IP addresses by the number of access to the external link embedded in our replies. All top 20 IP addresses are confirmed to be located

IP addresses	Percentage
Not in black/graylist	44.7%
Blacklisted	41.1%
Graylisted	14.2%

Table 3.4: IP addresses blacklisted by *Project Honey Pot* [18].

IP address	Country	# Times observed	Blacklisted?
41.211.193.XXX	Nigeria	298	-
41.203.67.XXX	Nigeria	241	Yes
41.203.67.XXX	Nigeria	204	Yes
41.203.67.XXX	Nigeria	160	Yes
41.211.198.XXX	Nigeria	93	-
<b>41.206.15.XXX</b>	Nigeria	89	Yes
41.184.21.XXX	Nigeria	89	Graylisted
<b>41.206.15.XXX</b>	Nigeria	88	Yes
41.211.201.XXX	Nigeria	85	-
<b>41.206.15.XXX</b>	Nigeria	79	Yes
<b>41.206.15.XXX</b>	Nigeria	77	Yes
41.220.68.XXX	Nigeria	73	Yes
41.203.67.XXX	Nigeria	71	Yes
<b>41.206.15.XXX</b>	Nigeria	68	Yes
<b>41.206.15.XXX</b>	Nigeria	64	Yes
<b>41.206.15.XXX</b>	Nigeria	60	-
<b>41.206.15.XXX</b>	Nigeria	58	Yes
<b>41.206.15.XXX</b>	Nigeria	57	Yes
<b>41.206.15.XXX</b>	Nigeria	56	Yes
<b>41.206.15.XXX</b>	Nigeria	56	Yes

Table 3.5: **Top 20 IP addresses by the number of times observed.** All top 20 IP addresses are from Nigeria. 11 of the 20 IPs (in bold) belong to the same class C subnet.

in Nigeria. More interestingly, 11 of them belong to the same class C subnet, 41.206.15.0/24, which strongly implies that they are part of the same scam factory.

We also found that 4 out of 19 IP addresses are not blacklisted yet.

## 3.5 Analysis of scammers' email accounts

### 3.5.1 Source, reply-to address discrepancy, and email account reuse.

Throughout the experiment, we collected 4,433 email accounts used for first responses of 13,215 scam threads, indicating average reuse counts of 3 per an email account. (min/max/median/standard dev = 1/101/1/5.5) The most frequently reused email account appeared in 101 threads. Figure 3.6 shows the distribution of email reuse counts. 2,410 email accounts, that is, 54.4% of total email accounts observed, were used in only one thread and about 10% were used in more than 6 threads. The majority of these single use only email addresses were initial inquiries about the product availability that never matriculated into further negotiations. However, many others were supporting emails used in the furtherance of the scam such as fake PayPal notifications, transportation agents, threats to contact the FBI (for when the product was not shipped) and similar emails. Some examples are posted in Appendix A.1.

We also observe that for 81.9% of the first responses received, the source email address is not the same as the reply-to address, and for 19.4% of the second responses received, the source email address is not the same as the reply-to address. This source and reply-to address discrepancy is shown in Table 3.6. The percentage of discrepancy was much higher, 97.6% for first responses, within the top 10 groups. The operating procedures of top tier organizations must account for the increased quantity of emails sent and received, both for management and security, which is

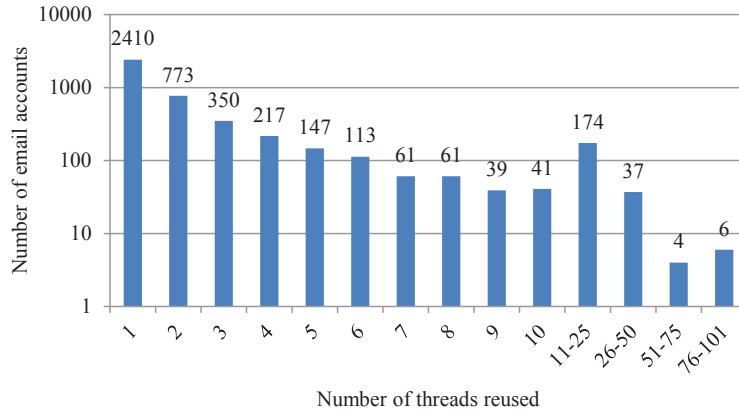


Figure 3.6: **Distribution of email account reuse count.** For total of 4,433 email accounts, 221 are used for more than 10 threads, and one email account is used for 101 threads.

<b>First responses</b>	13,125
<b>First responses with different source and reply-to addresses</b>	10,826 (81.9%)
<b>Second responses</b>	1,626
<b>Second responses with different source and reply-to addresses</b>	316 (19.4%)

Table 3.6: **Source and reply-to address discrepancy.** Source address and reply-to address are different in more than 80% of first responses, whereas the percentage is much lower for 2nd responses.

why they are more apt to split source and reply-to accounts.

Figure 3.7 shows that the set of source email addresses observed is much larger than the number of reply-to addresses observed. It is possible that the large pool of source addresses are disposable, and potentially automated, accounts that can be readily discarded and replaced as they are blacklisted. On the other hand the second, smaller tier of addresses are for more manageable monitoring and generally attempted to be kept “clean” for continued use over time. This intuition is supported by Table 3.7. We found a case that a single reply-to address was mapped to 108 source addresses. For total of 1,980 reply-to addresses, 141 were mapped to more

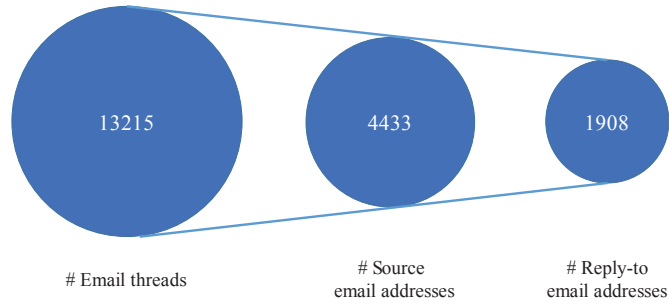


Figure 3.7: **Number of email threads, source email addresses and reply-to email addresses.**

Number of source addresses mapped	Reply-to addresses
108	1
Over 50	5
Over 20	64
Over 10	141
Over 5	246

Table 3.7: **Number of source addresses mapped to a single reply-to addresses.**

than 10 source addresses.

### 3.5.2 Email service provider.

Table 3.8 shows the proportion of each email provider the scammer uses, in comparison with the provider’s estimated world-wide market share as reported by Geekwire [19]. We find that the top email provider used by scammers is Gmail, which accounted for 65% of the scammer email accounts observed, followed by Microsoft (Hotmail and Live), which accounted for 10% of the scammer email accounts observed. Interestingly, in terms of world-wide market share, Yahoo is placed first, with a market share of 42.8% [19]. However, Yahoo only accounted for 3.5% of the email addresses we observed.

It is interesting to correlate this with underground market prices of bulk email accounts. Many PVA (Phone Verified Account) sellers of black market sell Gmail accounts for a price higher than other emails such as Yahoo or Microsoft. For example, one PVA seller<sup>3</sup> sells 1000 Gmail accounts for \$90, 1000 Yahoo accounts for \$15, 1000 Hotmail accounts for \$5 and 1000 AOL accounts for \$50.

Despite the most expensive market price, scammers seem to prefer Gmail over other email services. The reason for this might be the fact that Gmail is the only email service provider who supports free IMAP/SMTP and also hides IP address of email senders among top 4 email providers. IMAP/SMTP is imperatively necessary for scammers, since they deal with massive amount of emails. Also, since scammers are working underground, they might want to hide their IP addresses not to expose themselves and possibly to avoid filtering by email service provider of the recipient. Scammers' aversion to Yahoo can also be clearly explained, since Yahoo charges \$20 for SMTP/IMAP services. Although Microsoft supports free POP3/SMTP, those features were enabled recently in 2009. In this context, Gmail might be the reasonable solution for scammers.

### 3.5.3 Sample bad email addresses.

We also observed interesting “clusters” of bad email addresses. For example, the following email addresses we observed are close variants of each other:

This clearly shows that a scammer or a group of scammers create several email accounts to send out massive amount of scam emails.

---

<sup>3</sup><http://www.buybulkemailaccount.com/>

Email provider	Percentage	Est. market share	IMAP/SMTP	Hide sender IP?	Price for 1000 PVAs
Gmail	65.0%	25.0%	Yes	Yes	\$90
Microsoft	10.0%	20.3%	No (POP3/SMTP)	Yes	\$5
AOL	4.9%	11.9%	Yes	No	\$50
Yahoo	3.5%	42.8%	Yes (\$20)	No	\$15
Others	16.6%	—	—	—	—

Table 3.8: **Distribution of email service providers.** Gmail is preferred by majority number of scammers, despite the highest PVA(Phone Verified Account) price, possibly since Gmail is the only webmail service provider which supports free IMAP/SMTP and hides email senders’ IP addresses.

biglanre1@gmail.com    biglanre12@gmail.com    biglanre4@gmail.com  
biglanre10@gmail.com    biglanre13@gmail.com    biglanre5@googlemail.com  
biglanre11@gmail.com    biglanre14@gmail.com

### 3.6 Shipping Addresses and Phone Numbers

#### 3.6.1 Shipping Addresses.

153 distinct shipping addresses were identified throughout the study by threads that progressed far enough so that shipment of the product was expected. As with IP addresses discussed in 3.4, the majority, 70%, of the shipping addresses were located in Nigeria with 23% and 7% located in the United States and other foreign countries respectively (Table 3.8(a)). Some shipping addresses had multiple names, assumed to be aliases, associated with them. In one circumstance, seven names were associated with a single Nigerian address. For the classification of the threads into groups, emails with the same shipping address were assessed as belonging to the same group. Additionally, some addresses were in close proximity to each other. For example, three different apartment numbers for the same street address in Nigeria

were used as shipping addresses. In these circumstances, the threads were not assessed as belonging to the same group since being neighbors did not definitively indicate the occupants were part of the same organization.

### 3.6.2 Phone Numbers.

206 distinct phone numbers were identified during the study (Table 3.8(b)). Most were given either as part of the initial inquiry or during the follow-on negotiation emails, with only a few numbers withheld until the end of the purchase and then provided along with the shipping address. Diverging from the pattern seen with IP addresses and shipping addresses, the majority of the phone numbers, 91%, are registered within the United States and relatively balanced, but slightly in favor of, voice over internet protocol (VOIP) over cellular numbers. Of the 15 phone numbers identified as registered overseas, 12 were Nigerian, and all of these were associated with completed scam attempts and aligned with distinct Nigerian shipping addresses. Four phone numbers were either missing digits or area codes and therefore could not be categorized.

## 3.7 Scam Patterns

We report interesting patterns observed, including the distribution of scams received during various times of the day; response delay of the scammer; and what factors affect the scammer's response rate.



(a) **Shipping Addresses.**

Location	# Addresses	%
Nigeria	108	70%
USA	35	23%
Other	10	7%

(b) **Phone Numbers.** Unknown locations are due to missing digits or area codes.

Location	Service Type	Quantity
USA	VOIP	107
USA	Cellular	80
Nigeria	Unknown	12
Other Country	Unknown	3
Unknown	Unknown	4

Table 3.9: **Shipping addresses and phone numbers.**

### 3.7.1 When do scammers work?

The distribution of received time of first and second scam responses are illustrated in Figure 3.8. Figure 3.8(a) shows the distribution of posting time of effective advertisements and the received time of first scam responses. During the experiment, our automated data collection system posted advertisements evenly across the whole time of day. After removing the ads flagged by Craigslist, however, the number of effective ads varies slightly across different times of the day. Average number of effective advertisements at each hour is 26.2 (min/max/median/standard dev = 20/38/26/4).

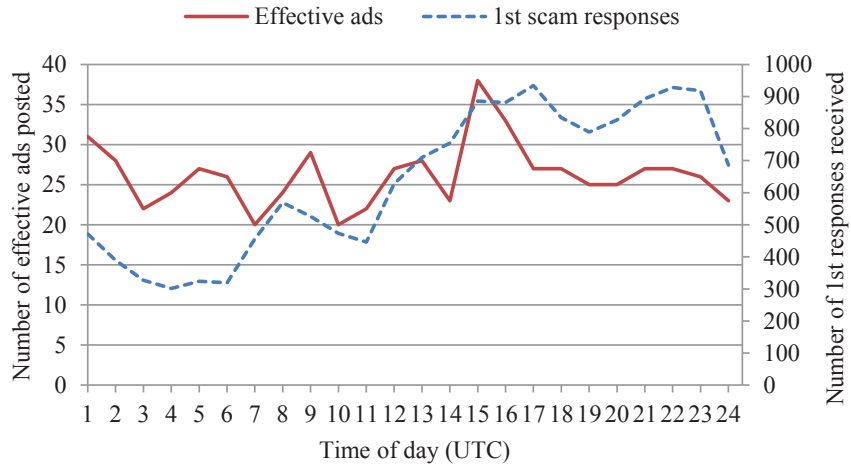
Figure 3.8(a) and Figure 3.8(b) show that both the first and second scammer responses peak during 11AM to 6PM, and from 8PM to 10PM UTC (Coordinated Universal Time), This corresponds to the time period between 12PM to 7PM WAT (West Africa Time), which largely overlaps with working time in Nigeria. Moreover,

the time period with the lowest scam responses is between 12AM and 6AM UTC, and it corresponds to 1AM to 7AM WAT. In addition, our first replies sent also peak during Nigeria's work hours, because our automated engine polls the emails every three hours and responds to the new emails. These observations support the result discussed in Section 3.4 that the majority of the collected IP addresses are from Nigeria.

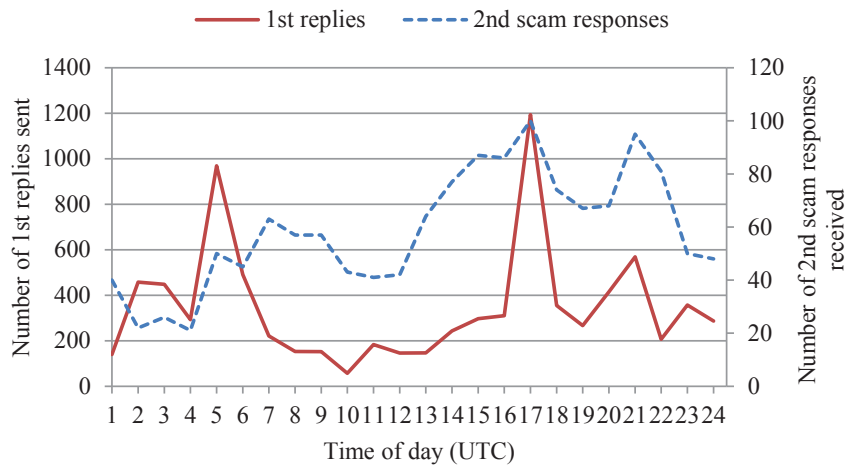
### 3.7.2 How fast do scammers respond?

Figure 3.9 shows the distribution of scammers' response time. Only 6.5% of first scam emails were received within 6 hour and about 36% were received within 24 hours. The response time can be an indication of the level of automation of scam factories. We will discuss scam process automation later in 3.8.

On the other hand, we can observe much faster response time for second scam responses. 26.5% of second scam responses were received within an hour from the first reply of ours, and about 90% were received within 24 hours. This is likely due to the fact that our automated engine sends replies to scammers no later than 3 hours from receipt of their scam email. Figure 3.8(b) shows our first replies peak during the work hours in Nigeria. This explains why scammers respond more quickly to our first replies.



(a) Ads posted and first responses received across different times of the day.



(b) First replies sent and second responses received across different times of the day. Our automated response engine sends replies within 3 hours upon arrival of a scammer email.

Figure 3.8: **Received time of scam responses.** The peak time of both first and second responses (11AM to 6PM UTC) largely overlaps the business hours in Nigeria.

### 3.7.3 Do product category and price affect scammers' response rate?

As shown in Figure 3.10, each of our ads attracted 2.2 to 19.8 scam trials, depending on the category of the advertisement. The number of scam trials per *Auto parts* advertisement is 17.6 while the plot shows 9.7 for *Cell phone*, 6.7 for

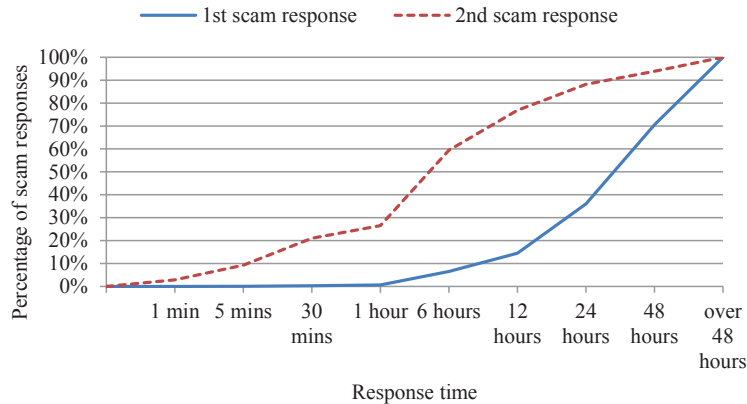


Figure 3.9: **Response time of scam emails.** Only about 36% of first scam responses were received within 24 hours from ad posting. However, about 60% and 90% of second scam responses were received within 6 hours and 24 hours from our first replies, respectively.

*Jewelry* and 5 for *Computer*.

An interesting pattern observed here is that scammers seem to prefer a specific product category, *Auto parts*, over others. We first tested if the distribution of number of scammers’ responses per ad in each product category follows normal distribution using *normaltest* function in Python SciPy package [20]. This function tests the null hypothesis that a given samples come from a normal distribution based on DAgostino and Pearsons test [21,22] that combines skew and kurtosis to produce an omnibus test of normality. P-values for each product category are presented in table 3.10. All P-values are lower than 0.05, hence the distribution of number of scammers’ responses per ad does not follow normal distribution.

	<b>Auto parts</b>	<b>Cell phone</b>	<b>Jewelry</b>	<b>Computer</b>
<b>P-value</b>	6.55e-11	4.28e-13	8.87e-17	1.42e-26

Table 3.10: **P-values of normal distribution tests.** P-values lower than 0.05 show that the distribution does not follow normal distribution.

We then ran non-parametric *KruskalWallis* test in Python SciPy package for

all product categories with the null hypothesis that the population median of all of the groups are equal. P-value of the test was  $6.85e-32$ , hence the null hypothesis is rejected.

Lastly we ran pairwise KruskalWallis tests with the null hypothesis that the population median of Auto parts and each of other product categories are equal. P-values for each pairwise test is presented in table 3.11. P-values for all pairs are lower than 0.05 showing that the population median of auto parts and other categories are not equal.

	<b>Auto parts/ Cell phone</b>	<b>Auto parts/ Jewelry</b>	<b>Auto parts/ Computer</b>
<b>P-value</b>	4.89e-12	3.42e-18	1.08e-26

Table 3.11: **P-values of pairwise KruskalWallis tests.** P-values lower than 0.05 show that the population median of given pair is not equal.

