
Dissertations, Theses, and Masters Projects

Theses, Dissertations, & Master Projects

2016

Investigating Fraudulent and Privacy Activities in Online Business.

Haitao Xu
College of William and Mary

Follow this and additional works at: <https://scholarworks.wm.edu/etd>



Part of the [Communication Technology and New Media Commons](#)

Recommended Citation

Xu, Haitao, "Investigating Fraudulent and Privacy Activities in Online Business." (2016). *Dissertations, Theses, and Masters Projects*. Paper 1593092112.
<https://dx.doi.org/doi:10.21220/m2-p516-h015>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Investigating Fraudulent and Privacy Activities in Online Business

Haitao Xu

Zhoukou, Henan, China

Master of Science, University of Chinese Academy of Sciences, 2010
Bachelor of Science, Zhengzhou University, 2007

A Dissertation presented to the Graduate Faculty
of the College of William and Mary in Candidacy for the Degree of
Doctor of Philosophy

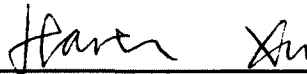
Department of Computer Science

The College of William and Mary
January, 2016

APPROVAL PAGE

This Dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy



Haitao Xu

Approved by the Committee, December, 2015

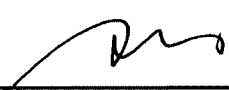


Committee Chair

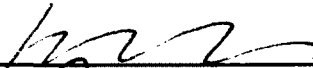
Professor Haining Wang, Electrical and Computer Engineering
University of Delaware



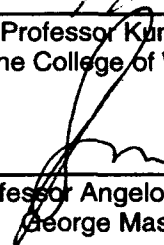
Professor Weizhen Mao, Computer Science
The College of William and Mary



Professor Qun Li, Computer Science
The College of William and Mary



Assistant Professor Kun Sun, Computer Science
The College of William and Mary



Associate Professor Angelos Stavrou, Computer Science
George Mason University

ABSTRACT

The continuous expansion of the Internet in the past 20 years has greatly facilitated the booming development of Internet business. Unfortunately, some unscrupulous participants in Internet business conduct fraudulent activities for their own profits at the expense of other parties. In this proposal, we present our study on the fraudulent activities in two kinds of major Internet businesses—online advertising and E-commerce.

Online advertising is leveraged by online advertisers to deliver marketing messages to potential customers. It serves as a significant source of revenue for web-based businesses and is crucial to a thriving Internet ecosystem. However, click fraud is posing a serious threat to online advertising systems. As the direct victims, advertisers still lack effective defense against click fraud. In this proposal, we present a novel approach for advertisers to detect click fraud without the helps from ad networks or publishers. Our proposed defense is effective in identifying both clickbots and human clickers, while incurring negligible overhead at both the server and client sides.

In an E-commerce market, a store's reputation is closely tied to its profitability. Sellers' desire to quickly achieve high reputation has fueled a profitable underground business, termed by us as a *seller-reputation-escalation (SRE)* market. An SRE market operates as a specialized crowdsourcing marketplace and facilitates online sellers to harness human laborers to conduct fake transactions for improving their stores' reputations. In this proposal, we characterize the SRE markets in terms of its prevalence, business model, market size, and the sellers and laborers involved. We also evaluate the effectiveness of the SRE services on reputation escalation.

An online photo could disclose much more information beyond what is visually depicted in the photo and what its owner expects to share. In this dissertation, we aim to raise public awareness of privacy risks resulting from sharing photos online. We first investigate the prevalence of privacy information among digital photos. Then we study the policies adopted by online media sites on handling the metadata information embedded in the photos they host. Finally, we introduce an attack vector not yet exploited before and demonstrate its surprising power in identifying a photographer with just one photo she ever took.

TABLE OF CONTENTS

Acknowledgments	vi
Dedication	vii
List of Tables	viii
List of Figures	xi
1 Introduction	2
1.1 Detecting Click Fraud in Online Advertising System	3
1.2 Understanding the Emerging Reputation-Escalation-as- a-Service in E-commerce System	5
1.3 Assessing Privacy Risks on Online Photos	8
1.4 Future Works	10
1.5 Organization	10
2 Click Fraud Detection on the Advertiser Side	11
2.1 Background	13
2.1.1 Clickbots	13

2.1.2	Human clickers	15
2.1.3	Advertisers	16
2.1.4	Web standards and feature detection techniques	16
2.2	Methodology	17
2.2.1	JavaScript support and mouse event test	18
2.2.2	Functionality test	19
2.2.3	Browsing behavior examination	22
2.3	Experimental Results	22
2.3.1	Running ad campaigns	23
2.3.2	Characterizing the click traffic	24
2.3.3	Validating detection approach	28
2.4	Discussion and Limitations	34
2.5	Related Work	35
2.6	Conclusion	38
3	E-commerce Reputation Manipulation: The Emergence of Reputation-Escalation-as-a-Service	39
3.1	Background	40
3.1.1	Taobao Overview	41
3.1.2	How a Typical SRE Market Works	41
3.2	Data Collection Methodology	46

3.2.1	Crawling Mechanism	46
3.2.2	Data Summary	49
3.2.3	Ethical Considerations	50
3.3	SRE Market Characteristics	51
3.3.1	SRE Market Popularity	51
3.3.2	Strategies to Evade Taobao Detection	52
3.3.3	Effectiveness of SRE Markets' Evasion Strategies	57
3.3.3.1	Taobao's Punitive Measures	58
3.3.3.2	Penalties Imposed on SRE Sellers	58
3.3.3.3	Summary	62
3.3.4	Seller Characteristics	62
3.3.5	Worker Characteristics	66
3.3.6	Estimating Revenue and Fake-Transaction Volume	68
3.4	Effectiveness of SRE services	69
3.4.1	Effect of Posting Tasks on SRE Market	70
3.4.2	An Emerging Service and Its Effectiveness	73
3.5	Potential Mitigation Strategies and Limitation	75
3.6	Revisit the SRE Ecosystem One Year Later	78
3.6.1	Current Statuses of SRE Markets	79
3.6.2	Current Activities of SRE Sellers on SRE Markets	80

3.6.3	Current Statuses of the EVIL Taobao Stores on the Taobao Marketplace	81
3.6.3.1	Reputation Growth of the Active EVIL Stores Across One Year	83
3.6.3.2	Reasoning Why EVIL Stores Become In- accessible Now	84
3.6.3.3	Summary	88
3.6.4	Summary	88
3.7	Related Work	88
3.8	Conclusion	90
4	Assessing Privacy Risks on Online Photos	92
4.1	Background	94
4.1.1	Metadata Information in a Photo	94
4.1.2	Potential Privacy Concerns Arising from Photo Meta- data	95
4.1.3	Three Stages of Digital Photos	97
4.2	Fresh Photos	97
4.2.1	Data Collection	97
4.2.2	Characterizing “Fresh” Photos	98
4.3	Intact Photos	100

4.3.1 Data Collection	101
4.3.2 Metadata Information Embedded	101
4.4 Wild Photos	102
4.4.1 Data Collection	102
4.4.2 Ethical Consideration	104
4.4.3 Metadata Information Embedded	104
4.4.4 Inferring Online Sites' Photo Handling Policies .	106
4.5 Re-Identification Attack	110
4.6 Discussion	115
4.7 Related Work	117
4.8 Conclusion	119
5 Conclusion and Future Work	121

ACKNOWLEDGEMENTS

My thanks first go to my advisor Dr. Haining Wang, without whom this dissertation would not have been accomplished.

I would also like to thank my committee members, Dr. Weizhen Mao, Dr. Qun Li, Dr. Kun Sun, and Dr. Angelos Stavrou for their invaluable feedback which truly helped me improve the dissertation.

A special thanks goes to the staff members in our department, including Vanessa Godwin, Jacquelyn Johnson, and Dale Hayes, for their help and support over the last five years.

This Ph.D. is dedicated to my parents and my wife for their endless love and support

LIST OF TABLES

2.1	Tested browsers, versions and release dates	19
2.2	Authentic feature set widely supported by modern browsers	20
2.3	Summary of our ad campaigns	24
2.4	Features extracted for each ad click	31
3.1	Qualification types	43
3.2	A typical task description	44
3.3	List of the SRE markets we infiltrated, the months monitored, total task postings, active sellers during the time frame, and Taobao ID identified sellers. *This market went down between 02/21 and 03/06.	50
3.4	Statistics of daily active sellers, daily new tasks, and the time to undertake a task on the five SRE markets. . .	51
3.5	Fraction (%) of tasks with restrictions on workers' Taobao accounts.	52
3.6	Fraction (%) of tasks with geographic preference and shipping address (SA) designated.	53

3.7 Fraction (%) of tasks with requirements for each kind of browsing behavior before checking out. “One or more” denotes the tasks with at least one required action.	54
3.8 Fraction (%) of tasks with requirements for e-Gift card payment or no credit card payment.	55
3.9 Fraction (%) of tasks declaring shipment of empty package.	56
3.10 Fraction (%) of tasks requiring leaving email address and phone number, confirmation on AliWangWang, and guarantee money.	57
3.11 Distribution of the reputation scores of the 4,109 identified Taobao sellers at the inception of using SRE services.	64
3.12 Estimated revenue and transaction volume	68
3.13 The store ages based on which we partition EVIL and BENI sellers. “< 1m” denotes a store age of less than 1 month while “< 1y” denotes less than 1 year.	70
3.14 TRUSTEE service expense standard: list of desired Taobao grade, corresponding reputation score increase, charged fees, and days needed to complete.	73

3.15 Statistics of total task postings, total active sellers, maximum daily new tasks, and maximum daily active sellers on the three SRE markets during 40 days.	80
3.16 Variation of the dynamism of the three SRE markets in the past one year.	80
3.17 Current status of the 8,473 sellers on the SRE markets.	81
3.18 Current statuses of <i>EVIL</i> and <i>BENI</i> stores.	83
4.1 List of metadata information typically included in a digital photo.	94
4.2 An example of modification information contained in a photo's metadata.	96
4.3 Demographic statistics of worker participants	99
4.4 List of the information typically contained in an account profile in each of the five OSNs. Note that the listed information represents the maximum amount of information available with public permissions of an OSN account.	113
4.5 Main functions of the browser extension prototype . . .	116

LIST OF FIGURES

2.1	How a clickbot works	14
2.2	Outline of click fraud detection mechanism	18
2.3	How the functionality test is performed by advertiser's web server.	21
2.4	A bait ad with the ad text of randomly selected English words	23
2.5	Distribution of click traffic vs. that of normal traffic by country	25
2.6	Distribution of click traffic by browser	26
2.7	Distribution of click traffic by publisher	27
2.8	Percentage of clicks without JavaScript support for the top 10 publisher websites contributing the most clicks .	28
2.9	Clients' execution time of JavaScript challenge code in milliseconds	33
3.1	Lifecycle of a fake-purchase task on the SRE market. .	42
3.2	Procedure of data collection.	46

3.3	A breakdown of when to confirm the receipt. Numbers in parentheses in the legend denote the fraction of the total 219,165 tasks crawled on the five SRE markets. .	56
3.4	CDF of the deducted reputation points of the 932 Taobao sellers suffering from the penalty V.	59
3.5	CDF of the reputation scores of the 89 heavily penalized sellers when beginning posting fake-transaction tasks on SREs.	59
3.6	CDF of the shop ages of the 89 sellers while being heavily penalized.	61
3.7	CDF of days taken from unusual reputation growth to heavy penalty.	61
3.8	CDF of shop start date of the identified Taobao sellers.	63
3.9	Top 5 main businesses run by the 4,109 Taobao sellers.	64
3.10	CDF of active duration of sellers on the COOL market.	65
3.11	Tasks posted daily per active seller on SRE markets over time. Numbers in parentheses in the legend denote the mean values of the number of tasks posted daily per active seller on each SRE market during our crawl interval.	66

3.12(a) CDF of active duration of the top workers on the COOL market. (b) CDF of average tasks undertaken daily by the top workers on the COOL market. (c) CDF of average daily earnings of the top workers on the COOL market.	67
3.13 Comparison of the reputation growth distribution between BENI and EVIL stores with varying store ages over the course of one month.	71
3.14 Reputation changes over time for the 12 Taobao sellers identified to use TRUSTEE service.	74
3.15 CDF of task postings per active seller per day in 2015 v.s. that in 2014	82
3.16 Comparison between currently active <i>EVIL</i> and <i>BENI</i> sellers in terms of reputation increase across one year. .	84
3.17 Comparison between currently active <i>EVIL</i> and <i>BENI</i> sellers in terms of reputation growth rate across one year. .	84
3.18 CDF of monthly reputation increase of the currently inaccessible <i>BENI</i> stores between 3/21/2014 and 4/21/2014. .	85
3.19 CDF of monthly reputation increase of the currently inaccessible <i>EVIL</i> stores between 2/21/2014 and 4/21/2014. .	85

3.20 CDF of monthly reputation growth rate of the currently inaccessible <i>BENI</i> stores between 3/21/2014 and 4/21/2014.	86
3.21 CDF of monthly reputation growth rate of the currently inaccessible <i>EVIL</i> stores between 2/21/2014 and 4/21/2014.	86
3.22 CDF of the monthly transaction volumes completed by the currently inaccessible <i>BENI</i> sellers in 2014.	87
3.23 CDF of the monthly transaction volumes completed by the currently inaccessible <i>EVIL</i> sellers in 2014.	87
4.1 Percentage of “fresh” photos containing metadata infor- mation.	99
4.2 Percentage of “fresh” photos tagged with GPS for smart- phone OS.	99
4.3 Percentage of “intact” photos containing metadata infor- mation. In each of four pairs of columns, the left black column represents <i>Flickr_p</i> while the right gray <i>Flickr_6</i> .	101
4.4 CDF of number of photos crawled from each site. . . .	105
4.5 Percentage of “wild” photos containing metadata infor- mation. In each of four pairs of columns, the left black column represents <i>GoogleImage</i> while the right gray <i>Top- SitesPhoto</i>	105

4.6 Percentage of sites estimated to resize their photos across the seven categories. 107

4.7 CDF of the percentage of photos containing metadata information on each site. 109

4.8 Percentage of sites estimated to preserve the photo metadata information across the seven categories. 110

4.9 CDF of the number of photos returned by *stolencamerafinder* for a given serial number. 112

4.10 Percentage of camera serial numbers (SNs) with camera owners' OSN accounts identified. 114

Investigating Fraudulent and Privacy Activities in Online Business

Chapter 1

Introduction

With the popularity of the Web, the Internet has been widely used for conducting business communications, collaboration, and transactions for almost two decades. A number of web-based online marketplaces have sprung up such as Amazon, eBay, and Taobao. In the recent years, the global Internet coverage has been further increased thanks to the explosion of mobile devices. The Internet-related economy today accounts for an increasingly significant portion of global Gross Domestic Product (GDP).

Accompanied by the booming development of Internet business, Internet fraud has become a serious and pervasive security problem. With a monetary motive, miscreants defraud victims to disclose their personal information or conduct fraudulent transactions by various means such as distributing malware, presenting fraudulent solicitations, and hacking websites. Internet fraud has resulted in annual global monetary loss of hundreds of billions of US dollars [1].

In this dissertation, we characterize and detect the fraudulent activities in online advertising systems and e-commerce marketplaces, and also assess privacy risks arising from online photo sharing activities. Specifically, we propose a novel de-

tection method to identify the fraudulent clicks on the ads from the perspective of online advertisers; we conduct an in-depth measurement study on newly discovered underground markets, which facilitate e-commerce sellers to harness human laborers to perform fake transactions for rapidly improving their stores' reputations; finally we investigate the prevalence of private sensitive information among digital photos and explore the potential privacy threats resulting from online photo sharing. For our future work, we plan to develop a novel detection system to capture those fake transactions conducted on e-commerce marketplaces.

1.1 Detecting Click Fraud in Online Advertising System

In online advertising systems, advertisers pay ad networks for each click on the former's ads, and ad networks in turn pay publishers a share of the revenue. As online advertising has evolved into a multi-billion dollar business [2], click fraud has become a serious and pervasive problem. For example, the botnet "Chameleon" infected over 120,000 host machines in the U.S. and siphoned \$6 million per month from advertisers [3].

Click fraud occurs when miscreants make HTTP requests for destination URLs found in deployed ads [14]. Such HTTP requests issued with malicious intent are called fraudulent clicks. The incentive for fraudsters is to increase their own profits at the expense of other parties. Typically a fraudster is a publisher or an advertiser. Publishers may put excessive ad banners on their pages and then forge clicks on ads to receive more revenue. Unscrupulous advertisers make extensive clicks on a competitor's ads with the intention of depleting the victim's advertising budget.

Click fraud is mainly conducted by leveraging clickbots, hiring human clickers, or tricking users into clicking ads [15]. In an act of click fraud, both an ad network and a publisher are beneficiaries while an advertiser is the only victim, under the pay-per-click model. Although the ad network pays out to the publisher for those undetected click fraud activities, it charges the advertiser more fees. Thus, the ad network still benefits from click fraud. Only the advertiser is victimized by paying for those fraudulent clicks. Therefore, advertisers have the strongest incentive to counteract click fraud. In this dissertation, we focus on click fraud detection from the perspective of advertisers.

We propose a novel approach for an advertiser to independently detect click fraud attacks conducted by clickbots and human clickers. Our approach enables advertisers to evaluate the return on investment (ROI) of their ad campaigns by classifying each incoming click traffic as fraudulent, casual, or valid. The rationale behind our design lies in two observed invariants of legitimate clicks. The first invariant is that a legitimate click should be initiated by a real human user on a real browser. That is, a client should be a real full-fledged browser, and hence it should support JavaScript, DOM, CSS, and other web standards widely followed by modern browsers. The second invariant is that a legitimate ad clicker interested in advertised products must have some level of user engagement on the advertised website. Based on the design principles, we develop a click fraud detection system mainly composed of two components: (1) a proactive functionality test and (2) a passive examination of browsing behavior. The functionality test challenges a client for its authenticity (a browser or a bot) with the assumption that most clickbots have limited functionality compared to modern browsers and thus would fail this test. The second component passively examines each user's browsing behaviors

on the advertised website. Its objective is to identify human clickers and those much advanced clickbots that may pass the functionality test. To evaluate the effectiveness of the proposed detection system, we build a prototype and deploy it on a large production web server. Then we run ad campaigns at one major ad network for 10 days. The experimental results show that our approach can detect much more fraudulent clicks than the ad network's in-house detection system and achieve low false positive and negative rates.

1.2 Understanding the Emerging Reputation-Escalation-as-a-Service in E-commerce System

Due to its convenience and ubiquitous nature, online shopping has become one of the primary means for purchasing goods. In many cases, due to its global nature, lower cost, and fast delivery, online shopping is the preferred or even the exclusive means of acquiring a product. To offer a means for buyers to give feedbacks on the products and the sellers, a large number of online shopping markets including Amazon and eBay have incorporated reputation systems. The reputation systems could encourage sellers to provide better products because through the scoring process they are rewarded with higher reputations, which in turn can attract more business. For instance, sellers with higher reputations are usually listed at the front by online market search engines, and shoppers are biased towards sellers with higher reputations [66]. Thus, online sellers have strong motivation to improve reputations as quickly as possible. In the majority of reputation systems, sellers' reputations are dominated by the number of transactions they complete and the number of positive customer ratings (or reviews) they receive.

However, depending on the popularity of a product, it usually takes a long time for a seller to accumulate high reputation. As a result, a non-negligible number of insincere sellers have attempted to subvert the reputation systems using opinion spams and artificial ratings, among others. Several recent works [62, 53, 63] have aimed to tackle these problems. However, we show that the known reputation manipulation techniques are only the tip of the iceberg of an emerging underground industry that employs sophisticated methods to cater to online sellers who want to quickly boost their store reputations. We refer to such enterprises as *seller-reputation-escalation (SRE) markets*.

SRE markets operate in the crowdsourcing model, where online sellers hire inexpensive human laborers to carry out fake purchase campaigns to accelerate reputation accumulation. By fake purchases, we mean purchases that although they appear legitimate and complete as far as the online system is concerned, no real product or at most an empty package is delivered by a seller. This approach is far more elaborate and much more difficult, if not infeasible, to detect because a buyer appears to have genuinely purchased a product as opposed to just leaving a review or score for the product and its seller. Moreover, multiple individuals that do not know each other are involved in the process.

For insincere sellers, fake purchases can significantly increase their transaction volumes, product ratings, and positive reviews. The boost in overall reputation attracts legitimate customers and at the same time cements the seller's ability to deal with negative reviews. Furthermore, the process is fairly scalable, and the seller may post up to hundreds of such tasks each day for quickly improving the store reputation. Therefore, SRE markets are seriously endangering existing reputation systems widely deployed in current e-commerce platforms. Although SRE markets

have appeared for several years, we still lack insights into the basic characteristics of this underground enterprise.

In this dissertation, we perform a detailed analysis of SRE markets by infiltrating five SRE markets specializing in providing reputation-escalation services to sellers on Taobao marketplace [27], the largest consumer-to-consumer (C2C) online marketplace in China. We conducted daily crawls of these five markets for two months, collected 219,165 tasks posted by more than 11,000 online sellers, and continuously monitored these sellers' activities on these markets. We characterized the sellers and workers involved, and we estimated the revenue generated and fake-transaction volume handled by these markets. Furthermore, we report a more threatening service recently launched by one SRE market where Taobao sellers can increase their reputation scores significantly during a single day¹ by paying less than \$100 to the SRE market operators. Given our insights into the SRE operations, we propose possible intervention approaches from the perspective of defenders. Finally, we revisit the SRE ecosystem one year later to reveal how the dynamism of the SRE markets changes over time and track the statuses of these Taobao sellers we identified to be using SRE services before.

In summary, our main contributions are threefold:

- To our knowledge, we are the first to analyze in depth the operations of SRE markets and perform the empirical study of them. We estimate that the five SRE markets we infiltrated generated at least \$46,438 in revenue and handled at least \$3,452,530 in fake-transaction volume during the two months we monitored.
- We investigate the effect of SRE markets on online store reputation esca-

¹A Taobao seller earns one reputation score for each completed transaction with good ratings.

tion. We find that online sellers using SRE services can increase their store reputations 10 times faster than legitimate ones, and about 25% of them are visibly penalized in the form of either having reputation scores deducted or zeroed, or being forcibly shut down.

- We revisit the SRE ecosystem one year later and observe that the SRE markets are not as active as before. There are evident declines in daily new task postings and daily active sellers on the SRE markets. Moreover, about 17% of these involved Taobao stores do not exist any more, probably due to heavy penalties imposed by Taobao for fake transactions.

1.3 Assessing Privacy Risks on Online Photos

With the proliferation of cameras, especially smartphone cameras, it is now very convenient for people to take photos whenever and wherever possible. Furthermore, the prevalence of online social networks and photo-sharing sites greatly facilitates people to share their digital photos with friends online. Every day around 1.6 million photos are shared on Flickr [72], one of the largest online photo sharing sites.

In their rush to share digital photos online, well-intentioned people unwittingly expose much hidden metadata information contained in the digital photos. The metadata information such as camera serial number may seem relatively innocent and trivial but could create significant privacy threats to photographers² and the people depicted in the photo. Unfortunately, one study [85] shows that up to 40% of high-degree participants do not even know the term metadata. The situation

²By photographer we mean the person who took the photo rather than a person who works as a photographer.

becomes worse concerning the fact that a photo could linger on the Internet for years.