A series of tests described above show Auto parts ads attracted statistically significantly more scammer responses than ads of other product categories. We are not able to state an explicit reason for the higher number of scam trials per ad for Auto part category. Reasoning for this observation would be left as a future work.

The number of advertisements posted by all Craigslist users is not a valid factor since we observed almost similar number of advertisements over 4 product categories used in our experiment. Also, price does not seem to be a valid factor since we are not able to find out any consistent pattern in 3.10 in terms of product price. We were not able to find out any correlation between number of scam trials and product price.

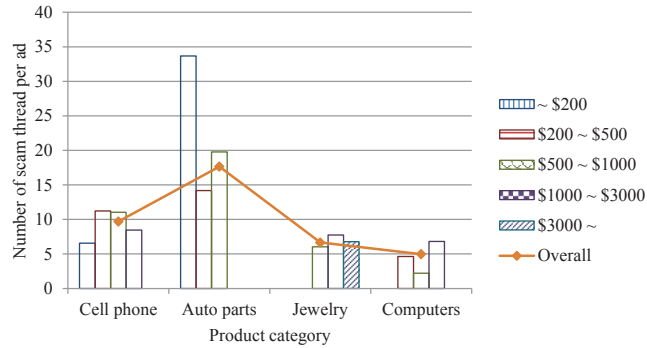


Figure 3.10: **Number of first scam responses per effective ad.** Each of Auto parts ad attracted 17.6 first scam responses in average while each of computer ad attracted 5 first scam responses. For this plot, we used a subset of the scam threads which we were able to link to an ad.

### 3.8 Level of Automation

One interesting question is whether the scam process is automated, and to what extent is it automated.

By combining various clues, such as inter-arrival time for the same email address and received email distribution across various times of the day, we can draw the conclusion that the scam process is automated to some extent, but not completely so. More details are provided in Table 3.12 and below.

#### 3.8.1 Signs of automation.

We observed clear signs of automation, including duplicate or templated responses observed in all stages of the scam process, outcome of broken scripts in subject lines, extremely short inter-arrival times for the same email address.

We now provide more explanation on the inter-arrival times for the same email address. Table 3.13 shows some of the largest email bursts received. An email burst

is a sequence of emails sent from the same email address within a very short time interval (no more than 15 seconds between emails). Table 3.13 suggests that the largest bursts observed consists roughly 8 to 15 emails, with a mean interarrival time of 2.5 to 5 seconds. These bursts were not directed solely at Craigslist ads within a single city. One burst sent 15 email messages to multiple ads across 3 different cities. More interestingly, two of the largest bursts observed consist solely of second replies.

We also observed many emails (including first response, second and later responses, as well as payment notifications) have exactly the same contents, or clearly use the same template to generate the content (Figure A.9 in Appendix A.1). We also observed outcomes of broken scripts in first responses received. These demonstrate that the scammer are using some (semi-)automated tools to automate their response process, and more interestingly, their tools sometimes broke and generated email subject lines that are not human readable (Figure A.3 in Appendix A.1).

### 3.8.2 Signs of manual labor.

On the other hand, we also observe signs of manual labor. First, according to Figure 3.8, scam responses peak during working hours in Nigeria, which accounted for 50% of the IP addresses we observed in Figure 3.4. Second, we received second and later scam responses containing curses — presumably the scammer was frustrated with us not shipping them the goods. Interestingly, we observed same curse emails occurring multiple times (Figure A.5 in Appendix A.1). It is likely in this

case that the angry scammer is copying and pasting the curse response. Also, in the latter stage of our data collection, some curse emails were received as a second response, before even reaching the payment stage — this could be a sign that the scammer actually started to detect our automated data collection.

The overall analysis of scam automation is outlined in [Table 3.12](#).



Stage	Signs of automation	Signs of manual labor	Conclusion
Reading in Craigslist ads First scam response	short inter-arrival time of first response broken scripts in email subject duplicate/templated email contents	received emails peak during work hours	Both first and second responses are partially automated Scammers may need to manually run or attend to automated tools
Second and later scam response	short inter-arrival time of second response duplicate/templated responses	scammers' curse emails (Figure A.3 in Appendix A.1), received emails peak during work hours	
Fake payment notification	duplicate/templated responses	Wrong email address/name in the notification	

Table 3.12: **Analysis of scam automation.**

# Emails in Burst	Mean Interarrival Time	# Cities
15	5.2 sec	3
11	4.5 sec	3
11	5.6 sec	4
8	5.5 sec	3
7	4.3 sec	7
6	3.7 sec *	6
5	2.4 sec *	5

Table 3.13: **Example Interarrival Times for Burst Traffic from a Single Email Address.** All emails in the same burst have exactly the same content. (\*: Second response emails)

### 3.9 Classification

In order to discover how prevalently these scammers/organizations infected Craigslist and determine the scope of their operations, we classified the email messages into groups based on similarities within their attributes.

#### 3.9.1 Conservative Classification Strategy.

We first used a very conservative clustering strategy to classify scam activities observed into scammer groups. Specifically, if two scam threads shared *exactly the same email addresses, shipping address, or phone numbers*, they are grouped as the same scammer group. Email addresses whose prefix were 90% identical were individually reviewed along with other attributes such as email textual content and IP addresses so series such as the `biglanreXX@gmail.com` addresses noted earlier were also grouped together when multiple attributes showed similarities. In this way, we are highly confident that two scam threads belong to the same scammer group when we place them into the same cluster.

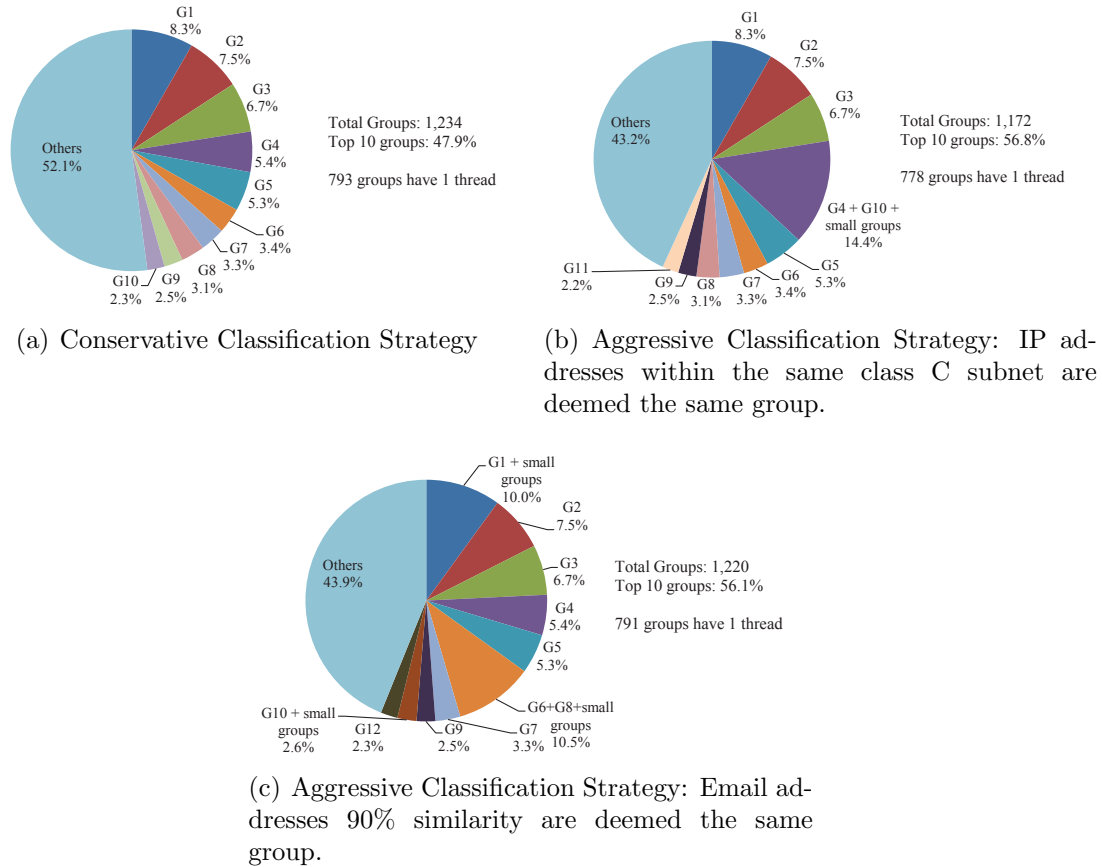


Figure 3.11: **Scammer Group by Number of Threads.** Small number of groups account for about half of scam threads.

### 3.9.2 Top 10 Groups.

Based on our very conservative classification strategy, we found that *the top 10 groups accounted for 48% of all received scam threads* (see Figure 3.11(a)).

Further analysis of the top 10 groups showed that they operate over (almost) all cities where we posted ads (Table 3.14), and most of them throughout the entire duration of our data collection (Table 3.15). Additionally, all groups responded to ads from all categories of products we advertised, *cell phone, computer, jewelry* and *auto parts*.

We give more detailed information about the top 10 groups below. Table 3.14 lists details including the number of threads associated with the group, source and reply-to email addresses, associated shipping addresses and phone numbers for the top 10 groups by number of threads. Almost all of the top 10 groups had an extremely high ratio (ranging from 86.9% to 100.0%) of threads whose source email address is different from the reply-to address, Group 10 had a 86.9% ratio which was anomalous until it was discovered that a single email address was used as both the source and reply-to address in 11.1% of the threads.

Interestingly, only one of the top groups had shipping addresses associated with the group. Our assessment is that 9 groups are more sophisticated in their separation of initial inquiries and transition to scam attempts, successfully segregating the two in order to keep their clean accounts and personal information off blacklists. One group on the other hand appears unconcerned with this separation, working on a volume basis using the same email addresses and phone numbers for finalizing scam attempts that are used throughout the negotiating process.

Group	# Threads	# Source Addresses	# Reply-To Addresses	% Source $\neq$ Reply-To	# Shipping Addresses	# Phone Numbers	# Cities	# Categories	Primary Category
1	1096	178	23	100.0%	0	0	18	4	Auto parts
2	993	270	64	98.7%	7	9	20	4	Balanced
3	885	313	48	95.8%	0	2	19	4	Jewelry
4	714	106	37	97.6%	0	0	20	4	Auto parts
5	700	52	11	98.6%	0	2	20	4	Balanced
6	449	182	30	98.0%	0	1	17	4	Auto parts
7	441	60	17	97.5%	0	1	20	4	Auto parts & Jewelry
8	416	103	10	100.0%	0	0	20	4	Auto parts
9	330	19	8	94.8%	0	0	19	4	Auto parts & Jewelry
10	306	71	23	86.9%*	0	0	20	4	Auto parts

Table 3.14: **Top 10 Groups.** Top 10 groups account for about 48% of emails threads. Scam emails of these groups were found in almost all of 20 cities and they covered 4 categories. They usually use a smaller number of reply-to addresses relative to the number of source addresses.

<b>Group</b>	<b>First Email</b>	<b>Last Email</b>	<b>Duration</b>
1	17 Apr 09:00	17 Jul 07:23	91 days
2	17 Apr 23:12	17 Jul 14:03	91 days
3	19 Apr 10:51	16 Jul 22:44	89 days
4	16 Apr 08:37	8 Jul 21:11	84 days
5	16 Apr 20:05	14 Jul 20:33	89 days
6	22 Apr 12:58	16 Jul 22:09	86 days
7	20 Apr 02:45	3 Jul 08:37	75 days
8	16 Apr 18:07	11 Jul 11:15	86 days
9	16 Apr 03:17	2 Jul 11:07	78 days
10	17 Apr 15:04	14 Jul 18:21	89 days

Table 3.15: **Top 10 Group Durations.**

Table 3.15 and Figure 3.12 show group activities over time. We make the following interesting observations. First, as mentioned earlier, all top 10 groups were active throughout the entire duration of the data collection. Second, Figure 3.12 shows that peak activities of a subset of the top 10 groups aligned with each other (e.g., the 5 plots on the left-hand side had aligned peaks and lulls). As mentioned later in Section 3.9.3, some of the top 10 groups merged when we used slightly more aggressive grouping criteria — therefore, it is likely that in reality, a subset of the top 10 groups are actually the same big group. The aligned peaks and lulls as shown in Figure 3.12 gives more evidence to support this hypothesis.

Finally, activities for most groups tapered off towards the end of the data collection, partly due to a combination of several reasons: the lack of new postings near the end of the research period, the expiration of postings created earlier, Craigslist’s flagging of some of our postings, and additionally, some scammers may have started detecting our automated data collection.

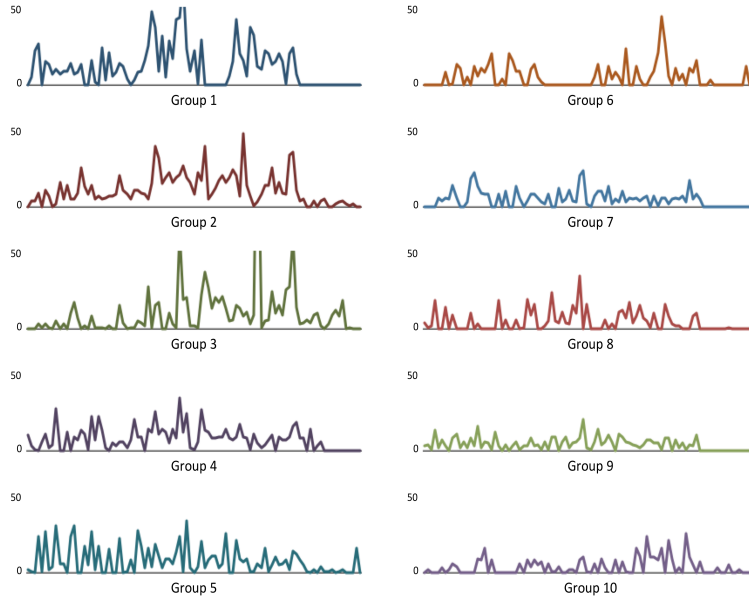


Figure 3.12: **Emails Per Day - Top 10 Groups.** A subset of the top 10 groups show aligned peaks and lulls. Activities tapered off towards the end, partly due the lack of new posts and expiration of existing posts.

### 3.9.3 More Aggressive Grouping Strategy and Findings.

Our first grouping strategy in section 3.9.1 was extremely conservative. There are many other attributes that we could have used in the grouping, but chose not to. For example, for similar email addresses, while we manually inspected a subset of them and marked them as being the same group, such as the aforementioned `biglanreXX@gmail.com`; for others such as `madelineXX@gmail.com` and `alexandraXX@gmail.com` we conservatively chose not to mark them as the same, since Madeline and Alexandra are common English names.

We observe, however, that if one applied a slightly more aggressive grouping criterion, some of the top 10 groups would have merged. For example, Figure 3.11(b) shows the grouping results if we additionally merged IP addresses from the same class C subnet to the same group. In this case, the 4th and 10th largest groups from

Figure 3.11(a) are merged, and is now the biggest group in Figure 3.11(b). The remaining ordering of the top 10 are preserved.

Figure 3.11(c) shows the grouping results if we additionally merged email addresses 90% similar to the same group. The similarity between two email addresses  $A$  and  $B$  (containing only the substring before ) is defined as:

$$\text{sim}(A, B) := 1 - \frac{\text{edit\_dist}(A, B)}{\min(\text{len}(A), \text{len}(B))}$$

Intuitively, a 90% similarity between two email addresses means every 1 out of 10 characters may be different. In Figure 3.11(c), the 6th and 8th largest groups from Figure 3.11(a) are merged, and is elevated to the largest group. The remaining ordering of the top 10 are preserved.

### 3.9.4 Classification Summary.

Summarizing the above, our classification effort clearly indicates that *a small number of scam groups account for half or more of the total scam activities we observed*. These scam groups work across all cities, and were continuously observed throughout the data collection.

### 3.10 Literature Review

There have been a number of previous studies that have looked at the structure by Smith [5], Buchaman and Grant [11] and estimated losses from advance fee fraud by Dyrud [23]. Whitty and Buchaman [4] and Rege [24] have investigated the



dynamics of online dating scams. More broadly, Stajano and Wilson [25] created a taxonomy of the different types of psychology motivations used by scammers. Garg and Nilizadeh [26] investigated whether economic, structural and cultural characteristics of a community affects the scams on Craigslist. Their work focuses on potential scammers' advertisements posted on Craigslist. Tive [27] introduced in his study various techniques of advance fee fraud. Herley [28] has argued that Nigerian scammers deliberately craft their messages to be unbelievable as a method of reducing the number of replies from people that are unlikely to fall victim to these scams. In contrast, our study aimed to be more focused on collecting empirical data to enable a data-driven analysis that does not rely in self reported statistics. Isacenkova et al. [29] identified a thousand scam groups from an existing scam email dataset with the help of a multi-dimensional clustering technique. This study also argued that scammers' email addresses and phone numbers are crucial factors of the clustering. Goa et al. [30] investigates the use of ontology-based knowledge engineering for Nigerian scam email text mining. Unlike previous studies, in our investigation we have focused on 1) understanding in great depth the prevalence and techniques, and 2) identifying the structure of larger scale groups of scammers that are engaged in attempting to defraud people posting goods for sale on Craigslist.

Another large body of recent work has set about conducting empirical measurements to understand the dynamics and economic underpinnings of different types of cybercrime. Much of this work has been focused on spam email [31, 32], illicit online pharmacies [33], and mapping out scam hosting infrastructure [34, 35]. Our work builds on this, but focuses deeply on the Nigerian scam problem in particular.

We have conducted, to our knowledge, the first large scale empirical measurement study of 419 scams. It provides us with insights into how these scams are organized and how they might be better deterred in the future.

### 3.11 Discussion

We have presented an in depth data-driven analysis of Nigerian scammers. This section will serve to provide a higher level view of our analysis to put it into context and discuss how our analysis might be used to deter these types of scams. In addition, we will describe future work that are planning to undertake that will improve our understand of these scams even further.

**Larger Organizations.** Our clustering results reveal that a large portion of the scam attempts are originating from a small set of groups that we can link together via their reuse of email addresses, shipping address, phone numbers, and similarity of the content in their messages. Our conservative estimate is that ten groups are responsible for about 50% of the scam attempts we received<sup>4</sup> This indicates that while this scam is highly prevalent that are only a relative small number of groups engaged in this activity. If these groups could be disrupted it would have a large impact on reducing the number of people targeted. As future work we plan to improve our measurement infrastructure to collect more information such as browser cookies that will enable us to more accurately cluster these groups. We will also work on improving our ability to estimate how many individuals are involved in this

---

<sup>4</sup>Note as we become less conservative with our clustering criteria some of the small group begin to merge into the larger groups and some of the larger groups begin to merge together.

scam and the division of labor among the individuals within each group.