During the spread of a digital photo, online social network (OSN) services and other media sites usually serve as the sink. Online media sites often compress and resize the photos they host for space saving. Media sites may even remove the metadata information in the photos they host. However, users usually do not know what the online services will do with their uploaded photos [85]. Thus, it is important to raise the public awareness of the potential privacy risks posed by metadata leakage and increase their knowledge of how online media sites handle the photos they uploaded.

To better describe contemporary digital photos, we create a taxonomy and classify digital photos into different stages based on the life cycle and the propagation process, which are: “fresh,” “intact,” and “wild.” “Fresh” photos are just freshly taken with a camera. “Intact” photos have been uploaded online but remain intact from the hosting sites. “Wild” photos may have been post-processed multiple times by the hosting sites. In this dissertation, we perform a data-driven assessment of privacy risks on contemporary digital photos. Specifically, we examine digital photos at the three stages in terms of metadata information contained and potential privacy risks, and we further explore the photo handling policies adopted by online media sites. Finally, we introduce a new attack vector and demonstrate that some other trivial looking metadata information could be exploited to launch re-identification attacks against photo owners.

1.4 Future Works

In our previous work, we investigated a new advanced technique of manipulating seller reputations, i.e., the seller-reputation-escalation (SRE) service, which enables sellers to hire thousands of cheap human laborers to conduct fake transactions. Based on our deep understanding of this advanced reputation manipulation technique, for our future work, we are going to develop a practical fake transaction detection system in cooperation with one large e-commerce marketplace. The proposed system will allow e-commerce marketplaces to detect all existing forms of fake transactions in real time.

1.5 Organization

The remainder of this dissertation is structured as follows. Chapter 2 presents our research efforts on detecting click fraud in an online advertising system from the perspective of advertisers. Chapter 3 presents our investigation of an emerging underground industry, termed as SRE markets, where a potentially unbounded number of inexpensive human laborers are hired by e-commerce sellers to conduct fake purchases for reputation inflation. Chapter 4 describes our investigation of online photo privacy issues. In Chapter 5, we conclude the dissertation and propose our future research work on developing a practical and real-time fake transaction detection system from the perspective of e-commerce marketplaces.

Chapter 2

Click Fraud Detection on the Advertiser Side

Click fraud detection is not trivial. Click fraud schemes have been continuously evolving in recent years [14, 23, 12, 15, 19]. Existing detection solutions attempt to identify click fraud activities from different perspectives, but each has its own limitations. The solutions proposed in [21, 20, 22] perform traffic analysis on an ad network's traffic logs to detect publisher inflation fraud. However, an advanced clickbot can conduct a low-noise attack, which makes those abnormal-behavior-based detection mechanisms less effective. Haddadi [18] proposed to leverage bait ads for blacklisting malicious publishers based on a predefined threshold. Motivated by [18], Dave et al. [15] proposed an approach for advertisers to measure click-spam ratios on their ads by creating bait ads. However, running bait ads increases advertisers' budget on advertisements.

In this chapter, we present our work on detecting click fraud in the online advertising system from the perspective of advertisers. Our proposed detection mechanism is composed of two components. The first component is proactive functionality

test. We test a client's functionality against web standards widely supported by modern browsers. Failing the test would induce all clicks generated by the client to be labelled as fraudulent. The second component is passive browsing behavior examination. Only if a client passes the functionality test and also shows enough browsing engagement on the advertised website, the corresponding click is labelled as valid. Otherwise, a click is labelled as casual if the corresponding client passes the functionality test but shows insufficient browsing behaviors. A casual click could be generated by a human clicker or by an unintentional user. We have no attempt to distinguish these two since neither of them is a potential customer from the standpoint of advertisers.

Our detection mechanism can significantly raise the bar for committing click fraud and is potentially effective in the long run after public disclosure. To evade our detection mechanism, clickbots must implement all the main web standards widely supported by modern browsers. And a heavy-weight clickbot will risk itself of being readily noticeable by its host. Likewise, human clickers must behave like real interested users by spending more time, browsing more pages, and clicking more links on the advertised sites, which contradicts their original intentions of earning more money by clicking on ads as quickly as possible. At each point, the net effect is a disincentive to commit click fraud.

The chapter is organized as follows. We provide background knowledge in Section 2.1. Then, we detail our approach in Section 2.2 and validate its efficacy using real-world data in Section 2.3. We discuss the limitations of our work in Section 2.4 and survey related work in Section 2.5. Finally, we conclude the chapter in Section 2.6.

2.1 Background

Based on our understanding of the current state of the art in click fraud, we first characterize clickbots and human clickers, the two main actors leveraged to commit click fraud. We then discuss the advertiser's role in inhibiting click fraud. Finally, we describe the web standards widely supported by modern browsers, as well as feature detection techniques.

2.1.1 Clickbots

A clickbot behaves like a browser but usually has relatively limited functionality compared to the latter. For instance, a clickbot may not be able to parse all elements of HTML web pages or execute JavaScript and CSS scripts. Thus, at the present time, a clickbot is instantiated as malware implanted in a victim's computer. Even assuming a sophisticated clickbot equipped with capabilities close to a real browser, its actual browsing behavior when connected to the advertised website would still be different from that of a real user. This is because clickbots are automated programs and are not sophisticated enough to see and think as human users, and as of yet, do not behave as human users.

A typical clickbot performs some common functions including initiating HTTP requests to a web server, following redirections, and retrieving contents from a web server. However, it does not have the ability to commit click fraud itself but instead acts as a relay based on instructions from a remote bot master to complete click fraud. A bot master can orchestrate millions of clickbots to perform automatic and large-scale click fraud attacks.

Figure 2.1 illustrates how a victim host conducts click fraud under the command of a botmaster. First, the botmaster distributes malware to the victim host by

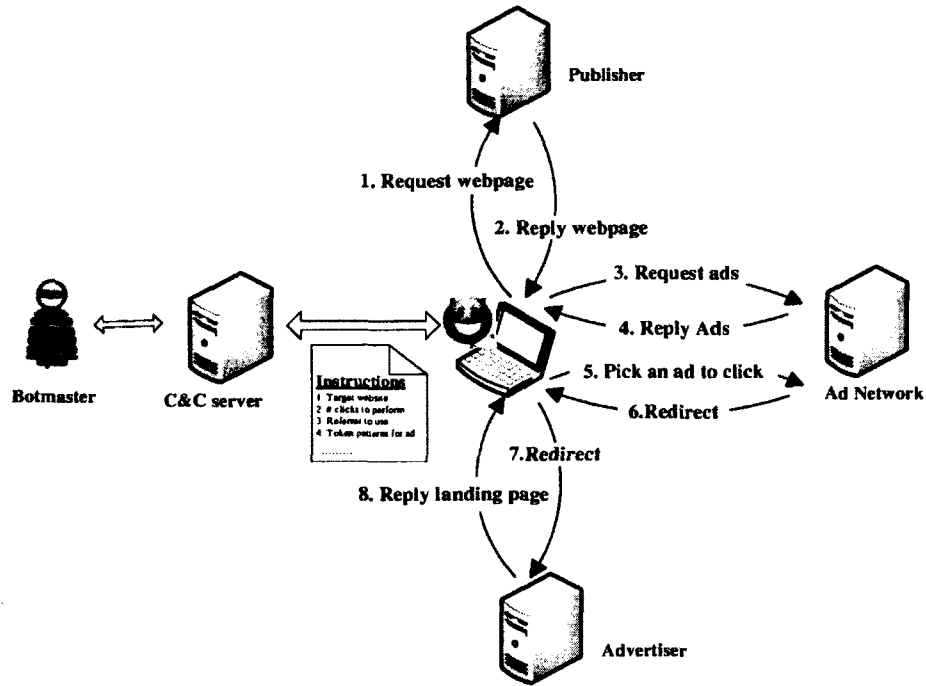


Figure 2.1: How a clickbot works

exploiting the host's security vulnerabilities, by luring the victim into a drive-by download or running a Trojan horse program. Once compromised, the victim host becomes a bot and receives instructions from a command-and-control (C&C) server controlled by the botmaster. Such instructions may specify the target website, the number of clicks to perform on the website, the referrer to be used in the fabricated HTTP requests, what kind of ads to click on, and when or how often to click [14].

After receiving instructions, the clickbot begins traversing the designated publisher website. It issues an HTTP request to the website (**step 1**). The website returns the requested page as well as all embedded ad tags on the page (**step 2**). An ad tag is a snippet of HTML or JavaScript code representing an ad, usually in an iframe. For each ad tag, the clickbot generates an HTTP request to the ad network to retrieve ad contents just like a real browser (**step 3**). The ad network returns

ads to the clickbot (**step 4**). From all of the returned ads, the clickbot selects an ad matching the specified search pattern and simulates a click on the ad, which triggers another HTTP request to the ad network (**step 5**). The ad network logs the click traffic for the purpose of billing the advertiser and paying the publisher a share, and then returns an HTTP 302 redirect response (**step 6**). The clickbot follows the redirection path (possibly involving multiple parties) and finally loads the advertised website (**step 7**). The advertiser returns back the landing page ¹ to the clickbot (**step 8**). At this point, the clickbot completes a single act of click fraud. Every time an ad is “clicked” by a clickbot, the advertiser pays the ad network and the involved publisher receives remuneration from the ad network. Note that a clickbot often works in the background to avoid raising suspicion, thus all HTTP requests in Figure 2.1 are generated without the victim’s awareness.

2.1.2 Human clickers

Human clickers are the people who are hired to click on the designated ads and get paid in return. Human clickers have financial incentives to click on ads as quickly as possible, which distinguishes them from real users who are truly interested in the advertised products. For instance, a real user tends to read, consider, think, and surf the website in order to learn more about a product before purchase. A paid clicker has few such interests, and hence tends to get bored quickly and spends little time on the site [13].

¹Landing page is a single web page that appears in response to clicking on an ad.

2.1.3 Advertisers

Advertisers are in a vantage point to observe and further detect all fraudulent activities committed by clickbots and human clickers. To complete click fraud, all fraudulent HTTP requests must be finally redirected to the advertised website, no matter how many intermediate redirections and parties are involved along the way. This fact indicates that both clickbots and human clickers must finally communicate with the victim advertiser. Thus, advertisers have the advantage of detecting clickbots and human clickers in the course of communication. In addition, as the revenue source of online advertising, advertisers have the strongest motivation to counteract click fraud.

2.1.4 Web standards and feature detection techniques

The main functionality of a browser is to retrieve remote resources (HTML, style, and media) from web servers and present those resources back to a user [4]. To correctly parse and render the retrieved HTML document, a browser should be compliant with HTML, CSS, DOM, and JavaScript standards which are represented by scriptable objects. Each object is attached with features including properties, methods, and events. For instance, the features attached to the DOM object include `createAttribute`, `getElementsByTagName`, `title`, `domain`, `url`, and many others. Every modern browser supports those features. However, different browser vendors (and different versions) vary in support levels for those web standards, or they implement proprietary extensions all their own. To ensure that websites are displayed properly in all mainstream browsers, web developers usually use a common technique called feature detection to help produce JavaScript code with cross-browser compatibility.

Feature detection is a technique that identifies whether a feature or capability is supported by a browser’s particular environment. One of the common techniques used is reflection. If the browser does not support a particular feature, JavaScript engines return null when referencing the feature; otherwise, JavaScript returns a non-null string. For instance, if the JavaScript statement “document.createElement” returns null in a specific browser, it indicates that the browser does not support the method createElement attached to the document object. Likewise, by testing a browser against a large number of fundamental features specified in web standards for modern browsers, we can estimate the browser’s support level for those web standards, which helps validate the authenticity of the execution environment as a real browser.

Feature detection techniques have three primary advantages. First, feature detection can be an effective mechanism to detect clickbots. A clickbot cannot “pass” the feature detection unless it has implemented the main functionality of a real browser. Second, feature detection stresses the client’s functionality thoroughly, and even a large pool of features can be used for feature detection in a fast and efficient manner. Lastly, the methods used for feature detection are designed to work across different browsers and will continue to work over time as new browsers appear, because new browsers fundamentally support reflection—even before implementing other features—and should also extend, rather than replace, existing web standards.

2.2 Methodology

Our approach mainly challenges a visiting client and its user engagement on the advertised site to determine whether the corresponding ad click is valid or not. To

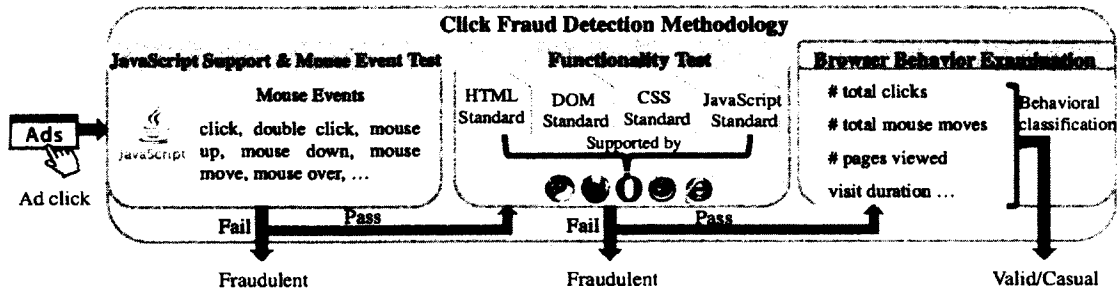


Figure 2.2: Outline of click fraud detection mechanism

maximize detection accuracy, we also check the legitimacy of the origin (client’s IP address) and the intermediate path (i.e., the publisher) of a click.

Figure 2.2 provides an outline of our approach. Our detection system consists of three components: (1) JavaScript support and mouse event test, (2) browser functionality test, and (3) browsing behavior examination.

For each incoming user, on the landing page, we test if the client supports JavaScript and if any mouse events are triggered. No JavaScript support or no mouse event indicates that the client may not be a real browser but a clickbot. Otherwise, we further challenge the client’s functionality against the web standards widely supported by mainstream browsers. The client failed the functionality test is labelled as a clickbot. Otherwise, we further examine the client’s browsing behavior on the advertiser’s website and train a behavior-based classifier to distinguish a really interested user from a casual one.

2.2.1 JavaScript support and mouse event test

One simple way to detect clickbots is to test whether a client supports JavaScript or not. This is due to the fact that at least 98% of web browsers have JavaScript enabled [5] and online advertising services usually count on JavaScript support.

Monitoring mouse events is another effective way to detect clickbots. In general,

a human user with a non-mobile platform (laptop/desktop) must generate at least one mouse event when browsing a website. A lack of mouse events flags the visiting client as a clickbot. However, this may not be true for users from mobile platforms (smartphones/pads). Thus, we only apply the mouse event test to users from non-mobile platforms.

2.2.2 Functionality test

A client passing the JavaScript and mouse event test is required to further undergo a feature-detection based functionality test.

Chrome(10)	1.0.154	2.0.173	4.0.223	5.0.307.1	8.0.552.215
	4/24/2009	6/23/2009	10/24/2009	1/30/2010	12/2/2010
	12.0.742.100	16.0.912.63	20.0.1132.47	24.0.1312.57	27.0.1453.94
Firefox(10)	2.0	3.0	3.5	3.6	4.0
	10/24/2006	6/17/2008	6/30/2009	1/21/2010	3/22/2011
	7.0	11.0	15.0	19.0.2	20.0.1
IE(5)	6.0	7.0	8.0	9.0	10.0
	8/27/2001	10/18/2006	3/19/2009	3/14/2011	10/26/2012
	3.1	3.2	3.2.2	4.0	4.0.5
Safari(10)	3/18/2008	11/14/2008	2/15/2009	6/18/2009	3/11/2010
	5.0.1	5.0.3	5.1	5.1.2	5.1.7
	7/28/2010	11/18/2010	7/20/2011	11/30/2011	5/9/2012
Opera(10)	8.50	9.10	9.20	9.50	10.00
	9/20/2005	12/18/2006	4/11/2007	6/12/2008	9/1/2009
	10.50	11.00	11.50	12.00	12.15
	3/2/2010	12/16/2010	6/28/2011	6/14/2012	4/4/2013

Table 2.1: Tested browsers, versions and release dates

To avoid false positives and ensure that each modern browser can pass the functionality test, we perform an extensive feature support measurement on the top 5 mainstream browsers [6]: Chrome, Firefox, IE, Safari, and Opera. To discern the consistently supported features, we uniformly select 10 versions for each browser vendor with the exception of 5 versions for IE. Table 2.1 lists the browsers we tested. As a result, we obtain a set of 153 features associated with web standards,

including browser window, DOM, and CSS (see Table 2.2). All those features are supported by both desktop browsers and their mobile versions. These features are commonly and consistently supported by the 45 versions of browsers in the past ten years. We call this set the authentic-feature set. We also create a bogus-feature set, which has the same size as the authentic-feature set but is obtained by appending “123” to each feature in the authentic-feature set. Thus, every feature in the bogus-feature set should not be supported by any real browser. Note that we just use the string “123” as an example. When implementing our detection, the advertiser should periodically change the string to make the bogus-feature set hard to evade.

Objects	Features
Browser Window (51)	closed, defaultStatus, document, frames, history, alert, blur, clearInterval, clearTimeout, close, confirm, focus, moveBy, moveTo, open, print, prompt, resizeBy, resizeTo, scroll, scrollBy, scrollTo, setInterval, setTimeout, appCodeName, appName, appVersion, cookieEnabled, platform, userAgent, javaEnabled, availHeight, availWidth, colorDepth, height, width, length, back, forward, go, hash, host, hostname, href, pathname, port, protocol, search, assign, reload, replace
DOM(26)	doctype, implementation, documentElement, createElement, createDocumentFragment, createTextNode, createComment, createAttribute, getElementsByTagName, title, referrer, domain, URL, body, images, applets, links, forms, anchors, cookie, open, close, write, writeln, getElementById, getElementsByName
CSS(76)	backgroundAttachment, backgroundColor, backgroundImage, backgroundRepeat, border, borderStyle, borderTop, borderRight, borderBottom, borderLeft, borderTopWidth, borderRightWidth, borderBottomWidth, borderLeftWidth, borderWidth, clear, color, display, font, fontFamily, fontSize, fontStyle, fontVariant, fontWeight, height, letterSpacing, lineHeight, listStyle, listStyleImage, listStylePosition, listStyleType, margin, marginTop, marginRight, marginBottom, marginLeft, padding, paddingTop, paddingRight, paddingBottom, paddingLeft, textAlign, textDecoration, textIndent, textTransform, verticalAlign, whiteSpace, width, wordSpacing, backgroundPosition, borderCollapse, borderTopColor, borderRightColor, borderBottomColor, borderLeftColor, borderTopStyle, borderRightStyle, borderBottomStyle, borderLeftStyle, bottom, clear, clip, cursor, direction, left, minHeight, overflow, pageBreakAfter, pageBreakBefore, position, right, tableLayout, top, unicodeBidi, visibility, zIndex

Table 2.2: Authentic feature set widely supported by modern browsers

How to perform the functionality test. Figure 2.3 illustrates how the functionality test is performed. For the first HTTP request issued by a client, the advertiser’s web server challenges the client by responding as usual, but along with a mixed set of authentic and bogus features. While the size of the mixed set is fixed (e.g.,100), the proportion of authentic features in the set is randomly decided.

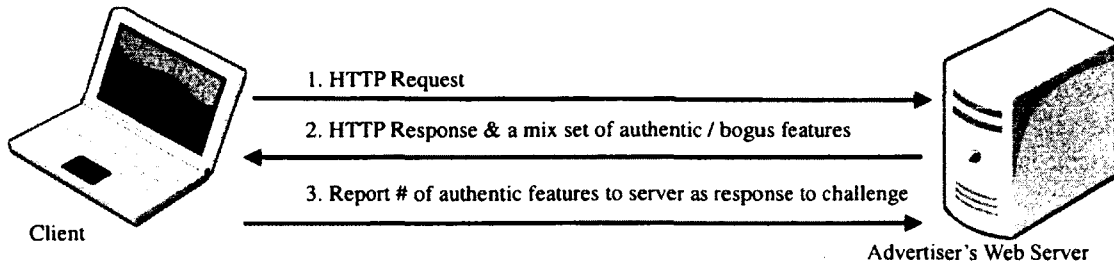


Figure 2.3: How the functionality test is performed by advertiser's web server.

Then, those individual authentic and forged features in the set are randomly selected from the authentic and bogus feature sets, respectively. The client is expected to test each feature in its environment and then report to the web server how many authentic features are in the mixed set as the response to the challenge.

A real browser should be able to report the correct number of authentic features to the web server after executing the challenge code, and thus passes the functionality test. However, a clickbot would fail the test because it is unable to test the features contained in the set and return the correct number. Considering some untested browsers may not support some authentic features, we set up a narrow range $[x - N, x]$ to handle this, where x is the expected number and N is a small non-negative integer. A client is believed to pass the test as long as its reported number falls within $[x - N, x]$. Here we set N to 4 based on our measurement results.

Evasion analysis. Assume that a client receives a mixed set of 150 features from a web server and the set consists of 29 randomly selected authentic features and 121 randomly selected bogus features. Thus, the expected number should fall into the range $[25, 29]$. Consider a crafty clickbot who knows about our detection mechanism in advance. The clickbot does not need to test the features, but just guesses a number from the possible range $[0, 150]$, and returns it to the server. In this case, the probability for the guessed number to successfully fall into $[25, 29]$ is

only 3%. Thus, the clickbot has little chance (3%) to bypass the functionality test.

2.2.3 Browsing behavior examination

Passing the functionality test cannot guarantee that a click is valid. An advanced clickbot may function like a real browser and thus can circumvent the functionality test. A human clicker with a real browser can also pass the test.

However, clickbots and human clickers usually show quite different browsing behaviors on the advertised website from those of real users. Click fraud activities conducted by clickbots usually end up with loading the advertiser’s landing page and do not show human behaviors on the site. For human clickers, their only purpose is to make more money by clicking on ads as quickly as possible. They tend to browse an advertised site quickly and then navigate away for the next click task. Instead, real interested users tend to learn more about a product and spend more time on the advertised site. They usually scroll up and down a page, click on their interested links, browses multiple pages, and sometimes make a purchase.

Therefore, we leverage users’ browsing behaviors on the advertised site to detect human clickers and advanced clickbots. Specifically, we extract extensive features from passively collected browsing traffic on the advertised website, and train a classifier for detection.

2.3 Experimental Results

In order to evaluate our approach, we run ad campaigns to collect real-world click traffic, and then analyze the collected data to discern its primary characteristics, resulting in a technique to classify click traffic as either fraudulent, casual, or valid.

Anchor Groundhog Estate
www.sawmillcreek.org
Variance Flock Accurate Chandelier
Cradle Naphtha Librettist Headwind

Figure 2.4: A bait ad with the ad text of randomly selected English words

2.3.1 Running ad campaigns

To obtain real-world click traffic, we signed up with a major ad network and ran ad campaigns for a high-traffic woodworking forum website. Motivated by the bait ad technique proposed in [18], we created three bait ads for the site and made the same assumption as the previous works [15, 16, 18], that very few people would intentionally click on the bait ads and those ads are generally clicked by clickbots and fraudulent human clickers. Bait ads are textual ads with nonsense content, as illustrated in Figure 2.4. Note that our bait ads were generated in English. In addition, we created two normal ads, for which the ad texts describe the advertised site exactly. Our goal of running ad campaigns is to acquire both malicious and authentic click traffic for validating our click fraud detection system. To this end, we set the bait ads to be displayed on partner websites of any language across the world but display normal ads only on search result pages in English to avoid publisher fraud cases from biasing the clicks on the latter normal ads. We expect that most, if not all, clicks on bait ads and normal ads are fraudulent and authentic, respectively.