**Locations of Scammers.** We also find that all of major groups are based in Nigeria based on IP and shipping addresses. However, some of these groups use shipping addresses in both Nigeria and the U.S. This indicates that they might have some limited ability to receive packages and reship them to Nigeria. As future work, we plan to identify shipping addresses associated with the major groups and ship them items with GPS tracking units embedded into the device. This will give us more insight into how and where the stolen goods are resold.

**Methods and Tools.** Our analysis offers many clues about the level of automation and sophistication of the tools used by these scammers. We have found strong evidence of automated tools that are run manually or require manual attending. In addition, we find that these tools are used to automate both initial responses and in some cases follow-up responses. Also, some of these tools are able to crawl multiple geographic regions on Craigslist and parse the posting's subjects and contents to include in the reply. However, some the tools are fairly limited and include static text in the body of the reply or cannot parse subjects of listings. This relative lack of sophistication in these tools indicate it would be possible to incorporate the static messages into spam filters and at least force the scammers to develop more advanced tools. As future work we plan to design experiments focused on gaining more insight into the tools being used and their limitations. This might include crafting messages that ask questions targeted at better understand which messages from the scammers are automated.

**Email Account Usage.** Our analysis of email accounts used be the scammers,

shows that they tend to use a large number of email address for the initial message that are quickly abandoned. However, they normally set the reply-to address in the initial message to a different email address that is reused often and longer lived. These longer lived email accounts offer a potentially better point of intervention, since it is conceivable these accounts can be blacklisted or banned before the scammer completes multiple email exchanges with the victim. In addition, as future work we will provide these email addresses to webmail providers and corporations such as PayPal to see if they have been used for other scams. This analysis might in turn help identify additional email accounts used by these groups.

**Filtering Messages.** In the course of our analysis we also identified many recurring themes in the content of the messages. These included the scammers claiming to be overseas military personnel to explain why they were located in Nigeria. Including religious content in their messages to gain the trust of their victim. Finally, they often used abusive language to coerce their victims into actually shipping the items. In addition, it would be feasible to exploit common linguistic features of scam email contents [30]. The combination of these patterns and the different reply-to address might be effectively used to improve the filtering of these messages. As future work will test this hypothesis by building improved filters that are more effective at detecting Nigerian scam messages.

### 3.12 Conclusion

In this chapter, we have presented a large scale empirical analysis of targeted Nigerian scams observed on Craigslist. From this we have learned valuable information on a variety of scam patterns such as scammers' working time and their response time to our ads and emails and discussed a degree of automation of scam process. Our analysis of IP addresses and shipping addresses indicates that the majority of scammers are located in Nigeria, but there is a smaller presence in the USA. We also found that around 10 groups account for almost half of scam attempts. Finally, we presented some higher level discussions based on our analysis and identify some potential points along this scam to intervene that might prove to be effective at deterring these scams.

## Chapter 4: Rental scams on Craigslist

### 4.1 Overview

In this chapter, we present our measurement study of rental scams on Craigslist. We conducted the first systematic empirical study of large-scale online rental scams on Craigslist. This study was enabled by a suite of techniques that allowed us to identify scam campaigns and our automated Scambaiter system introduced in chapter 3 that was able to collect additional information by conversing with scammers. Our measurement study sheds new light on the broad range of strategies different scam campaigns employ and the infrastructure they depend on to profit. We found that many of these strategies, such as credit report scams, were targeted directly at the rental domain and are structurally different from traditional advanced fee fraud found in previous studies. In addition, we measured the effectiveness of Craigslist's efforts to remove the initial scam postings. Finally, we found that many of the large-scale campaigns we detected depend on credit card payments, suggesting that a payment level intervention might effectively demonetize them.

## 4.2 Data Sets

This chapter focuses solely on *scams* and we consider spam, such as off-topic and aggressive repostings, as outside the scope of this chapter. Throughout this chapter, we define a rental listing as a scam if i) it is fraudulently advertising a property that is not available or not lawfully owned by the advertiser and ii) it attempts to extract money from replies using either advanced fee fraud or “bait-and-switch” tactics.

The basis of our study relies upon repeated crawls of the rental section on Craigslist in different geographic locations to collect all listings posted in these regions and detect listings that are subsequently flagged <sup>1</sup> We then use a combination of manual searching for reported rental scams and human-generated regular expressions to map fraudulent listings into scam campaigns. For a small subset of listings that are difficult to identify as scams or legitimate, we build an automated conversation engine that contacts the poster to determine the validity of the listing. Finally, we crawl five other popular rental listing sites to detect cloned listings that have been re-posted to Craigslist potentially by scammers.

### 4.2.1 Rental Listing Crawling

Our primary data set is based on listings collected from daily crawls of rental sections on Craigslist across 20 different cities and areas in the United States with the

---

<sup>1</sup>Craigslist detects and deletes user listings that violate the terms of use (e.g., excessively duplicate listings or malicious scam listings).

<b>Overview</b>	Duration	141 days (2/24/14-7/15/14)
	Cities/areas	20
<b>Rental ads</b>	Total posted	2,085,663
	Flagged for removal	126,898 (6.1%)
	Deleted (by user)	338,362 (16.2%)
	Expired*	1,620,403 (77.7%)

Table 4.1: **Dataset summary.** About 6% of rental ads are flagged for removal by Craigslist. Rental ads are considered to be expired 7 days after being posted.

largest population [36]: New York, Los Angeles, Chicago, Houston, Philadelphia, San Antonio, San Diego, Dallas, San Francisco (Bay area), Austin, Jacksonville, Indianapolis, Columbus, Charlotte, Detroit, El Paso, Memphis, Boston and Seattle. Our crawler revisited each crawled ad three days after the first visit to detect if they have been flagged by Craigslist. The crawler performed a final recrawl of any unflagged listings 7 days after the first visit to determine if they have been flagged or expired. We also collected rental ads from five additional major rental listing websites, *Zillow*, *Trulia*, *Realtor.com*, *Yahoo! Homes* and *Homes.com*.

Our crawler tracked all rental section ads on 20 cities/areas on Craigslist, for a total duration of 141 days, from 2/24/2014 to 7/15/2014. Table 4.1 shows the overall summary of this dataset. In whole, we collected over two million ads, among which 126,898 have been flagged by Craigslist.

## 4.2.2 Campaign Identification

Our crawling of Craigslist produced a large set of flagged and non-flagged ads that are potentially scam listings. We know that some of these ads are scams and that many of these are linked to a smaller number of distinct scam campaigns.



Due to the large number of ads in our data set a brute-force approach of manually analyzing a large set of ads would not be effective and would require a domain specific understanding of how scam ads differ from legitimate ads. In order to overcome these challenges, we bootstrap our knowledge of scam postings by finding a small number of suspicious ads in a semi-automated manner. To this end, we manually surveyed a broad range of user submitted scam reports online [37–39] to gain some initial insights about rental scams. Based on these insights, we constructed the following heuristics to identify an initial set of suspicious rental listings:

- Detect suspicious cloned listings by correlating listings posted to Craigslist with other rental listing websites, in particular, cloned ads from other sites that exhibit a substantial price difference (e.g., 20% difference).
- Detect posts with similar contents across multiple cities, e.g., posts with the same phone number or email addresses.
- Focus on ads flagged by Craigslist, and manually identify suspicious scam listings. As we will report in detail later, not all flagged posts are scam listings; and conversely, not all scam posts were flagged by Craigslist
- Identify ads that are similar to user-reported scams.

### 4.2.3 Campaign Expansion Phase: Latitudinal

For some of the campaigns we identified and hand labelled a small number of initial scam posts. Based on these we would like to identify other similar listings that are part of the same campaigns using automated and semi-automated methods. To this end, we used an approach that uses human-generated scam signatures.

#### 4.2.3.1 Human-generated scam signatures.

Our first approach is to manually inspect the handful of ads that we identified to be in the same campaign, and summarize a unique signature to identify this campaign. For example, one of the credit report scam campaigns have the following unique signatures: email accounts corresponding to the regular expression "[a-z]+(\s@\s)(yahoo)(\sdot\s)(com)" and no other contact information is included.

We then applied our signatures to all of our crawled ads, to identify additional ads that belong to the same campaign. As detailed in later sections, we will rely on a combination of human and automated verification techniques to confirm that scam ads identified by these signatures are indeed scams.

### 4.2.4 Campaign Expansion Phase: Longitudinal

For the initial scam postings we identified above, and the suspicious listings we identified in the latitudinal campaign expansion phase (Section 4.2.3), we wanted to confirm whether these are indeed scam messages. To this end, we built an automated conversation engine to converse with the suspected scammer, to see if the

conversation would lead to a phase where the scammer requested payment from us.

#### 4.2.4.1 Automated conversation engine.

We manually inspected the suspicious ads and found that some of them were clearly scams, e.g., the ads with a specific phone numbers that were reported as scams by many users. For others, while the ads appear highly suspicious, we were not sure whether they were scams as opposed to the more harmless spam posting from aggressive realtors or other service providers advertising their service/rentals.

We therefore modified the automated conversation engine described in previous Chapter 3.2.2 to i) verify whether a suspicious ad is a scam and ii) collect additional data. More specifically, we first selected a few suspicious ads and performed the email conversations manually. Then it was fairly straightforward to distinguish between legitimate users and malicious scammers during the email conversation. For example, clone ads scammers usually wanted to proceed with the rental process online since they were not in town for good purposes (e.g., serving in mission trip to Africa). From the preliminary conversations, we were able to generate a set of linguistic features (e.g., keywords such as “serving in mission” or rent application templates) and other types of features (e.g., embedded links to certain redirection servers) that distinguish rental scammers from other legitimate users.

The conversation starts with first emails sent by the conversation engine. First emails consist of up to three sentences excluding greeting and enclosure. Example sentences generated for the first emails are listed in Table 4.2. The conversation

engine composes the first email using 6 pre-defined sentences for the first sentence, 6 for the second sentence and 5 for the third sentence, making a total of 180 kind of email contents.

After sending the first email, the conversation engine waits for a response email from a scammer. Since there exist various kinds of rental scams, different strategies are deployed for each rental scam category. In case of “clone scam” which will be described later in Chapter 4.3.2, the conversation engine searches for specific rent application forms that the scammers usually ask their victims to fill in. Since scammers usually send more than one response emails (e.g., an email explaining scammer’s situation, an email with pictures of a house and another email with rent application form), the conversation engine focuses on the application form only and generate a second email by filling in the received rent application form. We found 18 kinds of rent application forms during the preliminary experiment and also during the main experiment. Figure A.12 in Appendix A.3 presents an example rent application form used by clone scammers. In case of other rental scam categories such as “credit report scams” (Chapter 4.3.1), the conversation engine does not send second email to the scammers since the scammers’ first response email usually contain their scam trials.

We ran the automated conversation engine only for the emails selected based on a predefined set of features. During the email conversations, we were able to collect additional data such as email accounts, IP addresses, phone numbers, links and payment information from the scammers. As in chapter 3, the automated conversation engine embedded an external image link into the emails. Once a scammer

Order	Example sentences
First	I'm looking for a place to stay and I saw your ad on Craigslist. I saw your ad on Craigslist and I'd like to ask if you are still looking for a tenant.
Second	I think your home is suitable for my family so I'd like to ask if I would be able to rent your house. Can we meet up so that I can take a look at your house?
Third	Please let me know any time soon. Please let me know if you have any questions about me.

Table 4.2: **Example sentence used in first emails generated by the conversation engine.**

clicks or loads the link in any way, the link leads the scammer to our private web server that logs the visitor's IP address. In this way, we were able to collect the IP addresses of the scammers from two sources: email headers and access logs to the web server.

#### 4.2.4.2 Ethics.

The longitudinal automation phase is the only part of the data collection that involved human subjects. We took care to design our experiments to respect common ethical guidelines and received approval from our institution's IRB for this study. As mentioned above, sometimes we rely on automated conversations to confirm (or disconfirm) whether scams we identify are truly scams. To minimize the inconvenience brought on legitimate users, we abided by the following guidelines. First, we only sent automated emails to ads that we suspected to be scams. Detailed methods are explained in Section 4.3.2 and 4.3.1. Second, we kept the automated conversations to a low volume. In the entire data collection, we sent out 2,855 emails,

from which we received 204 responses that were confirmed to be from scammers out of a total of 367 responses. The rest 163 responses were mostly from apartment leasing offices who posted several rental listings with different rental prices. Email conversation with them were usually terminated after one or two rounds of email exchanges since they usually asked to make phone calls for further help, and this kind of emails were not responded by our automated conversation engine. Although those people were not debriefed at the end of the conversation, we believe that we did not make any serious harm. From these initial results, we were able to improve our methods for detecting suspicious ads, which would further reduce the number of legitimate posters contacted. Finally, in some cases we called the phone number provided by the poster in order to collect additional information. These phone calls were all manually placed, restricted to low volumes and we only contacted suspected scam posters.

#### 4.2.5 Campaign Summaries

We present a high-level summary of the major scam categories and campaigns we identified in Table 4.3. For each campaign we assign it a name based on either the name of the company that is monetizing the scam when known or a feature used to identify the listings in the campaign. Applying our campaign identification methods from Section 4.3, we find seven distinct scam campaigns that account for 32K individual ads. For each campaign the table lists the monetization category of the scam, the raw number of listings associated with that campaign, the percentage

Scam category	Campaign	# Ads	% Flagged	City	Payment
Credit report	CreditReport_Yahoo	15,184	33.0%	20	Credit card
	CreditReport_Gmail	5,472	59.3%	9	Credit card
Rent	Clone scam campaigns	85	87.1%	17	Wire transfer
Realtor service	American Standard Online	3,240	62.4%	19	Credit card
	New Line Equity	3,230	43.3%	12	Credit card
	Search Rent To Own	1,664	77.5%	17	Credit card
Total		28,875	45.2%		

Table 4.3: **Major rental scam campaigns.** Rental scam campaigns of relatively large size in various rental scam types.

of ads that were flagged, the number of cities we found listings in out of the 20 total cities we monitored and the payment method used.

### 4.3 Analysis of Scam Campaigns

In this section, we will present our detailed findings for each campaign, including our insights on how the scams are organized, where they are geographically located and the degree of automation used by each campaign.

#### 4.3.1 Credit Report Scams

In a typical credit report scam, a scammer posts a false rental ad for a property not owned by the scammer. When a victim user replies to the rental ad, the scammer asks the victim to obtain their credit score by clicking on a link included in the email. When the victim clicks the link, a scammer-operated redirection server redirects the victim to a credit score company and includes a referral ID. If the victim pays for

the credit score service which accepts credit card payments, the scammer will be paid a commission by the credit score company through its affiliate program.<sup>2</sup>

#### 4.3.1.1 Data collection.

We identified initial postings for each campaign by manually examining the Craigslist-flagged ads, and correlating contact information and unique substrings included in the postings with user reports found on scam report sites [37–39]. In this manner, we identified two major campaigns, henceforth referred to as *CreditReport\_Yahoo* and *CreditReport\_Gmail* respectively, due to their usage of signature Yahoo and Gmail email addresses.

From the few examples that we found manually, we latitudinally expanded the campaign dataset through human-generated signatures. Using the human generated signatures, we were able to identify additional scam ads from the same campaigns. Craigslist had failed to flag many of the scam ads we identified. Specifically, for *CreditReport\_Yahoo* campaign, we found 15,184 scam ads of which 33.01% were flagged for removal by Craigslist. We also found 5,471 scam ads posted by *CreditReport\_Gmail* of which 59.27% were flagged. More details are provided in Table 4.3.

#### 4.3.1.2 Dataset sanity check.

We verified the suspicious ads identified by the signatures are indeed scams in two ways. First, we performed a sanity check by manually investigating 400 ran-

---

<sup>2</sup>According to the affiliate program of *Rental Verified*, which is used by one of the credit report campaigns we found, it pays up to \$18 per customer. <https://rentalverified.com/affiliates>



domly selected suspicious ads, 200 from each campaign. We considered a suspicious ad as a scam if 1) an ad contained no additional contact information such as name, phone number, street address or URL and 2) there existed same or similar ads with different email addresses in the same campaign. Through the manual inspection, we found only one false positive ad in CreditReport\_Yahoo campaign and two in CreditReport\_Gmail. The email addresses used in the false positive ads were also found in other suspicious ads, and we could also find out actual realtors who used those email addresses. Second, among a total number of 20,256 credit report scam ads we identified, we randomly selected 227 and 89 credit report scam ads from the CreditReport\_Yahoo and CreditReport\_Gmail campaigns respectively, and sent emails in response to the selected ads. Among the emails sent, we received 41 and 78 email responses and all of them were verified to be credit report scams.

#### 4.3.1.3 In-depth analysis.

We present further analysis results of the two credit report scam campaigns. Both credit report scam campaigns appear to be located in the United States. In particular, the CreditReport\_Gmail campaign appears to be located in New York city; while evidence described later (e.g., diverse IP addresses and short inter-arrival times within bursts) suggests that the CreditReport\_Yahoo campaign appears to rely on a botnet for their operation. We now provide an in-depth analysis of the IP addresses and email accounts of both campaigns. Table 4.4 lists the overview of two credit report scam campaigns we found during the experimental period.

	<b>CreditReport_Yahoo</b>	<b>CreditReport_Gmail</b>
Scam signature	“[a-z]+[ ]@[ ]yahoo.com” or “[a-z]+[ ]@[ ]yahoo[ ](dot)[ ]com” No other contact information	“[A-Z][a-z]+[A-Z][a-z]+ @gmail.com” No other contact information
Email account found	14,545 from 15,187 ads	1,133 from 5,472 ads
Affiliated websites	rentalverified.com, matchverification.com	freecreditnation.com, efreescore.com
IP addresses	69	30
IP addresses used once	65 (94.2%)	10 (33.3%)
Country	USA (100%)	USA (100%)
State	28 states	New York (100%)
ISP	Various	Verizon (100%)

Table 4.4: **Credit report scam campaigns.**

**IP address analysis.** For both campaigns, all the IP addresses observed are located in USA. However, two campaigns show completely different IP address usage patterns as shown in Table 4.4.

For CreditReport\_Yahoo, 69 IP addresses were found from 41 email conversations. The number of observed IP addresses are much larger than the number of corresponding email conversations since CreditReport\_Yahoo uses mostly different IP addresses for each round of conversations. In addition, they rarely reuse any IP addresses across different email conversations. 94.64% are used only in a single email conversation, and every IP address is used in at most two email conversations. The IP addresses are distributed over 24 states in USA and mapped back to residential ISPs. These observations, combined with others described later (e.g., level of automation), suggest that this campaign is potentially using a botnet for operation.