We ran our ad campaigns for 10 days. Table 2.3 provides a summary of our ad campaigns. Our ads had 2 million impressions², received nearly 11 thousand clicks and had a click-through rate (CTR) of 0.53% on average. Among these, 2.7 thousand clicks were considered by the ad network as illegitimate and were not charged. The

²An ad being displayed once is counted as one impression.

Set	Campaign	Clicks	Impressions	CTR	Invalid Clicks	Invalid Rate	Avg. CPC	Daily Budget	Duration (days)
1	bait1	1,011	417,644	0.24%	425	29.60%	\$0.08	\$15.00	10
2	bait2	4,127	646,152	0.64%	852	17.11%	\$0.03	\$15.00	10
3	bait3	5,324	933,790	0.57%	1,455	21.46%	\$0.04	\$15.00	10
4	normal1	288	68,425	0.42%	18	5.88%	\$0.40	\$20.00	10
5	normal2	224	20,784	1.08%	10	4.27%	\$0.48	\$20.00	10
Total	NA	10,974	2,086,795	0.53%	2,760	25.15%	\$0.06	\$85.00	10

Table 2.3: Summary of our ad campaigns

invalid click rate was 25.15%. The average cost per click (CPC) was \$0.06. Note that the two normal ads only received 512 clicks accounting for 4.67% of the total. The reason is that although we provided quite high bids for normal ads, our normal ads still cannot compete with those of other advertisers for top positions and thus received fewer clicks.

2.3.2 Characterizing the click traffic

We characterize the received click traffic by analyzing users' geographic distribution, browser type, IP address reputation, and referrer websites' reputations. Our goal, through statistical analysis, is to have a better understanding of both the users who clicked on our ads and the referrer websites where our ads were clicked. Although the ad network reported that our ads attracted close to 11 thousand clicks, we only caught on the advertised site 9.9 thousand clicks, which serve as data objects for both closer examination and validation of our approach.

Geographic distribution. We obtain users' geographic information using an IP geolocation lookup service [7]. Our 9.9 thousand clicks originate from 156 countries. Figure 2.5 shows the distribution of ad clicks by the top 10 countries which generate the most clicks. The distribution of normal daily visitors to the advertised site by country is also given in Figure 2.5. Note that the data form "X/Y" means

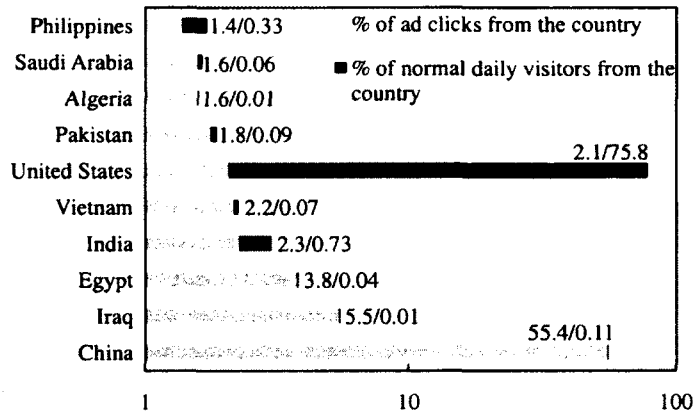


Figure 2.5: Distribution of click traffic vs. that of normal traffic by country

that $X\%$ of ad clicks and $Y\%$ of normal daily visitors are from that specific country. The top 10 countries contribute 77.7% of overall clicks. China alone contributes over 55% of the clicks, while the United States contributes 2.1%. This is quite unusual because the normal daily visitors from China only account for 0.11% while the normal visitors from the United States close to 76%. Like China, Egypt, Iraq, and other generally non-English countries also contribute much higher shares of ad click traffic than their normal daily traffic to the site. The publisher websites from these countries are suspected to be using bots to click on our ads. Even worse, one strategy of our ad network partner may aggravate the fraudulent activities. The strategy says that when an ad has a high click through ratio on a publisher website, the ad network will deliver the ad to that publisher website more frequently. To guarantee that our ads attract as many clicks as possible within a daily budget, the ad network may deliver our ads to those non-English websites more often.

Browser type. Next we examine the distribution of the browsers to see which browser vendors are mostly used by users to view and click on our ads. We extracted the browser information from the User-Agent strings of the HTTP requests to our advertised website.

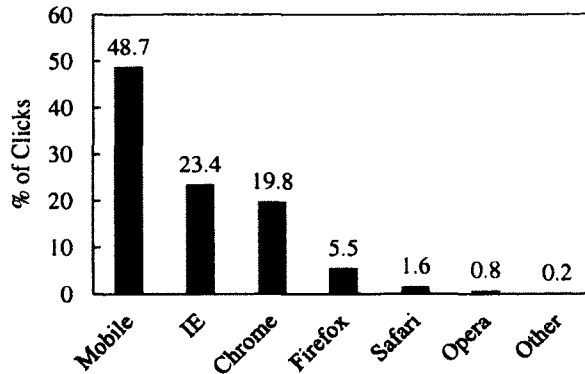


Figure 2.6: Distribution of click traffic by browser

Figure 2.6 shows the distribution of the browsers used by our ad clickers. IE, Chrome, Firefox, Safari, and Opera are the top 5 desktop and laptop browsers, which is consistent with the web browser popularity statistics from StatCounter [6]. Notably, mobile browsers alone contribute to nearly 50% of overall traffic, much larger than the estimated usage share of mobile browsers (about 18% [8]). Close scrutinization reveals that 40% of the traffic with mobile browsers originates from China. China generated over 50 percent of overall traffic, which skews the browser distribution.

Blacklists. A fraction of our data could be generated by clickbots and compromised hosts. Those malicious clients could also be utilized by fraudsters to conduct other undesirable activities, and are thus blacklisted. By looking up users' IP addresses in public IP blacklists [9], we found that 29% of the total hosts have ever been blacklisted.

Referrers. Another interesting question would be which websites host our ads and if their contents are really related to the keywords of our ads. According to the contextual targeting policy of the ad network, an ad should be delivered to the ad network's partner websites whose contents match the selected keywords for the ad.

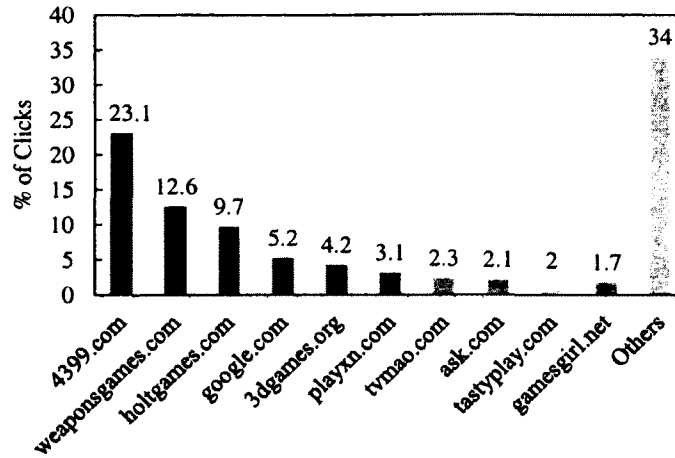


Figure 2.7: Distribution of click traffic by publisher

We used the Referer field in the HTTP request header to locate the publishers that displayed our ads and then directed users to our advertised website. However, we can only identify publishers for only 37.2% of the traffic (3,685 clicks) because the remaining traffic either has a blank Referer field or has the domain of the ad network as the referer field. For example, the Referer field for more than 40% of traffic has the form of doubleclick.net. We then examined, among those detected publishers, which websites contribute to the most clicks. Note that publishers could be websites or mobile apps. We identified 499 unique websites and 5 apps in total. Those apps are all iPhone apps and only generate 28 clicks all together. The remaining 3,657 clicks are from the 499 unique websites. Figure 2.7 shows the distribution of the click traffic by those 504 publishers. The top 3 websites with the most clicks on our ads are all small game websites, which contribute to over 45% of publisher-detectable clicks. Actually, the top 7 websites are all small game websites. Small game websites often attract many visitors, and thus the ads on those websites are more likely to be clicked on. However, our keywords are all woodworking-related and evidently, the contents of those game websites do not match our keywords. According to

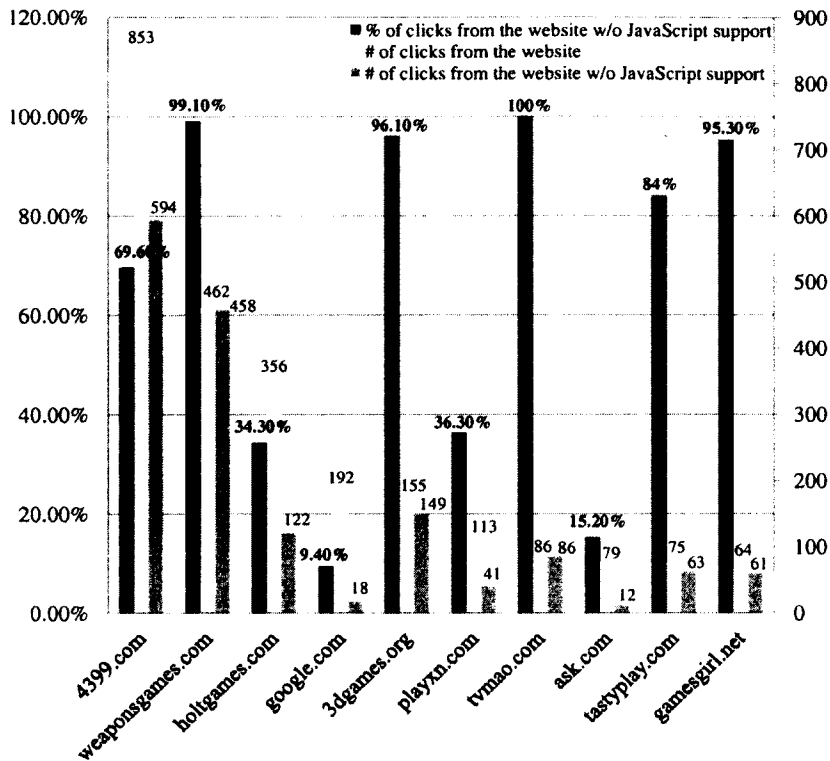


Figure 2.8: Percentage of clicks without JavaScript support for the top 10 publisher websites contributing the most clicks

the above mentioned contextual targeting policy, the ad network should have not delivered our ads to such websites. One possible reason is that from the perspective of the ad network, attracting clicks takes precedence over matching the ads with host websites.

2.3.3 Validating detection approach

As described before, our approach is composed of three main components: a JavaScript support and mouse event test, a functionality test, and a browsing behavior examination. Here we individually validate their effectiveness.