In the case of the CreditReport\_Gmail campaign, 30 IP addresses were found

# Emails in burst	Burst duration (sec)	Mean inter-arrival time (sec)	Min/Max/ Median/ Stdev (sec)	# Cities	# IP locations
7	62	10.3	3/24/8.5/7.4	5	7
4	67	22.3	0/57/10/30.4	3	4
4	74	24.7	8/57/9/28	3	4
3	9	4.5	0/9/4.5/6.3	3	3
3	11	5.5	1/10/5.5/6.3	3	3

Table 4.5: **Example inter-arrival time for burst email responses of CreditReport\_Yahoo.** Emails in the same burst have different content, although they contain a similar embedded link to a direction server.

from 78 email conversations. Of the 30 IP addresses, about 66.7% were reused in more than one email conversations and the maximum number of email threads that share the same IP address is 7. All the observed IP addresses of the CreditReport\_Gmail campaign are located in New York City, and map back to a single ISP, *Verizon Online LLC*.

**Level of automation.** We observed many signs of scam process automation, including extremely short inter-arrival time in a burst of emails and duplicate or templated email messages. Table 4.5 lists example email bursts received from CreditReport\_Yahoo campaign. Many email bursts consisting of up to 7 emails were observed and an average inter-arrival time between two emails ranges between 4.5 seconds and 24.7 seconds. Within each burst, emails were always sent from different IP addresses and therefore, usually sent from different cities. This observation also supports the use of the widely-deployed botnet. We also observed many duplicate or templated emails from both campaigns, which are also strong signs of automation. Example email message frequently observed during the whole experiment is shown in Figure A.13 in Appendix A.3.

On the other hand, we also observed signs of manual labor. One example is a distribution of time of day that we received email messages from scammers. In the case of CreditReport\_Yahoo, we never received any email response between 7 PM and 9 AM EST (Eastern Standard Time) and in case of CreditReport\_Gmail, there was no response between 8 PM and 7 AM EST.

### 4.3.2 Clone Scam

In clone scams, typically a scammer copies another legitimate rental ad from a different rental website, e.g., *realtor.com*. The cloned ad typically has the same street address and sometimes has the same description as the original ad. However, often the scammer lowers the rental price. This scam is typically monetized by the scammer requesting a money wire transfer or bank transfer for first months rent and a deposit.

#### 4.3.2.1 Data collection.

To detect clone scams, our crawler tracked rental posts on Craigslist and 5 other major rental websites. We compared these ads and identified Craigslist rental ads cloned from other websites.

Overall, we identified 22,852 cloned ads spanning all 20 cities on Craigslist – however, not all of these are necessarily scam ads. The majority of these appear to be legitimate users advertising their rentals on multiple websites. We then focused on the subset of 2,675 cloned ads with a price difference of at least \$300. These

ads are deemed to be suspicious, but we still cannot be sure whether they are truly scam ads. To verify whether the identified suspicious ads are truly scams, we sent 2, 517 emails to suspicious ads using our automated conversation engines. From the emails we sent, we received 237 responses among which 85 are verified to be scams.

#### 4.3.2.2 In-depth analysis of confirmed scams.

We now report statistics on the 85 confirmed clone scams. Our major insight is that most of these scams originate from Nigeria, and are likely operated by a small number of scam factories. To reach this conclusion, we performed a detailed analysis of the IP addresses, email addresses, wire transfer requests and bank account information contained in the scam attempts. We then performed a clustering algorithm based on identifying information.

**IP address analysis.** Excluding IPs from well-known web mail provider, such as Gmail and Microsoft, we observed a total of 89 unique IP addresses located in 7 countries. We used DB-IP database [40] to geolocate each IP address offline in order to prevent the leakage of the scammer’s IP address information that would result from using an online service. 59 out of 89 IP addresses (66.3%) are from Nigeria and 14 (15.7%) were from the U.S. The result shows fairly similar trend compared to the result of the previous study 3. which shows 50.3% and 37.6% of IP addresses of Nigerian sales scammers were from Nigeria and the U.S. Even though we consider the possibility of proxies or anonymous networks, the consistent results from two studies strongly imply that the major number of the scammers

were actually located in Nigeria.

**Payment Request Analysis.** From our conversations with clone ad posters, we collected a total of 12 unique payment requests and 8 duplicated requests for the same name or bank account. 5 out of 12 unique payment requests (41.7%) are located in the US while 4 (25%) requests are located in Nigeria. For a money transfer via Western Union or MoneyGram, a sender needs to specify the receiver's location information including street address, city and country. However, due to the small sample size of payment requests it is unclear if there is any bias in the subset of conversations that resulted in a payment request versus those for which we were able to collect an IP address.

**Phone number analysis.** We collected a total of 22 distinct phone numbers from 24 email threads. 14 (64%) of the observed phone numbers are registered in the USA, but half of these are identified as VoIP numbers. The rest 8 (36%) phone numbers were registered in Nigeria.

**Clustering.** In order to better understand how scammers are organized, we clustered the emails messages into groups based on similarities of their attributes. We used a conservative clustering strategy. Any two email threads are classified into a same group only if they shared one of the following: exactly the same email accounts, phone numbers, bank accounts, IP addresses or rent application templates. Since those attributes provide us with fairly explicit clues for clustering, we are highly confident of our clustering result. Due to the small size of the dataset, it may not be possible to derive a precise analysis. However, the result implies that these clone scammers may originate from *a small number of scam factories*, just as shown

Group	Ads (%)	Email accounts	Bank accounts	Phone numbers
1	31 (36%)	21	4	9
2	16 (19%)	16	2	3
3	6 (7%)	6	0	2
Others	32 (38%)	29	5	8
Total	85	70	11	22

Table 4.6: **Top 3 clone scam groups.**

in the clustering result of sales scammers in chapter 3.9. Through the clustering, we found a total of 15 scammer groups. Among them the top 3 groups account for 72% of all observed email threads. More detailed information of the top 3 groups are illustrated in Table 4.6. While IP addresses of the second and third groups are largely located in Nigeria, those of the first group are spread over Nigeria, the US, Malaysia and Egypt.

### 4.3.3 Realtor service scam

Realtor service scams involve a special type of realtor service, such as *pre-foreclosure rental* or *rent-to-own rental*. This type of rental is attractive to renters, since they may be able to own the property while paying monthly rent similar to the usual monthly rent of the same area. Realtor service scam campaigns usually request a victim to sign up for a private realtor service to get a list of rent-to-own rentals or pre-foreclosure rentals. To sign up for the service, the victim needs to pay up to \$200 initial fee and/or \$40 monthly fee.

While these businesses actually provide their customers with a list of homes, their rental ads are still considered scams since the ads are typically fake with unreasonably low rent prices, and/or for properties they do not own. Moreover,

many user scam reports claim that in most cases, the properties in the provided list are not even for rent or sale at all. In addition, the refund process is extremely difficult but this is not explained clearly before the customer signs up for their services.

#### 4.3.3.1 Data Collection.

As listed in Table 4.3, we found a total of 8,134 realtor service scam ads over all 20 cities of Craigslist, and about 57% of the ads were flagged by Craigslist. Through the manual inspection on the crawled Craigslist rental ads, we found several phone numbers and URLs observed frequently across multiple cities on Craigslist. We then extended the initial sets of phone numbers and URLs by correlating them with various user scam reports [37–39]. Based on the human generated signatures of phone numbers and URLs, we identified three large realtor services with advance fee campaigns: *American Standard Online*, *New Line Equity* and *Search Rent To Own*.

Among the three campaigns we found, two were identified by sets of phone numbers and the other campaign was identified by a set of URLs. For the soundness of the collected phone numbers, we manually called each number and confirmed a set of numbers actually belong to a same campaign. We confirmed that all phone numbers of a single campaign led us to the same automatic response system. Then we conversed with a representative over the phone and confirmed the business name of each campaign. Table 4.7 lists three large realtor services scam campaigns.



	<b>American Standard Online</b>	<b>New Line Equity</b>	<b>Search Rent To Own</b>
Scam Signatures	20 phone numbers	22 phone numbers	5 URLs
Payment	Initial fee (\$199)	Initial fee (\$9.95), Monthly fee (\$40.95)	Initial fee (\$109.95), Monthly fee (\$39.95)
BBB rating	F	Not found	Not found. (C/F*)

Table 4.7: **Realtor service with advance fee campaigns.** \*: BBB rating of the sibling websites.

#### 4.3.3.2 American Standard Online.

American Standard Online (ASO) was identified based on a total of 20 phone numbers. We gathered the set of phone numbers from our suspicious phone number detection method and many other sources such as *800notes.com*. Using the set of phone numbers, we found 3,240 rental ads posted by ASO over 19 cities on Craigslist. Among them, 62.34% were flagged for removal. Their ads offer rentals with much lower rent prices than other ads in the same area. However, a user is not able to get the information of the property from ASO representatives on the phone.

Because ASO is a registered company in the USA, we could find their record from *Better Business Bureau (BBB)*. BBB website shows that the company ASO has a total of 302 customer complaints and its rating is at the lowest ‘F’. The record obviously tells us that doing business with ASO could be highly risky. This also means that the Federal Trade Commission (FTC) could potentially investigate this company and enforce fines or criminal penalties that would de- monetize this campaign.

According to many user scam reports, the scam process of ASO is as follows. If a victim calls the number to ask about the rental ad, ASO never answers the questions about the rental ads. Instead, ASO requests a payment of \$199 for an initial fee to get an access to their pre-foreclosure (or rent-to-own) property database. Once the victim signs up for the service, ASO provides the victim with a property list. Due to the nature of the term “pre-foreclosure”, it is usually uncertain that the properties in the list are actually in the status of pre-foreclosure, and most of them turns out to be not for rent or sale.

At the time of contract, ASO lures victims by guaranteeing 100% refund after 90 days from the contract in case ASO does not satisfy their customers. However, their actual refund policy requires a wait of at least 90 days from the contract and at least 3 denial letters from the owners of the properties in the provided list. It is obvious that getting the multiple denial letters is extremely difficult.

#### 4.3.3.3 New Line Equity.

New Line Equity (NLE) is an another campaign which provides a special type of realtor service. We identified NLE based on 22 phone numbers observed over 12 cities. Based on the set of phone numbers, a total of 3,230 NLE rental scam ads were identified, and 43.34% of them were flagged.

Many user reports claim that the scam process of NLE is quite similar to that of ASO. A victim calls the number found in a Craigslist rental ads, and NLE requests an initial fee \$9.95 and monthly fee \$40.95. Once the victim makes a payment, NLE

provides him with a list of pre-foreclosure properties. In many cases, however, it turns out that most of the listed properties are not for rent or sale. We could not find a record of NLE from BBB, but there exists a record with a similar business name, *New Line of Equity* which has a BBB rating of ‘D’. Many user reports complain about the difficulties in terminating the monthly fee payment.

#### 4.3.3.4 Search Rent To Own.

We identified Search Rent To Own (SRO) based on five URLs frequently observed over 17 cities on Craigslist. Among the five URLs, one was used as the main URL and the rest were redirection links to the main URL. Based on the set of URLs, we identified 1,664 SRO rental scam ads of which 77.46% of them were flagged. Similarly to the other two campaigns, SRO posts false rental ads on Craigslist and ask the victims to sign up their services with initial and monthly fees. The BBB record of this campaign did not exist but we found the records of two sibling websites listed in SRO website. BBB rating of those two sibling websites were ‘F’ and ‘C’, which are poor ratings for legitimate businesses.

According to the user reports, SRO first lures a customer by offering 3-day free trial service. However, SRO does not fully explain that a \$39.95 monthly fee will be charged automatically after the free trial. We found many customer complains indicating that they were not notified upfront about the fact that monthly fee would be charged automatically after the free trial.

## 4.4 Flagged Ads Analysis

Currently, Craigslist relies on a flagging mechanism to filter out scam and spam ads. Our measurement study reveals that Craigslist currently flags only about 47% of all the scam ads that we identified. Further, for a subset of the scams (specifically, clone scams) that we closely monitored, the median time till flagging (for the ads that do get flagged) is about 13 hours – see Figure 4.1. The figure also shows that roughly 60% of clone scam ads remain active for more than 10 hours and 40% remain active for more than 20 hours.

For other scam categories, our data collection method did not allow us to obtain the time of flagging due to limitations of our measurement study: First, monitoring all ads on a per-hour basis would generate too much traffic, and our experiments were designed to keep our crawler’s traffic volume low. Second, detecting these unknown scams required some manual effort. Hence, for some scam categories, we did not identify the scam ads soon enough to allow us to monitor them on a per-hour basis.

Even though revisiting all ads on a per-hour basis is too aggressive, we were able to revisit all ads we crawled twice after three and seven days to determine whether they have been flagged. Table 4.8 presents a summary of the composition of the Craigslist-flagged ads. Of 126,898 Craigslist-flagged ads, we found about 10.2% are *Scams* where we found concrete proof of scams via automated email conversation. On the other hand, about 70.3% are classified as *Spams* which consists of local ads that are found usually within a single cities and a few renown legitimate real estate

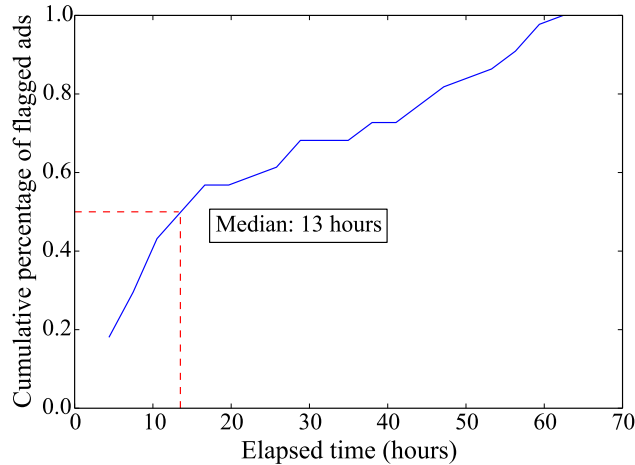


Figure 4.1: **Time taken to flag scam clone ads.** Our system monitored 85 clone scam ads and found that among the flagged ads, only 40% were flagged within 10 hours from ad posting time. In addition, about 60% were flagged within a day.

Category	Campaigns	# Ads (%)
Scams	Credit report scam	8,255 (6.51%)
	Clone scam	74 (0.06%)
	Realtor service scam	4,572 (3.60%)
Spams	Local ads	76,752 (60.48%)
	Credit repair ads	2,234 (1.76%)
	Legitimate rental ads	10,224 (8.06%)
Unidentified		24,787 (19.53%)
<b>Total</b>		126,898 (—%)

Table 4.8: **Flagged ads categorization.** 10.17% of flagged ads are identified as scams while 70.30% are identified as spams.

companies. This leaves 24,787 (19.5%) ads as *Unidentified* where we were not able to ascertain if the ads were benign or malicious. Some of these could be clone ads or other lower volume Rent scams, but they are unlikely to be part of a higher volume template-based campaign based on the diversity of their content.

## 4.5 Discussion

This section will serve to provide a higher level view of our analysis to put into context the value of this study and potential limitations.

### 4.5.1 Potential detection and conversation limitations.

The heuristics we used to detect and validate scam posting were highly accurate based on our analysis. However, it still leaves the question of how many scam listings we did not detect. Without ground truth we cannot provide an estimate for this question. A fair set of assumptions is that our heuristics performed well at detecting the majority of listings associated with larger templated campaigns and worse on the cloned and manually generated rent scam listings, due to the fact that it is difficult to detect these based on the contents of the listing. In spite of this, we do detect some cloned ads and are able to gain an understanding of the structure of their scams from our conversations with these scammers.

### 4.5.2 Similarities and differences with previous study.

Chapter 3 focused on understanding the structure of scammers posing as buyers on Craigslist. The study found that 70% of scammers provided physical shipping addresses located in Nigeria. Furthermore, the study found evidence of a largely manual work force that would respond to scams within 1-2 days during peak work hours in Nigeria.

In this study, we find a diversity of scams that depend on different sets of

infrastructure as well as rent scammers that are structured similarly to those that were encountered in their study. The credit report scams depend on credit report companies in the United States that operate affiliate programs and payout commissions for generating sales. The “bait-and-switch” campaigns depend on rental service businesses that are often incorporated in the United States and accept credit card payments with deceptive refund and re-billing policies.

## 4.6 Related Works

### 4.6.1 Advanced fee fraud.

There have been a number of previous studies that have looked at the structure by Smith [5], Buchanan and Grant [11] and estimated losses from advance fee fraud by Dyrud [23]. Whitty and Buchaman [4] and Rege [24] have investigated the dynamics of online dating scams. More closely related to our domain, Johnson [10] explored the offline methods of real estate scammers. More broadly, Stajano and Wilson [25] created a taxonomy of the different types of psychology motivations used by scammers. Garg and Nilizadeh [26] investigated whether economic, structural and cultural characteristics of a community affects the scams on Craigslist. Tive [27] introduced in his study various techniques of advance fee fraud. Herley [28] has argued that Nigerian scammers deliberately craft their messages to be unbelievable as a method of reducing the number of replies from people that are unlikely to fall victim to these scams. In contrast, our study aimed to be more focused on collecting empirical data to enable a data-driven analysis of rental scams that does not rely

on self reported statistics.

Goa et al. [30] investigates the use of ontology-based knowledge engineering for Nigerian scam email text mining. Isacenkova et al. [29] analyzed public scam email datasets mostly aggregated from numerous user reports. They identified over 1,000 different scam campaigns largely based on phone numbers. Huang et al. [41] measured romance scammer techniques on dating websites. Our work builds on this, but focuses on scammers that are posting fraudulent rental lists targeting people seeking housing on Craigslist. Unlike previous studies, our investigation we have focused on 1) understanding in-depth the *modi operandi* and infrastructure leveraged by rental scammers operating on Craigslist, and 2) identifying methods to detect larger-scale scam campaigns and scammers that are engaged in posting fraudulent rental lists.

#### 4.6.2 Underground studies.

Another large body of recent work has set about conducting empirical measurements to understand the dynamics and economic underpinnings of different types of cybercrime. Much of this work has been focused on spam email [31, 32], illicit online pharmacies [33], and mapping out scam hosting infrastructure [34, 35]. Our work builds on this, but focuses deeply on fraudulent rental lists in particular. We have conducted, to our knowledge, the first large scale empirical measurement study of fraudulent rental lists. It provides us with insights into how these scams are monetized and how they might be better detected in the future.



## 4.7 Conclusion

Rental scams on Craigslist are a real threat encountered by many people searching for housing online; we found about 29K rental scam postings on Craigslist across 20 major cities in 141 days. These fraudulent postings are designed to attract people interested in locating housing and target them with scams tailored to the rental domain. Based on our analysis of these scams we have identified a few potential chokepoints in rental scams that merit further investigation. We also note that analysis of online rental markets in other countries would be beneficial to improving our understanding of rental scams in other locations.

In this chapter, we presented a systematic empirical measurement study of rental scams observed on Craigslist. As part of this study we present techniques that are effective at identifying rental scam postings and classifying them into larger scam campaigns. In parallel, we contacted a subset of these scammers to gain detailed information about the infrastructure required for them to profit. In total we identify seven major rental scam campaigns of which five depend on credit card payments for deceptively advertised services and businesses that are often registered in the United States. Finally, we find that filtering efforts by Craigslist remove less than half of the listings we detected. We believe that our techniques for identifying scam campaigns and understanding of their infrastructure could provide more effective methods for disrupting rental scams.

## Chapter 5: Understanding and Deterring the Business of DDoS Services

### 5.1 Overview

DDoS (Distributed Denial-of-Service)-for-hire services, also known as *booters*, have commoditized DDoS attacks and enabled abusive subscribers of these services to cheaply extort, harass and intimidate businesses and people by taking them offline. However, due to the underground nature of these booters, little is known about their underlying technical and business structure. In this chapter, we empirically measure many facets of their technical and payment infrastructure. We also perform an analysis of leaked and scraped data from three major booters—Asylum Stresser, Lizard Stresser and VDO—which provides us with an in-depth view of their customers and victims. Finally, we conduct a large-scale payment intervention in collaboration with PayPal and evaluate its effectiveness as a deterrent to their operations. Based on our analysis, we show that these booters are responsible for hundreds of thousands of DDoS attacks and identify potentially promising methods to undermine these services by increasing their costs of operation.

## 5.2 Background

In this section we explain the high level business and technical structure of booter services as well as the underlining ethical framework for our measurements.

### 5.2.1 Distributed Denial-of-Service

Denial-of-Service (DoS) refers to an attack attempt to make a target machine or network unavailable to its intended users, that is, preventing legitimate users from accessing the target's online services such as website or web service [42]. The attack is usually conducted by transmitting excessive amount of traffic to the target machine. Distributed Denial-of-Service (DDoS) is a type of DoS attack where multiple attack sources are used. This is a coordinated attack on a target machine or network conducted through many number of compromised machines [45]. DDoS attack can be done more effectively with the use of amplification servers which receive small amount of traffic from an attacker and then send relatively larger amount of traffic to a target. For this purpose, an attacker should compromise amplification servers or locate misconfigured amplification servers. One of typical amplification servers is Domain Name Service (DNS) server. With DNS amplification, the ratio of query size to response size is 70:1– this means that an attacker can generate 70Gbps traffic to a target by just sending 1Gbps traffic to a DNS amplification server.

## 5.2.2 Booter Services

Booter services have existed since at least 2005 and primarily operate using a subscription-based business model. As part of this subscription model, customers or subscribers <sup>1</sup> can launch an unlimited number of attacks that have a duration typically ranging from 30 seconds to 1-3 hours and are limited to 1-4 concurrent attacks depending on the tier of subscription purchased. The price for a subscription normally ranges from \$10-\$300 USD per month depending on the duration and number of concurrent attacks provided. These services claim that they are only to be used by network operators to stress test their infrastructure. However, they have become synonymous with DDoS-for-hire.

These services can be found by visiting underground forums where they advertise and by web searches for terms, such as “stresser” and “booter.” The services are all in English; we did not find any evidence of similar services focused on other markets, such as Asia or Russia. They maintain frontend sites that allow their customers to purchase subscriptions and launch attacks using simple web forms. Their backend infrastructure commonly consists of databases that maintain subscriber information, and lists of misconfigured hosts that can be used for DDoS amplification. Rather than using botnets, most booter services rent high-bandwidth Virtual Private Servers (VPS) as part of their attack infrastructure. Ironically, booter services depend on DDoS-protection services, such as CloudFlare, to protect their frontend and attack infrastructure from attacks launched by rival competing booter services.

---

<sup>1</sup>We use these two terms interchangeably in this chapter.

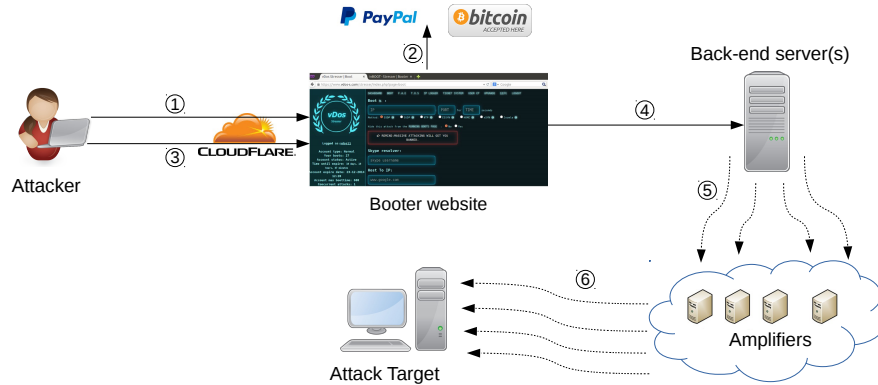


Figure 5.1: **Structure of booter services.**

Figure 5.1 provides a detailed illustration of the infrastructure and process of using a booter service. (1) The customer first locates a booter site and visits their frontend webservice, which is normally protected by CloudFlare. (2) The customer must next purchase a subscription using a payment method, such as Bitcoin or PayPal. (3) The customer then uses the frontend interface to request a DDoS attack against a victim. (4) This request is forwarded from the frontend server to one of the backend attack servers. (5) The backend server then sends spoofed request packets to a set of previously identified misconfigured amplification servers. (6) Finally, DDoS traffic in the form of replies is sent to the victim from the amplification servers.

### 5.2.3 Ethical Framework

As part of the ethical framework for our study, we consulted with our institution's general counsel and placed restrictions on the types of booter services we actively interacted with along with what we included in this chapter. First, we did

not engage with any DDoS service that advertised using botnets and ceased active engagement with any booter that we realized was using botnets. For example, in the case of Lizard Stresser, when we became aware that a botnet was being used, we immediately abandoned plans to collect active attack measurements from this service and restricted ourselves to passive measurements. Our victim server was connected by a dedicated 1 Gbs network connection that was not shared with any other servers. We also obtained consent from our ISP and their upstream peering points before conducting any DDoS attack experiments. We also minimized the attack durations, notified our ISP (Internet Service Provider) before launching any attack and had a protocol in place to end an attack early if it caused a disruption at our ISP.

There were no other methods for us to obtain measurements of their attack infrastructure, such as the set of amplifiers used and rate of usage, except for launching attacks. Our method did create some harm to amplifiers and their upstream peering points by consuming bandwidth resources which we quantify in Table 5.1. The largest amount of bandwidth consumed was 564.56 kbps for Character Generation Protocol (CHARGEN) [43] amplifiers and the least was 3.89 kbps for Simple Service Discovery Protocol (SSDP) [44] amplifiers. Over the course of our experiments we did not receive any complaints from the operators of these amplifiers. Based on our analysis, longer 1 hour attacks only discovered about 20% additional amplifiers over a one minute attack. Given this, we would recommend shorter one minute attacks for future self-attack based experiments to further minimize bandwidth consumed when measuring booters' attack infrastructure.

Type	Avg # of requests	Avg bandwidth
CHARGEN	22.76 (s)	564.56 (kbps)
NTP	1.07 (s)	231.43 (kbps)
DNS	0.71 (s)	12.71 (kbps)
SSDP	0.18 (s)	3.89 (kpbs)

Table 5.1: **Average number of requests per second and average bandwidth consumed in kbps for each amplifier.**

In order to profile the attack infrastructure used by booters and gain insights into how they operate we had to purchase subscriptions. When purchasing a subscription for a booter service, we selected the cheapest option to minimize the amount of money given to these services. In total, we spent less than \$140 and no individual booter service received more than \$19 in payments as part of the measurements in this study. Payments were made primarily using PayPal and we assumed that proper controls were put in place at PayPal to mitigate the risk of money flowing to criminal groups. Also, the 9 booters that overlapped with our payment intervention study likely lost larger sums of money due to our reporting of their PayPal accounts than we paid to them. As part of our design methodology, we minimized the amount of money paid and targeted a small set of booters to obtain valuable measures of their attack infrastructure.

We received an exemption from our Institutional Review Board (IRB), since our study did not include any personally identifiable information and was based on publicly leaked data and scraped data that was publicly accessible. The leaked data contained usernames that did not identify the true names of subscribers, email addresses that again did not directly reveal the real identify of subscribers, and the IP addresses of subscribers used to login into the service. We did not include any raw

data from these leaks or scrapes and we made no attempts to link this information to the real identities of subscribers. The leaks and scrapes also contained victim’s IP addresses. Again, we did not include any raw victim’s IP addresses and we did not mention any victims directly in the chapter. When dealing with publicly leaked and scraped data, our protocol was to create no additional harm from our analysis or what we included in the chapter.

### 5.3 Related Work

DDoS attack and defense techniques have been studied for close to two decades [45–52]. There have also been several empirical studies of DDoS attacks in the wild using backscatter analysis [53] which were revisited by Wustrow *et al.* [54]. More recent studies have measured Network Time Protocol (NTP) based DDoS attacks and conducted broader measurements of UDP amplifiers along with introducing methods to identify spoofing-enabled networks [56].

Other studies have explored the structure of botnet based DDoS attacks [57] and malware [58, 59]. However, the closest related work to ours in this vein is by Welzel *et al.* which measured the impact of DDoS attacks on victims by monitoring two DDoS botnet families and analyzing the service availability of the victims under attack [60] and an analysis of a leaked database from a single booter service, which presents the attack infrastructure and pattern, and the victims of the DDoS attacks, done by Karami and McCoy [61]. Our work differs from this previous work in that we are focused on holistically understanding the stakeholders and infrastructure



these booter services rely on to operate across a larger set of booter services.

Our study is in the same vein as prior work that views security problems through an economic lens [62]. We set out to understand the stakeholders and infrastructure of criminal DDoS-for-hire enterprises as has been done in other domains. Since booters are a criminal support service rather than the previously studied domain of abusive advertising [6, 33, 63], they operate under a different set of constraints. In this respect our work is more along the lines of studies focused on criminal support services, such as email spam delivery [31, 64], fake social links [65] and fake account creation [66].

In Section 4, we show that even though booter services are criminal-to-criminal enterprises their payment methods more closely resemble those of consumer-to-criminal. Using a methodology similar to that presented by Clayton *et al.* [67], we show that there is a concentration of booters that accept payments using PayPal. This indicates that the *follow-the-money* and payment intervention approach which has been demonstrated to be effective in previous studies [7] would be partially effective at undermining booter services.

The largest contribution of our study is in characterizing the ecosystem of subscription-based booter services, which has not been studied in much depth. We show that these booters are structured differently than traditional botnet based DDoS services that are rented for a fixed time period in terms of the underlying attack infrastructure, customer base, business model and payment methods. We believe that our findings enable a better understanding of the effectiveness of ongoing efforts to disrupt their attack infrastructure at the amplifier and hosting level. We

Booter	Period	All Users	Subscribers	Revenue	Attacks	Targets
Asylum Stresser	10/2011 - 3/2013	26,075	3,963	\$35,381.54	483,373	142,473
Lizard Stresser	12/2014 - 1/2015	12,935	176	\$3,368 <sup>†</sup>	15,998	3,907
VDO <sup>‡</sup>	12/2014 - 2/2015	11,975	2,779	\$52,773*	138,010	38,539
Total	-	50,985	6,918	\$91,522.54	637,381	184,919

Table 5.2: **Summary of Asylum Stresser and Lizard Stresser leaked databases and scraped VDO reported data.** <sup>†</sup> Revenue was converted from bitcoin to USD. \*Revenue is estimated based on subscription cost and number of paying subscribers. <sup>‡</sup> Domain name is abbreviated to the first three characters.

also offer a detailed analysis of the nature of these services, how they are structured and a preliminary evaluation of the potential effectiveness of a payment intervention.

## 5.4 Inside View of Booters

In this section, we analyze publicly leaked backend booter databases and scraped data. From this analysis we present some numbers to better understand the dynamics and scale of booter services. This includes, the amount of revenue generated, the number of users, the number of victims and the number of attacks initiated by the subscribers of these services.

### 5.4.1 Data Sets

Our datasets for this section are comprised of two leaked backend databases for Asylum and Lizard Stresser and scraped data from VDO. A summary of these data sets is included in Table 5.2. Before presenting our analysis, we will first describe each of these data sets in more detail.

**VDO Scraped Data.** At the time we started monitoring *VDO* to measure the scale of its operation in early December 2014, it was one of the top booter services on underground forums with a high rate of positive reviews. During an 8 weeks period ending in early February 2015, we crawled this booter on 10-minute intervals to collect data on users of the service and details of attacks launched by them. We found *VDO* to be unique in reporting a wealth of public data on its users and their attack details. This data includes a list of users logged into the service in the last 15 minutes where paid subscribers were distinguished from unpaid users. Also, the booter displayed a list of all currently running attacks and their details including the attack type, target, and duration. Users were able to optionally choose to remain anonymous when logging into the service and hide the IP address or URL of the target when initiating an attack. However, the default option was for all the information to be public and we found only less than 30% of scraped login records to be anonymous and the target was hidden for 39% of all the attack instances we observed during the 8 weeks monitoring period.

While we cannot fully vet this self-reported data, we did verify that the data representing our actions were reported accurately. We also validated that all NTP attacks reported for a day were accurate by sending monlist requests in 10-minute intervals to a set of 12 NTP amplifiers known to be abused by *VDO* and recorded the received responses. A total of 44 distinct victims were the target of NTP attacks as reported by *VDO* during that 24 hour time period and we were able to find matching records for all 44 targets in the monlist responses collected from the set of monitored NTP servers. This gives us some increased level of confidence that the

details of reported attacks and users are accurate.

**Asylum Stresser Backend Database.** Asylum Stresser was an established booter that was in operation for over two years before their backend database — containing 18 months of operational data that included user registrations, payments and attack logs — was publicly leaked. It ceased operation shortly after the leak and has not resumed operation. This leaked database has been vetted by many members of the anti-DDoS community that located their own test accounts in the user registration data and is believed to be authentic.

**Lizard Stresser Backend Database.** Lizard Stresser was launched in late December of 2014 by individuals calling themselves the *Lizard Squad*. This same group was responsible for DDoS attacks on Sony PlayStation and Microsoft Xbox networks on December 25, 2014. Their backend database covering their first two weeks of operation that included user registrations, payments and attack logs was publicly leaked. Since all payments were made in bitcoin and the associated wallet addresses are included, we could validate that the database is accurate. We have also checked for internal consistency within these leaked databases. While we cannot rule out that some of the data has been fabricated, it would take a fair amount of resources to create this high fidelity of a forgery.

## 5.4.2 Subscribers

We find that 15% of Asylum users and 23% of all VDO users purchased a subscription, compared to less than 2% of all Lizard Stresser users <sup>2</sup>. This might

---

<sup>2</sup>Note that Lizard Stresser did not offer free trial accounts.

be attributed to the fact that Asylum and VDO both accepted PayPal payments at least sporadically while Lizard Stresser only accepted Bitcoin as a payment method. It is difficult to attribute why the conversion rate of registered users to subscribers is much less for Lizard Stresser, since other factors, such as the media coverage, might have also driven many users to sign up out of curiosity. The Lizard Stresser's leaked database contains a total of 225 user support tickets. Out of these, 42 are related to user requests for purchasing subscriptions using PayPal. As one potential attacker wrote, "I want to pay via paypal real bad I'm a huge fan of and want to buy this ASAP but I don't have bitcoins."

### 5.4.3 Revenue

Asylum collected 99.4% (\$35,180.14) of their revenue through PayPal payments and only 0.6% (\$201.40) of their revenue was collected using their secondary payment method of MoneyBookers. Lizard Stresser collected all their revenue through their only supported payment method of Bitcoin and VDO accepted both PayPal and Bitcoin. They are presumably profitable, but these individual booters do not generate the profits required to pay the upfront capital, fees and potential fines for dedicated credit card merchant processing accounts. This amounted to around \$25-\$50K per an account, as was the case with illicit pharmaceutical and fake anti-virus groups that had revenues on the order of millions of USD dollars a month [33, 68].

#### 5.4.4 Attacks

From the leaked data we find that these three booters were responsible for over half a million separate attacks against over 100,000 distinct IP addresses. While the average attack from VDO only lasted 27 minutes, this data demonstrates the large-scale abuse problems and unwanted traffic generated by these services. Our analysis of victims finds that they are predominantly residential links and gaming-related servers, with a small number of higher profile victims, such as government, media and law enforcement sites. This matches previous analysis of victims from leaked databases [61]. For VDO our scraped data included the type of DDoS attack launched and our analysis of this data shows that amplified attacks, where the adversary attempts to exhaust the bandwidth capacity of the victim's connection, accounted for 72% of all attacks launch from VDO. The next most popular class of attacks were SYN flooding attacks, which made up only 16% of all attacks.

#### 5.5 Attack Infrastructure

Our measurements of booters' attack infrastructure are based on engaging with these services to understand what techniques and hosts are being actively used for attacks. Using this information might better inform defenders as to which ISPs and hosts to focus on for blacklisting, remediation and notification efforts. Our analysis of frontend servers finds a reliance on CloudFlare to protect booter's infrastructure from takedown and DDoS. In addition, we find that booters gravitate to using more stable amplifier infrastructure when possible. This differs from previous studies

that scan the Internet for the vulnerable populations of misconfigured amplification servers many of which might be transient and not be used for DDoS attacks. We also identify two hosting providers connected to the same ISP that are actively courting booter operators and providing stable high bandwidth attack servers that allow spoofing.

### 5.5.1 Data Set

Our first task was to identify booter services for this part of our study. Absent a centralized location for finding booters, we found services via search engines and advertisements on hacker forums. We selected 15 booter services that received the most positive feedback on underground forums for our attack infrastructure characterization. The number of booters was kept relatively small in order to minimize the amount of money we paid to these services. We make no claim about the coverage these booters provide of the entire ecosystem. Rather, we were looking to provide a sample of stable services ranked highly for search terms associated with booter services. In addition, these booters garnered the most positive replies to their advertisements on underground forums.

We purchased a one month subscription from each of the services which ranged from \$2.50-18.99 and focused on measuring amplification attacks based on our measurements of VDO that showed it was the most common type of attack. In addition, amplification attacks were the default attack type for all 15 booters. More precisely, we chose to measure the most common amplification reflection attack types offered

by the booters, which were SSDP, NTP, DNS and Chargen. Table 5.3 shows the set of booters, the four attack types that booter offered and the cost of a basic month subscription.

We conducted attacks directed at our target server from December 2014 - January 2015. The goal of these attacks was to map out the set of misconfigured hosts that were being used by each booter to amplify their reflection attacks. The configuration of our target system used for measuring the attacks was an Intel Xeon 3.3GHz server running Ubuntu with 32 GB of RAM and an isolated 1 Gbps dedicated network connection.

We used `gulp` [69], which is a lossless Gigabit packet capture tool to capture attack traffic. Each attack lasted for one hour total and was comprised of many shorter attack instances of 10 minutes each, which is the standard time limit for basic subscriptions. The reasoning behind the longer attack times was to increase our probability of identifying all the misconfigured reflection hosts used by a booter for each attack type.