JavaScript support and mouse event test. Among the 9.9 thousand ad

clicks logged by the advertised site, 75.2% of users do not support JavaScript. We labelled those users as clickbots. Note that this percentage may be slightly overestimated considering that some users (at most 2% [5]) may have JavaScript disabled. In addition, those visits without support for JavaScript do not correlate with visits from mobile browsers. We have checked that nearly all mobile browsers provide support for JavaScript despite limited computing power. We then focused on the top 10 publisher websites with the most clicks to identify potentially malicious publishers. Figure 2.8 depicts the percentage of clicks without script support from those top 10 publishers. Among them, the two non-entertainment websites google.com and ask.com have low ratios, 9.4% and 15.2%, respectively. In contrast, the other 8 entertainment websites have quite high click ratios without script support. There are 86 visits from tvmao.com and none of them support JavaScript. We believe that all 86 clicks are fraudulent and generated by bots. Similarly, 99.1% of clicks from weaponsgames.com, 96.1% of clicks from 3dgames.org, and 95.3% from games-girl.net are without JavaScript support either. Such high ratios indicate that the invalid click rate in the real-world ad campaigns is much larger than the average invalid rate of 25.15% alleged by the ad network for our ad campaigns, as shown in Table 2.3.

We observed 506 ad clicks (with JavaScript support) that result in zero mouse events when arriving at our target site. Of those, 96 are initiated from mobile platforms including iPad, iPhone, Android, and Windows Phone. The remaining 410 clicks are generated from desktop or laptop platforms. Those 410 ad clicks also have few other kinds of user engagement: no mouse clicks, no page scrolls, and short dwelling time. We labelled them as clickbots.

We further investigated the click traffic from 4399.com due to the fact that this

website generated the most clicks on our ads among all identified publishers. The following several pieces of data indicate the existence of publisher fraud. First, all 853 clicks from 4399.com were generated within one day. Notably, up to 95 clicks were generated within one hour. Second, several IPs were found to click on our ads multiple times within one minute using the same User-Agent, and one User-Agent was linked to almost 15 clicks on average. Third, close to 70% of clients did not support JavaScript. Hence we suspect that the website owner used automated scripts to generate fraudulent clicks on our ads. However, the scripts are likely incapable of executing the JavaScript code attached to our ads. In addition, they probably spoofed IP address and User-Agent fields in the HTTP requests to avoid detection.

Functionality test. The clickbots that cannot work as full-fledged modern browsers are expected to fail our functionality test. Among the logged 9.9 thousand clicks, 7,448 clicks without JavaScript support did not trigger the functionality test, and 35 of the remaining clicks with JavaScript support were observed to fail the functionality test and were subsequently labelled as clickbots. So far, 75.6% of clicks (7,483 clicks) had been identified by our detection mechanism to originate from clickbots. Among them, 99.5% (7,448 clicks) were simple clickbots without JavaScript support; and the rest 0.5% (35 clicks) were relatively advanced clickbots with JavaScript support yet failed the functionality test.

Browsing behavior examination. After completing the two steps above and discarding incomplete click data, 1,479 ad clicks (14.9 %) are left to be labelled. Among them, 1,127 ad clicks are on bait ads while the other 352 clicks are on normal ads. Here we further classify the click traffic into three categories—fraudulent, casual, and valid—based on user engagement, client IP, and publisher reputation

information.

Features. We believe that three kinds of features are effective to differentiate advanced clickbots and human clickers from real users. (1) How users behave at the advertised site, i.e., users’ browsing behavior information. (2) Who clicks on our ads, and a host with a bad IP is more likely to issue fraudulent clicks. (3) Where a user clicks on ads, and a click originating from a disreputable website tends to be fraudulent. Table 2.4 enumerates all the features we extracted from each ad click traffic to characterize users’ browsing behaviors on the advertised site.

Feature Category	Feature Description
Mouse clicks	# of total clicks made on the advertised site
	# of clicks made only on the pages excluding the landing page
	# of clicks exclusively made on hyperlinks
Mouse scrolls	# of scroll events in total
	# of scroll events made on the pages excluding the landing page
Mouse moves	# of mousemove events in total
	# of mousemove events made only on the pages excluding the landing page
Pages views	# of pages viewed by a user
Visit duration	How long a user stays on the site
Execution efficiency	Client’s execution time of JavaScript code for challenge
Legitimacy of origin	If the source IP is in any blacklist
Publisher’s reputation	If the click originates from an disreputable website

Table 2.4: Features extracted for each ad click

Ground truth. Previous works [18, 15, 16] all assume that very few people would intentionally click on bait ads and only clickbots and human clickers would click on such ads. That is, a click on a bait ad is thought to be fraudulent. However, this assumption is too absolute. Consider the following situation. A real user clicks on a bait ad unintentionally or just out of curiosity, without malicious intention. Then, the user happens to like the advertised products and begins browsing the advertised site. In this case, the ad click generated by this user should not be labelled as fraudulent. Thus, to minimize false positives, we partly accept the above common assumption, scrutinize those bait ad clicks which have shown rich human behaviors on the advertised site, and correct a-priori labels based on the following

heuristics. Specifically, for a bait ad click, if the host IP address is not in any blacklist and the referrer website has a good reputation, this ad click is relabelled as valid when one of the following conditions holds: (1) 30 seconds of dwelling time, 15 mouse events, and 1 click; (2) 30 seconds of dwelling time, 10 mouse events, 1 scroll event, and 1 click; and (3) 30 seconds of dwelling time, 10 mouse events, and 2 page views. We believe the above conditions are strict enough to avoid mislabelling the ad clicks generated by bots and human clickers as valid clicks.

Note that our normal ads are only displayed on the search engine result pages with the expectation that most, if not all, clicks on normal ads are valid. The ad campaign report provided by the ad network in Table 2.3 confirms this, showing that the invalid click rate for normal ads is only 5.08% on average. Based on our design and the ad campaign report, we basically assume that the clicks on normal ads are valid. However, after further manually checking the normal ad clicks, we found that some of them do not demonstrate sufficient human behaviors, and these normal ad clicks will be relabelled as casual when one of the following two conditions holds: (1) less than 5 seconds of dwelling time; (2) less than 10 seconds of dwelling time and less than 5 mouse events. The casual click traffic could be issued by human users who unintentionally click on ads and then immediately navigate away from the advertised site. From the advertisers' perspective, such a click traffic does not provide any value when evaluating the ROI of their ad campaigns on a specific ad network, and therefore should be classified as casual.

Actually, if there is no financial transaction involved, only a user's intention matters whether the corresponding ad click is fraudulent or not. That is, only users themselves know the exact ground truth for fraudulent/valid/casual clicks. For those clicks without triggering any financial transactions, we utilize the above

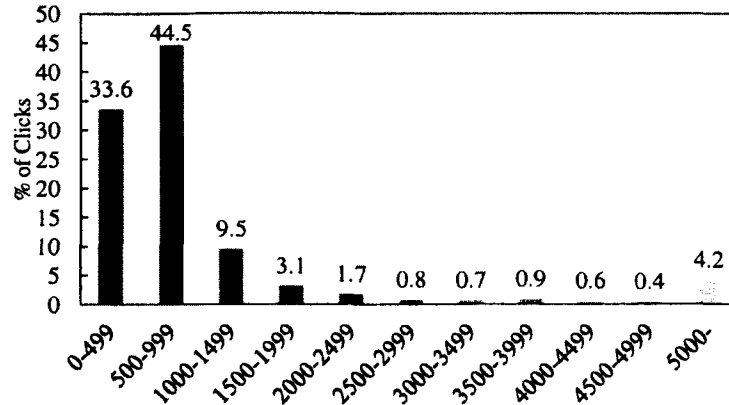


Figure 2.9: Clients’ execution time of JavaScript challenge code in milliseconds

reasonable assumptions and straightforward heuristics to form the ground truth for fraudulent/valid/casual clicks.

Evaluation metrics. We evaluated our detection against two metrics—false positive rate and false negative rate. A false positive is when a valid click is wrongly labelled as fraudulent, and a false negative is when a fraudulent click is incorrectly labelled as valid.

Classification results. Using Weka [10], we chose a C4.5 pruned decision tree [24] with default parameter values (i.e., 0.25 for confidence factor and 2 for minimum number of instances per leaf) as the classification algorithm, and ran a 10-fold cross-validation. The false positive rate and false negative rate were 6.1% and 5.6%, respectively. Note that these are the classification results on those 1,479 unlabelled clicks. As a whole, our approach showed a high detection accuracy on the total 9.9 thousand clicks, with a false positive rate of 0.79% and a false negative rate of 5.6%, and the overall detection accuracy is 99.1%.

Overhead. We assessed the overhead induced by our detection on the client and server sides, in terms of time delay, CPU, memory and storage usages.

The only extra work required of the client is the execution of a JavaScript chal-

challenge script and to report the functionality test results to the server as an AJAX POST request. We measured the overhead on the client side using two metrics: source lines of code (SLOC) and the execution time of JavaScript code. The JavaScript code is only about 150 SLOC and we observed negligible impact on the client. We also estimated the client's execution time of JavaScript from the server side to avoid the possibility that the client could report a bogus execution time. Note that the execution time measured by the server contains a round trip time, which makes the estimated execution time larger than the actual execution time. Figure 2.9 depicts the 9.9 thousand clients' execution time of the JavaScript challenge code. About 80% of clients finished execution within one second. Assuming that the round trip time (RTT) is 200 milliseconds, the actual computation overhead incurred at the client side is merely several hundred milliseconds.

We used the SAR (System Activity Report) [11] to analyze server performance and measure the overhead on the server side. We observed no spike in server load. This is because most of work involved in our detection happens on the client side, and the induced click-related traffic is insignificant in comparison with server's normal traffic.

2.4 Discussion and Limitations

In this project, we assume that a clickbot typically does not include its own JavaScript engine or access the full software stack of a legitimate web browser residing on the infected host. A sophisticated clickbot implementing a full browser agent itself would greatly increase its presence and the likelihood of being detected. A clickbot might also utilize a legitimate web browser to generate activities, and can thus pass our browser functionality test. To identify such clickbots, we could further detect

whether our ads and the advertised websites are really visible to users by utilizing a new feature provided by some ad networks. The new feature allows advertisers to instrument their ads with JavaScript code for a better understanding of what is happening to their ads on the client side. With this feature, we could detect if our ad iframe is visible at the client's front-end screen rather than in the background, and if it is really focused and clicked on.

In addition, compared to our user-visit related features (dwelling time, mouse events, scroll events, clicks and etc.), user-conversation related features³ are expected to have better discriminating power between clickbots, human clickers, and real users in browsing behaviors. However, our advertised site is a professional forum rather than an online retailer. If a user registers (creates an account) on the forum, it is analogous to a purchase at an online retailer. However, such conversion from guest to member is an event too rare to rely upon to enhance our classifier.

2.5 Related Work

Browser fingerprinting. Browser fingerprinting allows a website to identify a client browser even though the client disables cookies. Existing browser fingerprinting techniques could be mainly classified into two categories, based on the information they need for fingerprinting. The first category fingerprints a browser by collecting application-layer information, including HTTP request header information and system configuration information from the browser [17]. The second category performs browser fingerprinting by examining coarse traffic generated by the browsers [26]. However, both of them have their limitations in detecting clickbots. Nearly all the application-layer information can be spoofed by sophisticated

³Purchasing a product, abandoning an online cart, proactive online chat, etc.

clickbots, and browser fingerprints may change quite rapidly over time [17]. In addition, an advertiser often cannot collect enough traffic information for fingerprinting the client from just one visit to the advertiser. Compared to the existing browser fingerprinting techniques, our feature detection technique has three main advantages. First, clickbots cannot easily pass the functionality test unless they have implemented the main functionality present in modern browsers. Second, the client's functionality could be tested thoroughly at the advertiser's side even though the client visits the advertiser's landing page only once. Lastly, our technique works over time as new browsers appear because new browsers should also conform to the those web standards currently supported by modern browsers.