### 5.5.2 Frontend Servers

Booter services maintain a frontend website that allows customers to purchase subscriptions and launch DDoS attacks using convenient drop-down menus to specify the attack type and victim's IP or domain name. These frontend websites commonly come under DDoS attack by rival booters and are subject to abuse complaints from anti-DDoS working groups. All 15 booters in our study use CloudFlare's DDoS



<b>Booter</b>	<b>Attack Types</b>	<b>Cost</b>
ANO	DNS	\$6.60
BOO	NTP,Chargen	\$2.50
CRA	DNS,SSDP	£10.99
GRI	NTP,SSDP	\$5.00
HOR	NTP,SSDP	\$6.99
INB	DNS,NTP,SSDP	\$11.99
IPS	NTP,SSDP,Chargen	\$5.00
K-S	SSDP,Chargen	\$3.00
POW	SSDP	\$14.99
QUA	DNS,SSDP	\$10.00
RES	DNS,NTP	\$10.00
SPE	DNS,NTP,SSDP,Chargen	\$12.00
STR	DNS,SSDP	\$3.00
VDO	DNS,NTP,SSDP	\$18.99
XR8	DNS	\$10.00

Table 5.3: **List of booter services we measured, the attack types offered, and the cost of the least expensive one-month subscription.**

protection services to cloak the ISP hosting their frontend servers and to protect them from abuse complaints and DDoS attacks.

As part of this study, we contacted CloudFlare’s abuse email on June 21st 2014 to notify them of the abusive nature of these booters. As of the time of writing this dissertation, we have not received any response to our complaints and the still active subset of booters continue to use CloudFlare. This supports the notion that at least for our set of booters CloudFlare is a robust solution to protect their frontend servers. In addition, *crimeflare.com* has a list of over 100 booters that are using CloudFlare’s services to protect their frontend servers.

### 5.5.3 Attack Servers

Renting back-end servers to generate attack traffic is the primary expense for operators of booter services. We did some research to get a broad sense of the

<b>Provider</b>	<b>VPS IP</b>	<b>Uplink speed</b>	<b>Bandwidth</b>	<b>Monthly cost</b>
CaVPS Host	192.210.234.203	3.5 Gbps	Unmetered	\$35
Spark Servers	96.8.114.146	949 Mbps	10 TB	\$60

Table 5.4: **Spoofing enabled VPS services.**

market availability and cost of back-end servers that allow the source IP address to be spoofed. Being spoof friendly, fast uplink speed and high caps or unmetered bandwidth usage are the key requirements of a server appropriate for supporting the operation of a booter service. Providers of spoof friendly Virtual Private Servers (VPS) can be located on the same underground forums as where booters advertise their services. These VPS providers often explicitly advertise the ability to spoof source IP addresses as one of their key features.

In order to understand if these services delivered on their claims of allowing spoofing and providing the bandwidth they advertised we rented VPS from two hosting providers that advertised on underground forums. We rented one of the spoof-friendly virtual servers directly from a booter service included in our study.

Table 5.4 summarizes the services that we purchased. Both of the VPSs we purchased were connected to the same ISP (ColoCrossing) in the US. We also verified that both VPSs allowed spoofing and measured their actual link speeds. One VPS provided around 1Gbps uplink bandwidth and the other one provided up to 3.5Gbps. Due to budget limitations we could only rent these two VPSs and did not rent any higher end dedicated servers. However, our initial results show that this is a potentially effective method of mapping out abusive hosting and we plan to scale this part of our measurements as future work.

Booter	Chargen		DNS		NTP		SSDP	
	#	Overlap %	#	Overlap %	#	Overlap %	#	Overlap %
ANO	-	-	1,827	73%	-	-	-	-
BOO	370	65%	-	-	1,764	86%	-	-
CRA	-	-	43,864	56%	-	-	64,874	46%
GRI	-	-	-	-	1,701	72%	10,121	60%
HOR	-	-	-	-	8,551	58%	242,397	30%
INB	-	-	38,872	55%	4,538	92%	170,764	54%
IPS	1,636	44%	-	-	1,669	85%	90,100	29%
K-S	1,422	30%	-	-	-	-	5,982	76%
POW	-	-	-	-	-	-	1,424,099	11%
QUA	-	-	10,105	85%	-	-	39,804	67%
RES	-	-	2,260	82%	27	100%	-	-
SPE	2,358	38%	26,851	61%	6,309	35%	258,648	24%
STR	-	-	93,362	53%	-	-	7,126	74%
VDO	-	-	16,133	82%	6,325	82%	150,756	62%
XR8	-	-	44,976	52%	-	-	-	-
Total	4,565	23.46%	181,298	35.30%	17,599	42.31%	2,145,015	11.84%

Table 5.5: **Number of total amplification servers and percentage of overlap with amplification servers used by other booters.**

#### 5.5.4 Attack Techniques

Due to their effectiveness, amplified volume-based attacks are the default attack technique offered by most booter services. We focused our analysis on SSDP (more commonly known as Universal Plug and Play (UPnP)), DNS, NTP and Chargen. These attacks depend on servers running misconfigured UPnP, DNS resolvers, NTP and Chargen services that enable attackers to amplify attack traffic by sending spoofed packets with the victim’s source address in the IP header and having these services respond with a larger amount of traffic directed to the victim.

### 5.5.5 Amplifiers

As part of our measurements we can map out the set of amplifiers that are being abused to magnify the traffic volume of attacks. This sheds light on the population of hosts that are not only potential amplifiers, but are actively being used as amplifiers for DDoS attacks. Table 5.5 shows that the set of abused Chargen and NTP servers are smaller and more highly shared between two or more services, whereas there is an ample supply of DNS and SSDP servers that are used as amplifiers. However, the overlap of DNS servers used by two or more booter services is still relatively high suggesting that these DNS resolvers might be more stable, have higher bandwidth connections and be in more limited supply.

### 5.5.6 Amplifier Location

As demonstrated by Table 5.6, both the geolocation and AS of amplifiers used by booters are fairly diffuse. There are a few notable exceptions, such as the concentration of Chargen amplifiers in China with three Chinese ASs connecting 34% of these amplifiers. In addition, there is a slight concentration of abused NTP servers connected to one Taiwanese AS and two United States network operators. This might indicate a potential to focus notification and patching efforts on these networks, given the limited pool of hosts used for Chargen and NTP attacks from Table 5.5. Feeds of these actively abused servers could also be distributed to these network operators and to DDoS mitigation services.

CC	%	AS	%
<b>Chargen</b>			
CN	48.78%	4134 (Chinanet)	14.46%
US	12.51%	37963 (Hangzhou Alibaba Advertising)	10.47%
KR	5.50%	4837 (CNCGROUP China169 Backbone)	6.88%
RU	4.58%	17964 (Beijing Dian-Xin-Tong Network)	2.61%
IN	2.56%	7922 (Comcast Cable Communications)	2.61%
<b>DNS</b>			
US	12.38%	4134 (Chinanet)	2.68%
RU	11.58%	3462 (Data Communication Business Group)	2.15%
BR	9.19%	18881 (Global Village Telecom)	1.46%
CN	6.84%	4837 (CNCGROUP China169 Backbone)	1.45%
JP	3.61%	7922 (Comcast Cable Communications)	1.27%
<b>NTP</b>			
US	31.47%	3462 (Data Communication Business Group)	14.01%
TW	15.29%	46690 (Southern New England Telephone)	12.35%
CN	10.68%	7018 (AT&T Services)	4.84%
KR	5.50%	4134 (Chinanet)	3.58%
RU	4.74%	4837 (CNCGROUP China169 Backbone)	2.18%
<b>SSDP</b>			
CN	36.26%	4837 (CNCGROUP China169 Backbone)	18.98%
US	19.37%	4134 (Chinanet)	11.16%
EG	6.83%	8452 (TE Data)	6.61%
AR	5.37%	22927 (Telefonica de Argentina)	5.13%
CA	5.36%	7922 (Comcast Cable Communications)	4.60%

Table 5.6: **Top country locations and autonomous systems for amplifiers.**

### 5.5.7 Amplifiers Churn

In order to measure the stability of these amplifiers we probed them periodically for 13 weeks to understand how many were still located at the same IP and misconfigured. As shown in Figure 5.2, the set of DNS resolvers were the most stable with nearly 80% still misconfigured and located at the same IP after one month, and over 60% were still misconfigured after 13 weeks. This result is counter to the previous results of churn based on Internet wide scanning that found a 50-60% churn rate for DNS servers after one week [56]. It potentially indicates that booters have gravitated to using a more stable set of DNS resolvers and that focusing mitigation

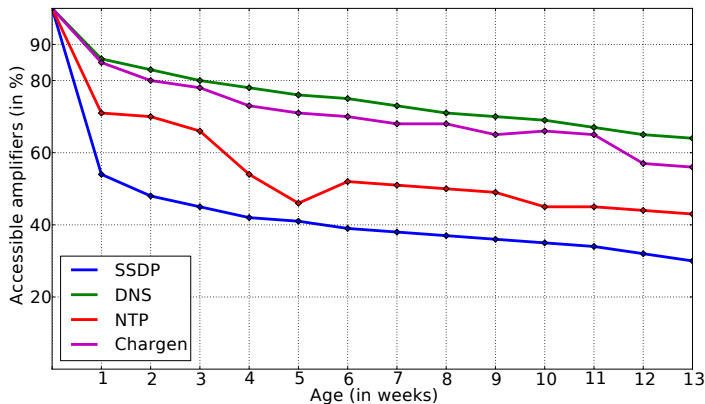


Figure 5.2: **IP churn of amplifiers.**

efforts on these might cause these DNS attacks to be less efficient and require additional bandwidth and cost. However, our measurements were collected after those in the previous study making direct comparisons challenging.

### 5.5.8 Amplification Factor

One of the few direct costs incurred for every attack a booter service launches is the bandwidth sent from their rented attack servers. In order to reduce this cost amplification attacks are used for volume-based flooding attacks. Our amplification factor measurements largely agree with the lower-end bandwidth amplification factor numbers reported in a previous study [56], with NTP attacks resulting in an average amplification factor of 603 times. Chargen was the next largest at 63 times, DNS resulting in an average of 30 times amplification and SSDP generating the smallest average amplification factor of 26 times. This and the limited number of NTP amplifiers confirms that the communities focus on prioritizing notification and patching of misconfigured NTP servers is the correct approach. We also suggest

that some effort be placed into notifying operators of servers with misconfigured and abused Chargen services, since these are the next largest threat and there is also a potentially constrained supply.

## 5.6 Payment Intervention

As part of our study we sought out opportunities to understand and also measure the effectiveness of undermining DDoS Services. In this section, we present our measurements of a payment intervention that was conducted in collaboration with PayPal and the FBI.

We find that reporting accounts to responsive payment service providers, such as PayPal, can have the desired effect of limiting their ability and increasing the risk of accepting payments. This technique requires constant monitoring of the booters and drives them to move to more robust payment methods, such as Bitcoin.

### 5.6.1 Payment Ecosystem

At the onset of our study, the majority of booter services accepted credit card payments via PayPal as their primary mechanism for receiving funds from their customers. In addition, some booters accepted bitcoin payments and a limited number accepted credit card payments using Google Wallet<sup>3</sup> and virtual currencies, such as WebMoney and Perfect Money. These last two prohibit customers from the United States from opening an account and using their platform.

---

<sup>3</sup>Google phased out their digital goods payment processing at the start of March 2015 — <https://support.google.com/wallet/business/answer/6107573>.

We identified a larger set of 60 booter services <sup>4</sup> that accepted PayPal and created custom crawlers to monitor their payment methods and merchant accounts for about 6 weeks from April 22, 2014 through June 07, 2014. To receive their payments using PayPal, booter services redirect customers to the PayPal website where existing PayPal users can login and complete a transaction. After logging into PayPal, our crawlers were able to collect the merchant account identifier of the corresponding booter service from the HTML source of the page without completing a transaction.

The set of booters selected for monitoring were located from underground forum advertisements and web searches for terms commonly associated with booter services. Again we make no claim about the coverage these booters provide of the entire ecosystem. To minimize the effect of unstable booters on our study, the final set of booters included in our analysis was limited to the 23 stable booter services that were able to successfully use PayPal to receive funds for at least half of the time before the PayPal intervention and used at least one PayPal account after the intervention. Nine of these 23 booter services overlapped with the set of 15 booters measured in section 5.5. Among the reasons that six booters were not included in this measurement is that some did not accept PayPal and others did not accept PayPal over half the time in the first 6 week period.

After collecting our initial data on the stability of their PayPal merchant accounts, we reported these booter's domains and accounts to PayPal. The or-

---

<sup>4</sup>This set is larger than the previous set, since we did not have to pay for a subscription in order to monitor their payment accounts.



ganization then began to monitor merchant accounts linked to these domains and suspended them after an investigation. Note that PayPal will initially limit merchant accounts that are found to violate their terms of service by accepting payments for abusive services until they perform an investigation of the account. Once an account is limited the merchant cannot withdraw or spend any of the funds in their account. This will result in the loss of funds in these accounts at the time of freezing and potentially additional losses due to opportunity cost while establishing a new account. In addition, PayPal performed their own investigation to identify additional booter domains and limited accounts linked to these domains as well. This had the affect of a large-scale PayPal payment disruption for the majority of booter services.

In order to further understand the effectiveness of our payment intervention, we monitored underground forums where these booters advertise their services and news feeds from booters we joined to discover qualitative data on the effectiveness of PayPal's payment intervention.

### 5.6.2 Usage Pattern of PayPal Accounts

Based on our observations, booter services will generally use only one PayPal account at a time to receive payments. Once a limit is put on an account, they will change it. At times, they will also proactively change accounts to reduce the risk of having limits imposed. We used the dataset collected during the initial monitoring period to understand how frequently booter services were changing their PayPal

Booter	accounts before	accounts after	Status
ANO *	6 (8.2)	7 (2.9)	✓
AUR	6 (7.2)	6 (2.7)	✗
BOO •	6 (8.3)	11 (2.7)	✓
CRI †	4 (9.0)	1 (2.0)	✗
DAR ◊	4 (6.0)	5 (4.8)	✓
DIA †	3 (15.7)	0 (-)	✗
GET	2 (14.0)	1 (4.0)	✓
GRI *	4 (10.5)	1 (6.0)	✓
HAZ ◊	4 (12.2)	5 (5.6)	✓
IDD ◊	3 (7.7)	2 (9.0)	✓
IPS ◊	3 (7.3)	5 (5.4)	✓
POW	5 (4.5)	9 (5.0)	✓
PRI	6 (8.8)	2 (1.0)	✗
QUA	11 (4.3)	22 (1.8)	✓
RAG •	13 (3.9)	4 (2.0)	✓
REB	2 (11.5)	9 (2.3)	✗
RES *	6 (8.2)	7 (2.9)	✓
SNO	1 (-)	0 (-)	✗
STA	5 (13.0)	3 (5.3)	✓
STR	1 (47.0)	1 (4.0)	✓
TIT ◊	12 (5.3)	17 (2.9)	✓
XR8	4 (10.5)	11 (1.6)	✓
XRS •	8 (5.2)	4 (2.5)	✗
	119 (7.84)	133 (3.07)	

Table 5.7: **Number of PayPal accounts used by monitored booters before and after the intervention.** The numbers within the () are the average lifespan (in days) of the accounts used by that booter. Accounts that are active both before and after are counted only in the before and not included when computing the average lifespan. Matching symbols indicate that this set of booters shared at least one PayPal account. These shared accounts might be instances of a third party agreeing to accept payments for these services.

accounts. Note that our age measures are both right and left-censored. For the booter’s initial account our data is left-censored and for the last account our data is right-censored. However, we believe our age measurements accurately represent the effects of the PayPal intervention based on our interactions with and postings from the booters themselves.

Table 5.7 provides an overview of the PayPal accounts observed by our crawler

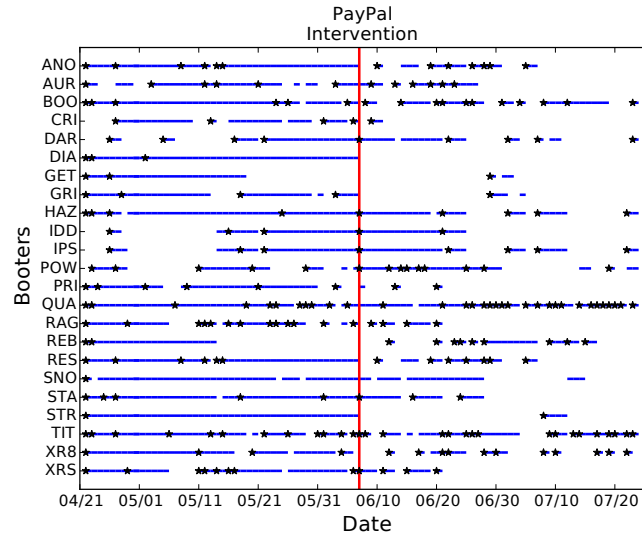


Figure 5.3: **PayPal account usage over time.** Black asterisks denote a new PayPal account and gaps in the blue line represent PayPal unavailability for that time period. The red vertical line indicates when the reporting of accounts started.

broken down by each service monitored. As Table 5.7 shows, accounts had an average lifespan of slightly over a week before the intervention with STR and SNO using a single account that remained active during the entire 47-day initial observation period and SNO’s account remaining active for 37 days after the intervention began. On the other end of the spectrum, QUA, RAG and TIT changed accounts every 4-5 days before the intervention. The impact of the intervention can be visually seen in Figure 5.3.

Once the target intervention begins the average lifespan of an account drops to around 3.5 days with many booter’s PayPal accounts only averaging around two days before they are no longer used again. Figure 5.3 visually shows the impact of the payment intervention on the lifespan of booter’s PayPal accounts and provides some indication of the time period that elapsed between a new PayPal account being actively used to accept payments and it ended either due to PayPal action

or a booter’s proactive replacement. The length and number of PayPal outages increase after the intervention, with only QUA and TIT avoiding major PayPal outages by resorting to aggressively replacing accounts. Note that this replacement strategy was not fully effective, since our monitoring infrastructure detected and reported these accounts.

We use the Kaplan-Meier estimator to compare the lifespan of PayPal accounts before and after the intervention. The lifespan duration of an account is defined as the time difference in days between the first usage of the account and its last date of usage. The accounts that were first seen before the intervention date and were still in active use on the intervention date are labeled as censored. The same applies to the accounts used by booter services in the time period after the intervention and still in active use at the end of the data collection period (07/24).

Figure 5.4 shows that the lifespan decreased after the intervention. The shaded areas represent the 95% confidence bounds. The result of a log-rank test with 99% confidence limits indicates a significant difference between the two survival curves.

### 5.6.3 Booters’ Status

As part of our daily monitoring of the 23 booter services, we recorded if the service could accept PayPal payments and if the site was functional. This enabled us to better understand the impact of the payment intervention on the booter’s ability to accept PayPal payments and operate the booter. Each booter was categorized by one of the following statuses each day, based on the results of our crawl.