Revealed click fraud. Several previous studies investigate known click fraud activities, and clickbots have been found to be continuously evolving and become more sophisticated. As the first study to analyze the functionality of a clickbot, Daswani et al. [14] dissected Clickbot.A and found that the clickbot could carry out a low-noise click fraud attack to avoid detection. Miller et al. [23] examined two other families of clickbots. They found that these two clickbots were more advanced than Clickbot.A in evading click fraud detection. One clickbot introduces indirection between bots and ad networks, while the other simulates human web browsing behaviors. Some other characteristics of clickbots are described in [15]. Clickbots generate fraudulent clicks periodically and only issue one fraudulent click in the background when a legitimate user clicks on a link, which makes fraudulent traffic hardly distinguishable from legitimate click traffic. Normal browsers may also be exploited to generate fraudulent click traffic. The traffic generated by a normal browser could be hijacked by currently visited malicious publishers and be further converted to fraudulent clicks [19]. Ghost click botnet [12] leverages DNS changer

malware to convert a victim's local DNS resolver into a malicious one and then launches ad replacement and click hijacking attacks. Our detection can identify each of these clickbots by actively performing a functionality test and can detect all other kinds of click fraud by examining their browsing behavior traffic on the server side.

Click fraud detection. Metwally et al. conducted an analysis on ad networks' traffic logs to detect publishers' non-coalition hit inflation fraud [21], coalition fraud [20], and duplicate clicks [22]. The main limitation of these works lies in that ad networks' traffic logs are usually not available to advertisers. Haddadi in [18] and Dave et al. in [15] suggested that advertisers use bait ads to detect fraudulent clicks on their ads. While bait ads have been proven effective in detection, advertisers have to spend extra money on those bait ads. Dave et al. [16] presented an approach to detecting fraudulent clicks from an ad network's perspective rather than an advertiser's perspective. Li et al. [19] introduced the ad delivery path related features to detect malicious publishers and ad networks. However, monitoring and reconstructing the ad delivery path is time-consuming and difficult to detect click frauds in real time. Schulte et al. [25] detected client-side malware using so-called program interactive challenge (PIC) mechanism. However, an intermediate proxy has to be introduced to examine all HTTP traffic between a client and a server, which would inevitably incur significant delay. Like [18, 15], our defense works at the server side but does not cause any extra cost for advertisers. Our work is the first to detect clickbots by testing their functionalities against the specifications widely conformed to by modern browsers. Most clickbots can be detected at this step, because they have either no such functionalities or limited functionalities compared to modern browsers. For the advanced clickbots and human clickers, we scrutinize

their browsing behaviors on the advertised site, extract effective features, and train a classifier to identify them.

2.6 Conclusion

In this project, we have proposed a new approach for advertisers to independently detect click fraud activities issued by clickbots and human clickers. Our proposed detection system performs two main tasks of proactive functionality testing and passive browsing behavior examination. The purpose of the first task is to detect clickbots. It requires a client to actively prove its authenticity of a full-fledged browser by executing a piece of JavaScript code. For more sophisticated clickbots and human clickers, we fulfill the second task by observing what a user does on the advertised site. Moreover, we scrutinize who initiates the click and which publisher website leads the user to the advertiser's site, by checking the legitimacy of the clients' IP addresses (source) and the reputation of the referring site (intermediate), respectively. We have implemented a prototype and deployed it on a large production website for performance evaluation. We have then run a real ad campaign for the website on a major ad network, during which we characterized the real click traffic from the ad campaign and provided advertisers a better understanding of ad click traffic, in terms of geographical distribution and publisher website distribution. Using the real ad campaign data, we have demonstrated that our detection system is effective in the detection of click fraud.

Chapter 3

E-commerce Reputation

Manipulation: The Emergence of Reputation-Escalation-as-a-Service

In this chapter, we present our study on a newly emerging underground industry, so called SRE markets, in which a potentially unbounded number of inexpensive human laborers are hired by e-commerce sellers to conduct fake purchases for reputation inflation. By fake purchases, we mean purchases that although they appear legitimate and complete as far as the online system is concerned, no real product or at most an empty package is delivered by the seller. This approach is far more elaborate and much more difficult, if not infeasible, to detect because the buyer appears to have genuinely purchased the product as opposed to just leaving a review or score for the product and the seller. Moreover, multiple individuals that do not know each other are involved in the process.

To provide an empirical analysis of the prevalence of the SRE markets, in this project, we answer some quantitative and qualitative questions about their current

operations and structure: How popular are SRE markets with online sellers? What strategies have SRE markets forged for online sellers to evade fake transaction detection in online marketplaces? What kind of online sellers are involved in this shadowy ecosystem? How active are sellers on SRE markets? What is their actual effectiveness? Has the sellers' reputation been escalated as a result? What is the worker population? How much can a worker earn daily? What amount of fake-transaction volumes is handled by SRE markets monthly? How much revenue is generated by them each month? Are there any opportunities for disrupting the value chain of the black economy?

We organize the chapter as follows. Section 3.1 provides a brief introduction to Taobao and an overview of the business model of SRE markets. Section 3.2 describes our data collection methodology and data we collected. Section 3.3 presents the results from our infiltration of five SRE markets and our insights into the shadowy business. We evaluate the effectiveness of SRE services in Section 3.4, and discuss possible defensive interventions and limitations in Section 3.5. We revisit the SRE ecosystem one year later and present our new findings in Section 3.6. We review the related work in Section 3.7, followed by our conclusions in Section 3.8.

3.1 Background

The increasingly thriving e-commerce has drawn forth a large number of online sellers. For instance, eBay, the online marketplace giant, has 25 million sellers globally [28]. SRE markets have emerged to satisfy online sellers' demands for high reputation. We identified five SRE markets which provide SRE services exclusively to online sellers on the Taobao marketplace. In this section, we first briefly introduce the Taobao online marketplace and then provide an overview of how a typical SRE

market works.

3.1.1 Taobao Overview

Taobao, launched by Alibaba Group [29] in 2003, is the largest consumer-to-consumer (C2C) online marketplace in China with more than 8.5 million sellers, over one billion product listings, and around 500 million registered users as of March 2015 [30, 31]. Taobao has achieved great success with 60 million daily visitors and 50,000 sales per minute, and is ranked 9th globally by Alexa [32] as of this writing. In 2013, the total gross merchandise traded on Alibaba was more than Amazon and eBay’s gross sales combined [33].

To facilitate shopping on Taobao, Alibaba operates Alipay and AliWangWang. Alipay serves as an online payment system which provides escrow services for buyers — holding buyer’s payment until the buyer is happy with the goods received. AliWangWang is an embedded instant messaging program, commonly used for Taobao buyers to communicate with sellers prior to the purchase.

The great success of Taobao makes it an ideal host for miscreants such as the operators of SRE markets, which accumulate wealth by providing reputation-escalation services to Taobao sellers. Though not yet reported, the other major online marketplaces such as Amazon and eBay may also suffer from fake transactions conducted through SRE markets.

3.1.2 How a Typical SRE Market Works

SRE markets operate in the crowdsourcing mode and are at the center of the shadowy ecosystem, which connects insincere online sellers who desire for high reputation with people who want to earn extra money. However, unlike other crowdsourcing

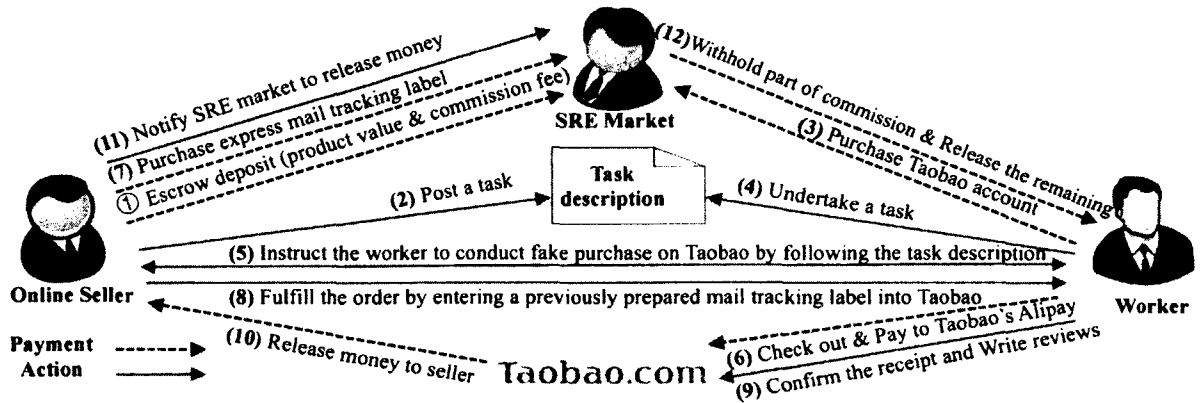


Figure 3.1: Lifecycle of a fake-purchase task on the SRE market.

markets such as Amazon Mechanical Turk (MTurk) [42], these SRE markets only accommodate one kind of task: conducting fake transactions on the specified Taobao stores. According to the terminology used on crowdsourcing markets, online sellers on SRE markets act as task requesters while people undertaking fake-purchase tasks act as task workers.

Figure 3.1 illustrates a typical lifecycle of a fake-purchase task on a SRE market. We classify the lifecycle into five distinct stages: *task creation*, *task undertaken*, *conducting fake purchase on Taobao*, *order fulfillment*, and *commission realization*.

Task Creation. The lifecycle of a fake-purchase task begins with an online seller creating a task on the SRE market. To this end, the seller must first deposit money into the SRE market at the amount equivalent to the sum of the goods' value designated in the task and the commission fee calculated by the SRE market (**Step 1**). Then, the seller creates a task and customizes an associated qualification requirement to limit which workers are qualified to undertake this task (**Step 2**).

Qualification Requirement. The qualification requirement is composed of several qualification types predefined by an SRE market. Table 3.1 lists the qualification types provided by an SRE market. Each one describes a qualification that

Qualification Types

Q1: The worker must have an amount of guaranteed money held by the SRE market.

Q2: The Taobao account used by the worker for taking tasks must be verified and pre-aged.

Q3: The worker's Taobao account must not be used a lot for fake purchases on the SRE market.

Q4: The worker should be proficient, reflected by her score value on the SRE market.

Q5: The worker should be located in the geographic region specified in the task.

Q6: The worker never undertakes prior tasks posted by this seller.

Table 3.1: Qualification types

a worker must have in order to take on the task and is believed to help the seller reduce the risk of being penalized by Taobao for fake transactions. Qualification Q1 requires a worker to have an amount of money held by the SRE market in case a dishonest worker intentionally complains about the seller to Taobao and requests a refund for her purchase even though she gets rewarded for making the fake purchase. Q2 requires the Taobao buyer account used by the worker for undertaking tasks to be verified and pre-aged¹, which allows the seller to evade simple detection heuristics used by Taobao for suspending freshly minted accounts based on weak signs of misbehavior. Q3 requires the worker's Taobao account to not get involved in too many fake purchases because such accounts are probably being closely monitored by Taobao. Q4 requires a worker to be familiar with the task flow. Q5 makes the requirement for workers' geographical distribution represented by IP address to make the fake purchase appear more real. Q6 reflects the seller's effort to diversify the workers to avoid triggering any Taobao alarms.

Task Undertaken. In light of the qualification requirements, a professional worker usually has several pre-aged Taobao accounts on hand and takes turns using them to avoid using one account too often. The demand for Taobao accounts has inspired another service provided by SRE markets. Each SRE market serves as an

¹Pre-aged Taobao account refers to the account that has been created for some time and ever used for real purchases.

account merchant that stockpiles a multitude of Taobao accounts and can sell them to workers at a whim. The prices for Taobao accounts range from \$0.2 to \$0.5 each depending on the account’s age and purchase history. After purchasing a specific number of Taobao accounts from the SRE market (**Step 3**), the worker chooses a qualified task to work on (**Step 4**).