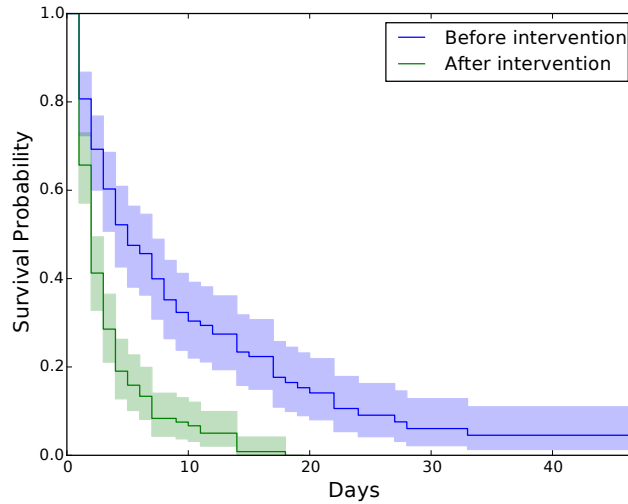


Figure 5.4: **Lifespans of PayPal accounts before and after the intervention.**

**Active:** The booter is able to successfully use a PayPal account to receive payments from its customers.

**Unreachable/Broken:** Either the booter’s frontend website was not responding to HTTP requests, the booter service had closed or the frontend site was not functional.

**PayPal Disabled:** The booter’s frontend website is active, but the service has either removed PayPal as a payment option, or the PayPal account linked to the booter website is limited and therefore unable to receive payments.

Figure 5.5 shows the status of booter services over time. The vertical line represents the date on which we started sharing our data with PayPal and PayPal started to independently investigate the reported accounts and take action against them. As observed in Figure 5.5, the percentage of active booters quickly drops from 70-80% to around 50% within a day or two following the intervention date and continues to decrease to a low of around 10%, before fluctuating between 10-30%. This resulted in an increase in PayPal unavailability from 20% before the

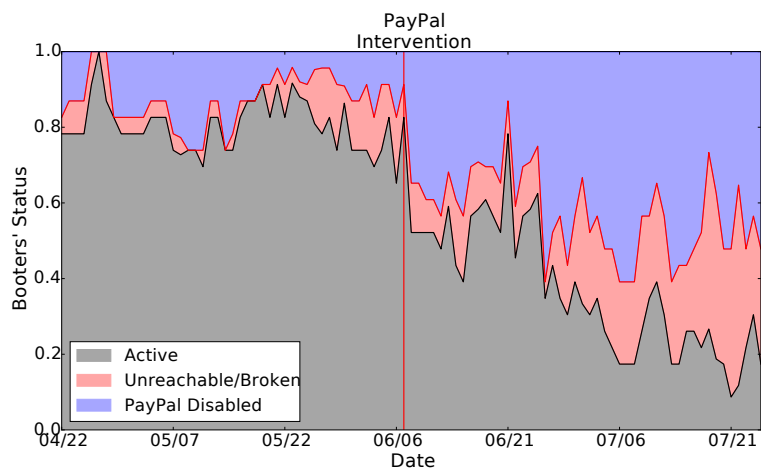


Figure 5.5: **Status of booters over time.**

intervention to 63% during the intervention. In addition, we observed 7 booter services in our study shut down their businesses and most of the remaining services switched to alternative payment methods, such as Bitcoin.

#### 5.6.4 Qualitative Assessments

In addition to our quantitative measurements, we also have qualitative evidence of the efficacy of PayPal’s payment intervention. By monitoring the underground forums where these services advertise we can witness the impact of these account limitations. Wrote one booter operator during the intervention, “So until now 5 time my 5 PayPal Accounts got Limited on My stresser is other stresser have same Problem with the f\*\*\*ing Paypal ? is there any solution what we should do about f\*\*\*ing Paypal ?” Similarly, customers vented their frustration at being unable to purchase a booter service using PayPal. Wrote one booter customer, “when i go to buy a booter it normally says i can’t buy because their PayPal has a problem.”

In a number of cases, booters directly link their closures to loss of funds due

to PayPal merchant account limitations. This message was posted on the front page of a defunct booter service, “It’s a shame PayPal had to shut us down several times causing us to take money out of our own pocket to purchase servers, hosting, and more.”

### 5.6.5 Booter Response

As with any intervention, the target, in our study booters, will respond by adapting to the pressure. In this case, we do not have enough quantitative measurements to assess the effectiveness or the full range of responses to our attempt to undermine booter’s payment infrastructure. However, we have identified several common classes of adaptations in response to the intervention.

**Alternate payment methods.** Most booters have added Bitcoin as an alternate payment method and have posted links to services that allow customers to purchase bitcoins using credit cards or PayPal. In addition to bitcoins some switched to Google Wallet and others added the option to pay using virtual currencies, such as Webmoney and Perfect Money. By all accounts, these have resulted in reduced customer bases unless the booter can directly accept credit card payments. Evidenced by the fact that many booters continued to replace their PayPal accounts even when previous ones were limited and their funds were lost. Assuming that alternative payment methods did not result in reduced revenue or higher costs there would be little incentive to continue using risky payment methods, such as PayPal.

**Referrer anonymizing services.** We have noticed that some booters have stopped

directly linking to PayPal and are now linking to an intermediary site and then redirecting the customer from this intermediary domain to PayPal's site. This intermediary redirection site is used to hide the booter's real domain name in the referrer field from PayPal. A subset of booters have also started to replace this intermediary domain every time they replace a PayPal account. The effect of this is that it requires active crawling and measurements of booter sites to identify a new PayPal account. This bypasses passive methods PayPal could use to linking accounts, such as by using referrers, which results in increasing the difficulty of monitoring booter's PayPal accounts and effort required to investigate these accounts.

**Offline payment.** Finally, in some cases booters have required customers to open a ticket to pay using PayPal. This method increases the effort to monitor the booter for new accounts, since instead of an automated crawler someone must now interact with the booter service manually. However, this method also requires the booter service to manually activate each account and drives away customers that are seeking automated subscription purchasing systems.

## 5.7 Discussion and Future Work

We have gathered a few key points from ours and the community's efforts to understand and undermine these DDoS services. Most of these potential strategies involve driving up costs of operating booter services and reducing the convenience of subscribing.

**Reducing scale and accessibility.** Limiting access to convenient payment meth-



ods, such as PayPal, had an impact on the scale of booter services based on our quantitative and qualitative analysis. This also results in a financial loss to the cybercriminals since they cannot withdraw the funds in their limited accounts at least during a certain amount until PayPal completes an internal investigation on the limited accounts.

In addition, disabling PayPal payment system of booter services may be able to result in reduced accessibility to the booter services. Our analysis of the revenue generated via PayPal and other payment methods (e.g., Bitcoin and MoneyBookers) implies that the customers of the booter services prefer PayPal. Hence, the intervention on PayPal payment system of the booter services would result in limited accessibility to their services.

However, based on the short duration of the intervention it is unclear if this approach would continue to be effective in the longer-term. As future work, we plan to understand how to improve the effectiveness of these interventions and make them sustainable. This in part requires developing more robust monitoring tools that better mitigate countermeasures being deployed to make their payment methods more robust to interventions.

**Reducing effectiveness of attacks.** We plan to continue our monitoring efforts of the amplification servers used by booters and begin sharing this information with existing patching efforts, such as the OpenResolverProject [70] and OpenNTPProject [71]. Along with this, we plan to experiment with active notifications sent to the ISP and abuse contact for the server. There is some indications that active notification improves patching rates in context of patching vulnerable services [72, 73].

**Increasing costs.** This might be achieved with an increased effort to locate and blacklist or de-peer low-cost hosting services that cater to DDoS attacks by providing the ability to spoof and unlimited bandwidth. This might force these services to pay a premium for bullet-proof hosting attack servers, which would result in reduced profitability or be passed along to subscribers in the form of increased subscription costs.

In addition, convincing CloudFlare and other free anti-DDoS services to prohibit these booter services would increase their costs by forcing them to build and pay for anti-DDoS services that cater to these abusive booters. Admittedly these suggestions will likely not result in large cost increases unless tremendous amounts of pressure were placed on these parts of their infrastructure.

**Increasing risk to operators.** Our analysis of data provided by PayPal suggests that much of this activity is occurring in the United States. If this is the case there is the potential that increased law enforcement efforts could have a direct impact in arresting key operators of these services and increasing the perceived risk of operating and using these services. In the case of operators it is likely they could be replaced by overseas operators. However, in the case of customers it might be difficult to find a new subscriber based for these services that is located outside the United States and Western Europe if the perceived risk of using these services increased. To this end we plan to work with law enforcement to understand how effective this type of intervention is on mitigating the threat of booter services.

## 5.8 Conclusion

Unfortunately, there is no silver bullet that will mitigate the threat posed by booter services overnight. These booters have grown in scale due to the perceived low-risk nature, their profitability and increasing demand for DDoS attacks.

In this chapter we have mapped out a range of support infrastructure that booters depend on in terms of advertising, attack, hosting and payment. We demonstrated that payment interventions, which undermine the accessibility of convenient payment methods, such as PayPal, can potentially have an impact on reducing the scale of these services. Our hope is that by continuing to explore new methods for understanding and undermining booters, we can identify increasingly effective methods of adding friction, cost and risk to these ventures that further erodes their scale and profitability over time.

## Appendix A: Scam Examples

### A.1 Example fake payment scam emails

Frequently Observed Emails	
[Body:]	HELLO..IS THE ITEM HELLO, IS THE ITEM POSTED ON CL LISTED ABOVE STILL FOR SALE?? KINDLY GET BACK TO ME WITH THE LAST PRICE AND PRESENT CONDITION.THANKS
# Times observed:	295
[Body:]	HELLO, IS THE ITEM POSTED ON CL LISTED ABOVE STILL FOR SALE??
# Times observed:	293
[Body:]	Is it still available?
# Times observed:	182
[Body:]	Good day as i come across your listing on craigslist and i would like to know if its still for sale.
# Times observed:	180

Figure A.1: **Recurring fake payment scam emails.** Usually observed in first scam responses.

### Belligerence/Threats

- [words omitted] Please respond to this mail before the penalty decision is taken against you. You are warned. We are waiting for your mail before we can credit your account and this is due to the large increase in the rate of the online scams recorded in the previous years. [words omitted]
- [words omitted] i think if i did not hear back from you within next 24hrs iwill have to contact FBI about your actions on Craigslist.org [words omitted]
- [words omitted] i will report you to paypal an FBI i give you 12h to get it ship [words omitted]
- Hey man what is going on i getting the FBI involve in this; is getting irritating [words omitted]

Figure A.2: **Sample emails with belligerent tones.** After making fake PayPal payment, scammers urge the victim to send the goods immediately. Note that some sentences are omitted for brevity or to remove personally identifiable information.

### Broken Subject & Body

[Subject:] <span class=i h data-id=0:00x0x\_6FP1-zGRnuHX> </span>  
[Body:] Is your <span class="i h" data-id="0:00x0x\_6FP1zGRnuHX"></span> still available for sale??  
i will reply right away. Thanks  
Sent from Devon's iPhone

[Subject:] <span class=i h data-id=3Fa3Le3I45Nd5I-c5Gcd624761cea879bf1558.jpg></span>  
[Body:] <html><head><META http-equiv="Content-Type" content="text/html; charset=utf-8"></head>  
<body>still for sale?; feel free to email me at darrenamos69@gmail.com  
</body></html>

Figure A.3: **Sample broken subject and body lines.** Similar broken subjects were observed 3316 times in total. This observation implies scammers might be using some automated tools.

### Email Bursts

[Body:] I'm interested in buying the posted item & your price Get back to me with your email welchcarrie619@gmail.com

Date/Times: 6/23/2013 3:04 - 4:47 PM

# Times observed: 36

[Body:] I'm interested in buying the posted item & your price Get back to me with your email sanjossmith@gmail.com

Date/Times: 6/29/2013 2:08 - 3:57 PM

# Times observed: 14

[Body:] I'm interested in buying the posted item & your price Get back to me with your email robert.waddick@gmail.com

Date/Times: 6/30/2013 3:56 - 6:04 PM

# Times observed: 19

Figure A.4: **Sample recurring emails bursts.** Bursts of emails are frequently observed. This observation also implies scammers might be using the automated tools.

### Curses Indicating Manual Operation

- [curse word omitted] you stupid scammer. [words omitted]
- ARE YOU TRYING TO SCAM ME OR WHAT?
- WTF are you sending to me again ? i have already transfer the money into your PayPal Account and the money has already been deducted from my account;get the iphone 5 ship out via usps express mail and get back to me with the tracking number immediately you shipped. [words omitted]

Figure A.5: **Sample emails indicating manual operation.** In some cases, scammers detected us and sent this kind of curses in second responses.

### Invoking God for Sincerity/Empathy

- [words omitted] I need you to be honest with the sale as I am a God fearing person. [words omitted]
- [words omitted] Note: I will be paying you extra money to cover the shipping cost through USPS EXPRESS MAIL. Also i wanted you to consider this sold to me and please remove the post from the craigslist site.Thank you and God Bless [words omitted]
- Do you still have this item for sale? GOD BLESS AMERICA.....
- [words omitted] God bless as you do ship and i hope to do more business with you. [words omitted]

Figure A.6: **Sample emails invoking God.**

### Capitalized Text

- LET ME KNOW IF THE ITEM STILL AVAILABLE FOR SALE.
- HELLO..IS THE ITEM HELLO, IS THE ITEM POSTED ON CL LISTED ABOVE STILL FOR SALE?? KINDLY GET BACK TO ME WITH THE LAST PRICE AND PRESENT CONDITION.THANKS

Figure A.7: Sample emails with CAPITALIZED text.

### Recurring Themes in Emails

#### **Military member unable to come view product**

- [words omitted] My mode of payment would be in CERTIFIED CHECK and i will arrange for a local pick up as soon as you get the check; because that is the only convenient means for me and due to my work frame i can not be able to get there and i promise everything will go smoothly.I really wish to be there to check out the item but i don't have chance cause am very busy person (US MARINE). And am already back to camp but i will get home very soon [words omitted]
- [words omitted] i have no problem with the amount as am a US marine i work for the United State Marine Corps (USMC) but am currently hospitalized so am on a treatment in New york [words omitted]
- [words omitted] Am willing to buy and am a serious buyer but am not around now so i won't be able to come to have a look because am in camp now I'm a Marine(US MARINE).payment will be done by BANK CERTIFIED CHECK [words omitted]

#### **Present for family member and buyer overseas**

- [words omitted] i want you to know you are also in safe hands and i want you to assure me that i won't be disappointed with it cos am getting it for my cousin the issue is that am not around i would have come and see it [words omitted]
- [words omitted] i wanted to buy this for my Cousin; but the issue is am currently out of state on a Contract Project .The contract is strictly no call due to the lack of reception in the area. [words omitted]
- [words omitted] im arranging it for my cousing birthday who live in OKLAHOMA USA.im off shore and Right now the only way i can make the payment is via paypal as i don't have access to my bank account online and theres no way i can issue out a check or something here [words omitted]

Figure A.8: Sample themes in emails. Usually observed in second or later responses. Conversation leads to fraud attempt through fake PayPal payment or bogus check.

## A.2 Example rental scam ads

\*Come See this stunning 2 bedroom 2.5 bathroom home\*  
Come See this quiet 3 bedroom 2 bathroom rental property\* 2014 Deal  
Come See this lovely 1 bed / 1.5 bath rental\* Discount for 2014  
Come Lay your eyes upon this wonderful 2 bedroom 1.5 bathroom property\*  
Come Lay your eyes upon this gorgeous 1 bed 1 bath place\*  
Come Lay your eyes upon this gorgeous 1 bed / 1.5 bath rental\*

Figure A.9: Example ad titles with sophisticated templates used by CreditReport\_Yahoo campaign.

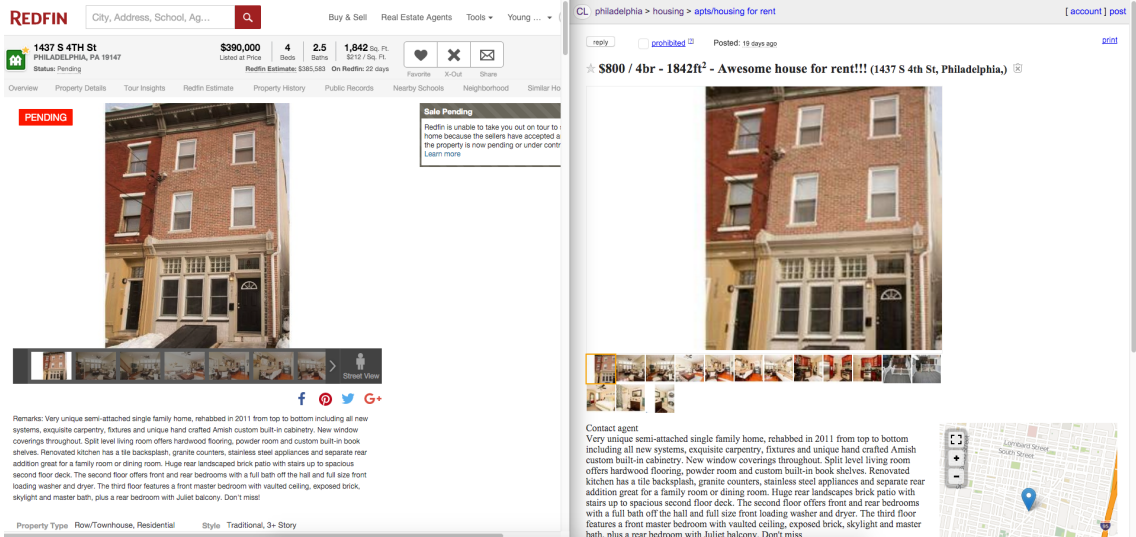


Figure A.10: Example rental clone scam ad. Left figure is an original legitimate ad posted on Redfin.com for sale. Right figure is a clone scam ad with extremely low rent price, \$ 800.



### A.3 Example rental scam emails

```
Good Morning,  
How are you doing? I had to leave Bloomington, MN in a hurry due to my recent  
promotion and it involves me working in different states.  
...  
I am currently working as a counselor for the United States Aid for  
International Development (US-AID) here in Salinas, CA. Please visit  
http://www.usaid.gov/, you can also support our mission.  
...  
The rent is $600 while the security deposit is also $600. The house is ready  
for move in..UTILITIES INCLUDED AND LEASE IS 6 MONTHS TO A YEAR..  
...
```

(a) Example scam email received from KenStaley campaign. (User scam report)

```
Good evening,  
How are you and your family? I really hope i am not getting back to you very  
late about the house you want to rent, I had to leave Houston, TX in a hurry  
due to my recent promotion and it involves me working out of the Houston, TX.  
...  
I am currently working as a counselor for the United States Aid  
for International Development (US-AID) here in AL. Please visit  
http://www.usaid.gov/, you can also support our mission.  
...  
The rent is $600 while the security deposit is also $600. The house is ready  
for move in.. YES UTILITIES INCLUDED AND LEASE IS 6 MONTHS TO A YEAR WITH  
OPTION OF LEASE TO OWN..  
...
```

(b) Example scam email received from a clone scam campaign.

Figure A.11: **Example rental scam emails.**

```

PLEASE TELL US ABOUT YOURSELF
Full Name_____
Home Phone ( )_____
Cell Phone ( ) _____
Date of Birth_____
Current Address_____
City_____State_____ Zip_____
Reasons for Leaving_____Rent $_____
Are you married_____
How many people will be living in the house_____
Do you smoke_____
Do you have a pet_____
Do you have a car_____
Occupation_____
Move In Date_____
How soon can you make the payment_____
How soon do you want to receive the keys and the document_____

```

Figure A.12: **Example rent application template.** Clone scam campaigns usually request a victim to fill out their rent application form.

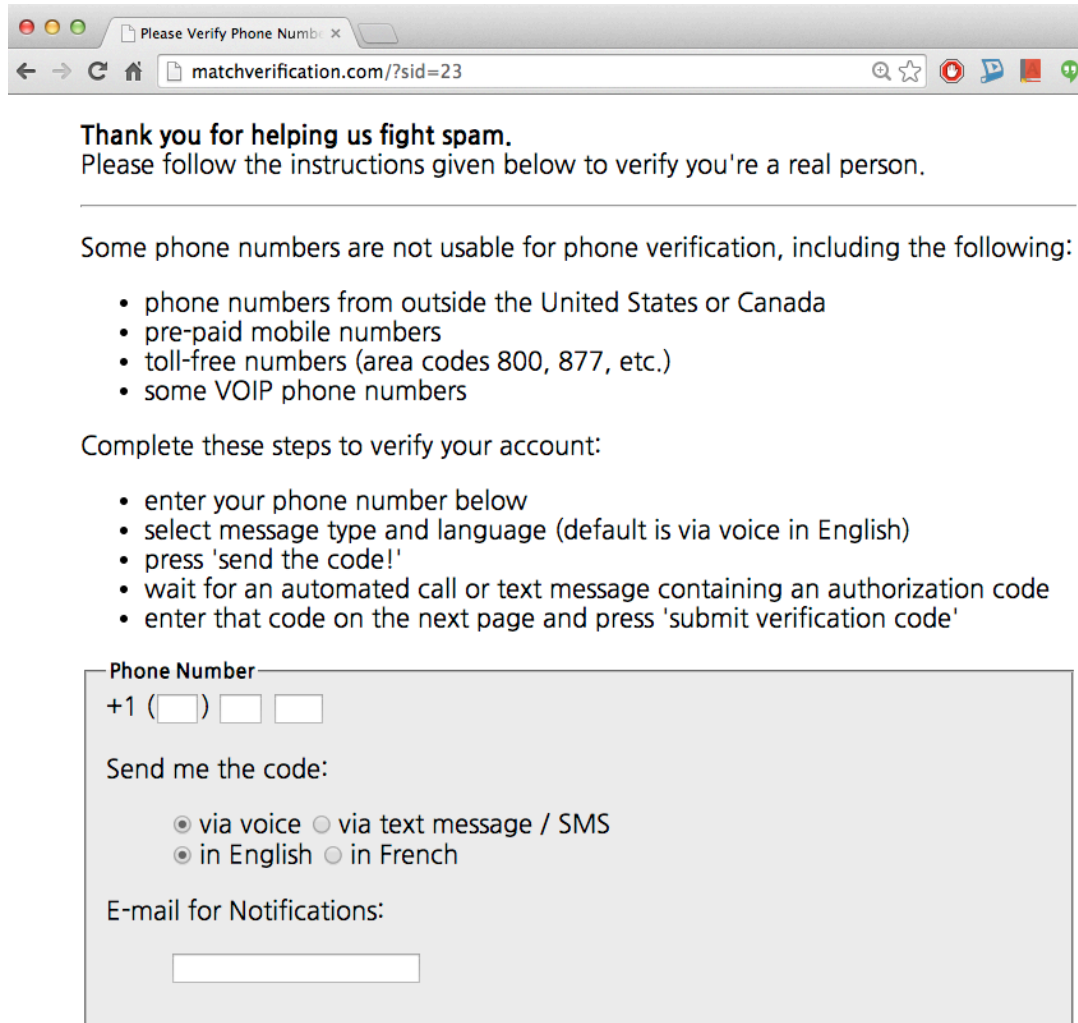
```

Hello,
I hope you are having a wonderful day. Here's some good news: the apartment's
still available!
...
When you're ready for a personal appointment, then please go to the link
below and grab your free credit score. We recommend this site because all
of our tenants used it and never had any problems. Just fill out the form and
indicate that you want the score. What is in the report isn't important to us,
it's more of a formality to have it on file, to make sure there are no previous
property related issues. You can get your free credit score at CLICK HERE
Remember, we only need to see the page about the rental history. That's all
we need to see at the showing. We typically waive the security deposit with a
score of 560+.
...

```

Figure A.13: **Example credit report scam email.**

## A.4 Misc.



Please Verify Phone Number x

matchverification.com/?sid=23

**Thank you for helping us fight spam.**  
Please follow the instructions given below to verify you're a real person.

---

Some phone numbers are not usable for phone verification, including the following:

- phone numbers from outside the United States or Canada
- pre-paid mobile numbers
- toll-free numbers (area codes 800, 877, etc.)
- some VOIP phone numbers

Complete these steps to verify your account:

- enter your phone number below
- select message type and language (default is via voice in English)
- press 'send the code!'
- wait for an automated call or text message containing an authorization code
- enter that code on the next page and press 'submit verification code'

Phone Number

+1 (  )

Send me the code:

via voice  via text message / SMS

in English  in French

E-mail for Notifications:

Figure A.14: **Craigslist phone verification scam.** The web page shows the Craigslist's original phone verification web page to victims. The victims unintentionally help the scammer get phone-verified Craigslist accounts.

## Bibliography

- [1] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge university press Cambridge, 2010.
- [2] Federal Bureau of Investigation, Internet Crime Complaint Center (IC3) annual reports. <https://www.ic3.gov/media/annualreports.aspx>.
- [3] Kathleen Fearn-Banks. *Crisis communications: A casebook approach*. Routledge, 2006.
- [4] Monica T Whitty and Tom Buchanan. The online romance scam: A serious cybercrime. *CyberPsychology, Behavior, and Social Networking*, 15(3):181–183, 2012.
- [5] Andrew Smith. Nigerian scam e-mails and the charms of capital. *Cultural Studies*, 23(1):27–47, 2009.
- [6] Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, Damon McCoy, Nicholas Weaver, Vern Paxson, Geoffrey M. Voelker, and Stefan Savage. Click trajectories: End-to-end analysis of the spam value chain. In *IEEE Symposium on Security and Privacy*, 2011.
- [7] Damon McCoy, Hitesh Dharmdasani, Christian Kreibich, Geoffrey M. Voelker, and Stefan Savage. Priceless: The role of payments in abuse-advertised goods. In *Proceedings of the 2012 ACM Conference on CCS, CCS '12*, 2012.
- [8] Youngsam Park, Jackie Jones, Damon McCoy, Elaine Shi, and Markus Jakobsson. Scambaiter: Understanding Targeted Nigerian Scams on Craigslist. In *NDSS*, 2014.
- [9] Youngsam Park, Damon McCoy, and Elaine Shi. Understanding craigslist rental scams. 2012.

- [10] Creola Johnson. Fakers, breachers, slackers, and deceivers: Opportunistic actors during the foreclosure crisis deserve criminal sanctions. *Capital University Law Review*, 40(4), December 2012.
- [11] Jim Buchanan and Alex J Grant. Investigating and prosecuting Nigerian fraud. *United States Attorneys' Bulletin*, 49(6):39–47, 2001.
- [12] Mohammad Karami, Youngsam Park, and Damon McCoy. Stress testing the booters: Understanding and undermining the business of ddos services. 2016.
- [13] Michaela Beals, Marguerite DeLiema, and Martha Deevy. Framework for a taxonomy of fraud. [http://fraudresearchcenter.org/wp-content/uploads/2015/07/FFRC\\_Taxonomy\\_FullReport\\_7-22-15.pdf](http://fraudresearchcenter.org/wp-content/uploads/2015/07/FFRC_Taxonomy_FullReport_7-22-15.pdf).
- [14] National Fraud Authority. <https://www.gov.uk/government/organisations/national-fraud-authority>.
- [15] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [17] Karen Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21, 1972.
- [18] Project Honey Pot. <http://www.projecthoneypot.org/>.
- [19] GeekWire. <http://www.geekwire.com/2011/stats-hotmail-top-worldwide-gmail-posts-big-gains/>.
- [20] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed 2016-04-14].
- [21] Ralph B d'Agostino. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2):341–348, 1971.
- [22] RB DAugustino and ES Pearson. Testing for departures from normality. *Biometrika*, 60:613–622, 1973.
- [23] Marilyn A. Dyrud. I brought you a good news: An analysis of Nigerian 419 letters. In *Proceedings of the 2005 Association for Business Communication Annual Convention*, 2005.
- [24] Aunshul Rege. What's love got to do with it? Exploring online dating scams and identity fraud. *International Journal of Cyber Criminology*, 3(2):494–512, 2009.

- [25] Frank Stajano and Paul Wilson. Understanding scam victims: seven principles for systems security. *Communications of the ACM*, 54(3):70–75, 2011.
- [26] Vaibhav Garg and Shirin Nilizadeh. Craigslist Scams and Community Composition: Investigating Online Fraud Victimization. In *International Workshop on Cyber Crime*. IEEE, 2013.
- [27] Charles Tive. *419 scam: Exploits of the Nigerian con man*. iUniverse, 2006.
- [28] Cormac Herley. Why do Nigerian Scammers say they are from Nigeria? In *WEIS*, 2012.
- [29] Jelena Isacenkova, Olivier Thonnard, Andrei Costin, Davide Balzarotti, and Aurelien Francillon. Inside the scam jungle: A closer look at 419 scam email operations. In *Security and Privacy Workshops (SPW), 2013 IEEE*, pages 143–150. IEEE, 2013.
- [30] Yanbin Gao and Gang Zhao. Knowledge-based Information Extraction: a case study of recognizing emails of Nigerian frauds. In *Natural Language Processing and Information Systems*. Springer, 2005.
- [31] Chris Kanich, Christian Kreibich, Kirill Levchenko, Brandon Enright, Geoffrey M Voelker, Vern Paxson, and Stefan Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the 15th ACM conference on CCS*. ACM, 2008.
- [32] Brett Stone-Gross, Thorsten Holz, Gianluca Stringhini, and Giovanni Vigna. The underground economy of spam: a botmaster’s perspective of coordinating large-scale spam campaigns. In *Proceedings of the 4th USENIX conference on Large-scale exploits and emergent threats, LEET’11*, pages 4–4, Berkeley, CA, USA, 2011. USENIX Association.
- [33] Damon McCoy, Andreas Pitsillidis, Grant Jordan, Nicholas Weaver, Christian Kreibich, Brian Krebs, Geoffrey M. Voelker, Stefan Savage, and Kirill Levchenko. Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs. In *USENIX Security symposium*, 2012.
- [34] Maria Konte, Nick Feamster, and Jaeyeon Jung. Dynamics of online scam hosting infrastructure. In *Passive and Active Network Measurement*, pages 219–228. Springer, 2009.
- [35] Brett Stone-Gross, Andy Moser, Christopher Kruegel, Engin Kirda, and Kevin Almeroth. FIRE: FInding Rogue nEtworks. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)*, Honolulu, HI, December 2009.
- [36] United States Census Bureau. <http://www.census.gov/>.
- [37] Ripoff Report. <http://www.ripoffreport.com/>.

- [38] 800notes. <http://800notes.com/>.
- [39] Report craigslist Scams. <http://reportcraigslistscams.com/>.
- [40] DB-IP. <http://db-ip.com/>.
- [41] JingMin Huang, Gianluca Stringhini, and Peng Yong. Quit playing games with my heart: Understanding online dating scams. In *Detection of Intrusions and Malware, and Vulnerability Assessment , DIMVA, 2015*.
- [42] David Karig and Ruby Lee. Remote denial of service attacks and countermeasures. *Princeton University Department of Electrical Engineering Technical Report CE-L2001-002*, 17, 2001.
- [43] Character Generator Protocol. <https://tools.ietf.org/html/rfc864>.
- [44] Simple Service Discovery Protocol. <https://tools.ietf.org/html/draft-cai-ssdp-v1-03>.
- [45] Stephen M Specht and Ruby B Lee. Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures. In *Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems*, pages 543–550, 2004.
- [46] Vern Paxson. An Analysis of Using Reflectors for Distributed Denial-of-service Attacks. *SIGCOMM Comput. Commun. Rev.*, 31(3):38–47, July 2001.
- [47] Stefan Savage, David Wetherall, Anna Karlin, and Tom Anderson. Practical Network Support for IP Traceback. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '00, pages 295–306, 2000.
- [48] Yang Xiang, Ke Li, and Wanlei Zhou. Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics. *IEEE Transactions on Information Forensics and Security*, 6(2):426–437, June 2011.
- [49] John Ioannidis and Steven M. Bellovin. Implementing Pushback: Router-Based Defense Against DDoS Attacks. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2002, San Diego, California, USA*, 2002.
- [50] Srikanth Kandula, Dina Katabi, Matthias Jacob, and Arthur Berger. Botz-4-sale: Surviving Organized DDoS Attacks That Mimic Flash Crowds. In *Proceedings of the 2ND Conference on Symposium on Networked Systems Design & Implementation*, NSDI'05, pages 287–300. USENIX Association, 2005.
- [51] Christian Rossow. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. In *Proceedings of the 2014 Network and Distributed System Security (NDSS) Symposium*, February 2014.

- [52] Marc Kührer, Thomas Hupperich, Christian Rossow, and Thorsten Holz. Hell of a Handshake: Abusing TCP for Reflective Amplification DDoS Attacks. In *Proceedings of the 8th USENIX Workshop on Offensive Technologies*, August 2014.
- [53] David Moore, Colleen Shannon, Douglas J. Brown, Geoffrey M. Voelker, and Stefan Savage. Inferring Internet Denial-of-service Activity. *ACM Trans. Comput. Syst.*, 24(2):115–139, May 2006.
- [54] Eric Wustrow, Manish Karir, Michael Bailey, Farnam Jahanian, and Geoff Huston. Internet Background Radiation Revisited. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, pages 62–74. ACM, 2010.
- [55] Jakub Czyz, Michael Kallitsis, Manaf Gharaibeh, Christos Papadopoulos, Michael Bailey, and Manish Karir. Taming the 800 Pound Gorilla: The Rise and Decline of NTP DDoS Attacks. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 435–448. ACM, 2014.
- [56] Marc Kührer, Thomas Hupperich, Christian Rossow, and Thorsten Holz. Exit from Hell? Reducing the Impact of Amplification DDoS Attacks. In *Proceedings of the 23rd USENIX Conference on Security Symposium*, SEC'14, pages 111–125. USENIX Association, 2014.
- [57] Armin Büscher and Thorsten Holz. Tracking DDoS Attacks: Insights into the Business of Disrupting the Web. In *Proceedings of the 5th USENIX Conference on Large-Scale Exploits and Emergent Threats*, LEET'12, pages 8–8, Berkeley, CA, USA, 2012. USENIX Association.
- [58] Vrizlynn L Thing, Morris Sloman, and Naranker Dulay. A Survey of Bots Used for Distributed Denial of Service Attacks. In *New Approaches for Security, Privacy and Trust in Complex Environments*, pages 229–240. Springer, 2007.
- [59] Evan Cooke, Farnam Jahanian, and Danny McPherson. The zombie roundup: Understanding, detecting, and disrupting botnets. In *Proceedings of the Steps to Reducing Unwanted Traffic on the Internet*, SRUTI'05. USENIX Association, 2005.
- [60] Arne Welzel, Christian Rossow, and Herbert Bos. On Measuring the Impact of DDoS Botnets. In *Proceedings of the 7th European Workshop on Systems Security (EuroSec 2014)*, April 2014.
- [61] Mohammad Karami and Damon McCoy. Understanding the Emerging Threat of DDoS-as-a-Service. In *USENIX Workshop on Large-Scale Exploits and Emergent Threats*, Berkeley, CA, 2013. USENIX.
- [62] Tyler Moore, Richard Clayton, and Ross Anderson. The Economics of Online Crime. *Journal of Economic Perspectives*, 23(3):3–20, 2009.



- [63] David Y. Wang, Matthew Der, Mohammad Karami, Lawrence Saul, Damon McCoy, Stefan Savage, and Geoffrey M. Voelker. Search + Seizure: The Effectiveness of Interventions on SEO Campaigns. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 359–372, 2014.
- [64] Brett Stone-Gross, Thorsten Holz, Gianluca Stringhini, and Giovanni Vigna. The Underground Economy of Spam: A Botmaster’s Perspective of Coordinating Large-scale Spam Campaigns. In *Proceedings of the 4th USENIX Conference on Large-scale Exploits and Emergent Threats*, LEET’11, 2011.
- [65] Gianluca Stringhini, Gang Wang, Manuel Egele, Christopher Kruegel, Giovanni Vigna, Haitao Zheng, and Ben Y. Zhao. Follow the Green: Growth and Dynamics in Twitter Follower Markets. In *Proceedings of the 2013 Conference on Internet Measurement Conference*, IMC '13, pages 163–176, 2013.
- [66] Kurt Thomas, Damon McCoy, Chris Grier, Alek Kolcz, and Vern Paxson. Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse. In *Proceedings of the 22nd Usenix Security Symposium*, 2013.
- [67] R. Clayton, T. Moore, and N. Christin. Concentrating correctly on cybercrime concentration. In *Proceedings (online) of the Fourteenth Workshop on the Economics of Information Security (WEIS)*, Delft, Netherlands, June 2015.
- [68] Brett Stone-Gross, Ryan Abman, Richard Kemmerer, Christopher Kruegel, Douglas Steigerwald, and Giovanni Vigna. The Underground Economy of Fake Antivirus Software. In *Economics of Information Security and Privacy III*, pages 55–78. Springer, 2013.
- [69] Corey Satten. Lossless Gigabit Remote Packet Capture With Linux. <http://staff.washington.edu/corey/gulp/>, 2008.
- [70] Open Resolver Project. <http://OpenResolverProject.org/>.
- [71] Open NTP Scanning Project. <http://OpenNTPProject.org/>.
- [72] Orcun Cetin, Mohammad Hanif Jhaveri, Carlos Ganan, Michel van Eeten, and Tyler Moore. Understanding the role of sender reputation in abuse reporting and cleanup. In *Workshop on the Economics of Information Security*, 2015.
- [73] Zakir Durumeric, James Kasten, David Adrian, J. Alex Halderman, Michael Bailey, Frank Li, Nicolas Weaver, Johanna Amann, Jethro Beekman, Mathias Payer, and Vern Paxson. The Matter of Heartbleed. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 475–488, 2014.