Class	Detail
Goods	Type (physical or virtual), Selling price
Commission	Commission fee offered for this task
Browsing behavior	Search first on Taobao by the keywords given, randomly choose three other stores to browse before finally entering the seller’s store. Like the store and add it to favorites. Stay on the page for 5 minutes and scroll down to the bottom before adding to cart. Feign chat with the seller via Taobao’s built-in IM program AliWangWang.
Payment method	The worker pays either for herself or using the e-Gift card provided by the seller.
Shipping address	Use the shipping address designated by the seller for the order placed.
Confirmation & reviews	Confirm the delivery and leave good ratings and positive reviews after a pre-defined waiting time.

Table 3.2: A typical task description

Conducting Fake Purchase on Taobao. After attaching a Taobao account to the task, the worker first starts chatting with the seller posting the task through a third-party Instant Messaging (IM) program on the SRE site. After further checking the worker’s Taobao account against the qualification requirements, the seller instructs the worker to follow the task description and behave like a real buyer on Taobao (**Step 5**). Table 3.2 enumerates a typical task description. It describes the type of goods to purchase (physical or virtual), its selling price, and the commission fee offered. It also details the required browsing behaviors on Taobao before checking out, the payment method, shipping address, as well as the timing for delivery confirmation and leaving positive reviews.

Order Fulfillment. After finishing all the required actions listed in the task description, the worker gets to the checkout step and provides payment to Taobao’s escrow-based system Alipay using either her own form of allowed payment or the

e-Gift card provided by the seller (**Step 6**). Then the seller arranges to fulfill the order. For virtual goods such as software and prepaid phone cards, the order is directly fulfilled via the Internet, and the worker is required to confirm the receipt and leave good ratings immediately after checkout. For physical goods such as clothes, the seller never ships out the ordered goods but is required by Taobao to provide a mail tracking label for package tracking. To evade detection, for each task with physical goods, the seller purchases one express mail tracking label from the SRE market at a price of \$0.4-0.7, depending on the shipment companies (**Step 7**). SRE markets usually partner with shipment companies to get a stable supply of fresh and unscanned express tracking labels. With the label purchased, the seller inputs the label number into Taobao and hence fulfills the order (**Step 8**). Some sellers may ship an empty package to the designated shipping address while most ship nothing.

Commission Realization. After a predefined wait time elapses, the worker confirms the receipt of the goods on Taobao (**Step 9**). In addition, the worker must rate the seller with a full score and write positive reviews with the contents either specified in the task description or composed by the worker. Then the worker requests Alipay to release money from her Alipay account to the seller's Alipay account (**Step 10**). Subsequently, the seller notifies the SRE market to release the money pre-deposited on the market when posting the task to the worker (**Step 11**). Upon request, the SRE market withholds a portion (typically 20%) of the commission fee offered for the task and then releases the remaining money (the remaining 80% commission fee, along with the reimbursement if the worker paid for the goods with her own money) to the worker (**Step 12**). To this point, the lifecycle of a typical fake-purchase task is completed.

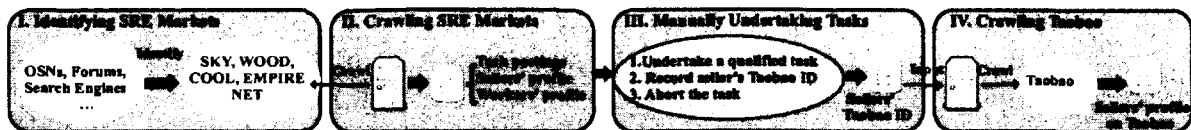


Figure 3.2: Procedure of data collection.

In summary, the worker typically needs a verified and aged Taobao account, invests several minutes and tens of US dollars in purchasing goods, and gets a reward of several dollars after about 3 days; the seller needs to purchase an express mail tracking label if the goods in the task is physical and pays a small commission fee in return for accumulating a transaction and a good rating (review); the SRE market earns money from each task by withholding a part of the commission fee, selling mailing labels to sellers and Taobao accounts to workers. In the lifecycle of a task, two escrow-based payment services — Taobao’s Alipay and the SRE market’s payment system — play a key role in guaranteeing that the worker completes all required actions to earn the commission and that the seller pays a commission fee to the worker for the fake transaction.

3.2 Data Collection Methodology

In this section, we describe our crawling mechanism and summarize the dataset, followed by a discussion on ethical concerns.

3.2.1 Crawling Mechanism

Figure 3.2 briefly illustrates the procedure of our data collection. First, we identify SRE markets (I). Then web crawlers are developed to automatically crawl the identified SRE markets for task postings and the profiles of users (i.e., sellers and

workers) on the SRE markets (II). In order to recognize those sellers' IDs on the Taobao marketplace, we manually undertake tasks for one month (III). For those sellers with their Taobao IDs recognized, we then crawl Taobao for their sales and reputation information to evaluate the impact that their acquisition of both transactions and feedbacks conducted through the SRE markets have upon their Taobao stores (IV). Note that steps I and III are manual operations and the others are fully automated. We detail these steps as follows.

Identifying SRE Markets. SRE markets usually advertise themselves to attract new Taobao sellers and workers through online social networks, public forums, IM group chat, search engines, etc. We investigated these common haunts and identified five SRE markets in total. They are SKY [34], WOOD [35], EMPIRE [36], COOL [37], and NET [38]. All five markets have nearly identical web layout and source code. Whois domain lookup reveals that they have been founded for two to four years. We do not claim that our study covers all SRE markets, which is very challenging if not impossible. However, we believe that the five SRE markets we studied represent a reasonable coverage since they are some of the most active and popular SRE markets.

Crawling SRE Markets. In early February 2014, we first registered an account on each market, then performed a few test crawls, and finally developed automated crawlers which exploit the cookies stored locally by SRE markets to bypass their CAPTCHA mechanisms and login prompts. During the two months from February 21, 2014 to April 21, 2014, we conducted a comprehensive crawl of the five chosen SRE markets. Our crawled data includes task postings and the profiles of both sellers and workers on the SRE markets.

In an attempt to record all task postings, we had to crawl the SRE markets

continuously 24-hours/7-days because a task may immediately become invisible once undertaken by a worker. A task posting specifies what kind of workers are qualified for the task (see Table 3.1 for a list of qualification types) and provides a detailed description of the task including the goods involved, its price, commission fee offered, browsing behaviors required, payment method, and many more listed in Table 3.2. From the collected task postings, we extracted the involved sellers' usernames on a SRE market and further employed the application programming interface (API) provided by the SRE market to crawl once daily for the sellers' profiles on the market. A profile mainly contains the number of tasks ever posted on the market by the seller and when the seller first and last posted tasks. However, the publicly accessible worker-related data is only restricted to a list of the top 10 workers on each market, which is published and daily updated as an excitation mechanism. We crawled once daily for the list of the top 10 workers and further collected their profile information. A worker's profile includes the total number of tasks undertaken, as well as the first and the last time to undertake a task.

Undertaking Tasks. In order to examine the impact that a seller's posting fake-purchase tasks on a SRE market has upon her online store in the Taobao marketplace, we have to figure out the corresponding Taobao ID of the SRE seller. We soon realized that only undertaking her task postings allows us to record her Taobao ID. Unfortunately, we cannot recognize a seller's associated Taobao ID unless we undertake her tasks *manually*. To undertake a task, a Taobao account is required and we used our own legitimate Taobao account. During the month between February 21, 2014 and March 21, 2014, we conducted numerous attempts to undertake a task and then abort the task immediately after the associated Taobao seller ID is recorded. In this way, we were able to identify more than 4,000 Taobao seller

IDs. We failed to identify more seller IDs because either our Taobao account, SRE market account, or geographic IP does not satisfy the qualification requirements of many tasks.

Crawling Taobao. With the identified Taobao seller IDs, we were able to monitor the daily variation in transaction volume and reputation of each of those Taobao stores. To this end, we developed another web crawler and employed the API provided by Taobao to crawl Taobao once daily for those sellers' profile information. A seller's profile on Taobao mainly contains the following information: her seller ID, the major business she runs, the store start date, the current store reputation score, the transaction volume in the recent week (month, semi-year, and year), and customers' rate.

We took a similar recipe as in [47] to make sure our continuous crawl was not noticed by both SRE market operators and Taobao marketplace. Specifically, neither our IP nor accounts on SRE markets were blocked during the period of measurement. Also, we were not contacted by any operator or Taobao to inquire about our browsing activities. So, we believe that our crawled data is valid and not tainted by SRE market operators. With the collected data from SRE markets and Taobao, we were able to examine the SRE market characteristics, evaluate the impact of SRE services upon Taobao stores, and offer insights for designing a robust fake-transaction detection mechanism.

3.2.2 Data Summary

Table 3.3 summarizes the dataset collected on the five SRE markets we infiltrated. Specifically, it enumerates the measurement period, the total number of task postings, the number of active sellers, and the number of sellers with Taobao IDs suc-

cessfully identified. As a result of the two-month collection of the five SRE markets, we collected 219,165 tasks in total, contributed by 11,130 Taobao sellers. Of them, 4,162 sellers' Taobao IDs were identified through our manually undertaking tasks for one month.

Market	Period	Task Posts	Active Sellers	Identified Sellers
SKY	02/21-04/21	63,343	2,789	1,332
WOOD	02/21-04/21	54,824	2,968	1,232
EMPIRE	02/21-04/21	48,120	2,016	706
COOL	03/07-04/21*	39,823	2,419	657
NET	02/21-04/21	13,055	938	235
Total	—	219,165	11,130	4,162

Table 3.3: List of the SRE markets we infiltrated, the months monitored, total task postings, active sellers during the time frame, and Taobao ID identified sellers. *This market went down between 02/21 and 03/06.

A comparison of Taobao seller IDs across the five markets shows that 52 of 4,162 sellers posted tasks on more than one SRE markets. Excluding the overlapping seller IDs, we identified 4,109 unique Taobao seller IDs altogether.

3.2.3 Ethical Considerations

In our study, we identified sellers' Taobao IDs by manually undertaking tasks. We emphasize that we did not purchase or register any fraudulent Taobao accounts but used our own legitimate Taobao accounts to undertake tasks. Furthermore, we never completed a single task but aborted a task immediately after we recorded the seller's Taobao ID. Thus, we did not participate in fake transactions and strictly abided by Taobao's terms and policies. In addition, we did not expose any Taobao IDs identified in this project, and all data crawled from SRE markets and Taobao is publicly available. Therefore, our work will not introduce any additional risk to SRE market operators or their sellers.

3.3 SRE Market Characteristics

Now we present our measurement results of the five SRE markets. We first examine the popularity of SRE markets. Then we investigate the strategies formulated by SRE markets to circumvent Taobao’s detection of fake transactions. Next we characterize two key players on SRE markets: sellers and workers. Finally, we estimate the generated gross revenue and the total fake-transaction volume handled by the five SRE markets during our two-month monitoring.

3.3.1 SRE Market Popularity

Market	Daily active sellers		Daily new tasks		Time to undertake (seconds)	
	Avg.	Max	Avg.	Max	Avg.	Min
SKY	224	313	951	1481	230	2
WOOD	297	381	816	1132	260	1
EMPIRE	222	310	689	1035	243	1
COOL	233	517	663	1843	95	2
NET	59	102	138	276	288	1

Table 3.4: Statistics of daily active sellers, daily new tasks, and the time to undertake a task on the five SRE markets.

We first attempt to measure how attractive SRE markets are to Taobao sellers, in terms of daily active sellers, daily task postings, and how fast a task is undertaken. By active sellers we mean those sellers who post at least one task on a specific day. Table 3.4 lists the statistics of these metrics for each of the five SRE markets. All five markets but NET have more than 200 active sellers per day on average, and as many as 517 active sellers can be observed on the COOL market on a single day. In addition, hundreds of new tasks are posted every day on each market, and the average number of new tasks on the SKY market has almost reached 1,000 per day. The peak number of new daily tasks is observed on the COOL market, with 1,843 posts. Moreover, a newly posted task is usually undertaken very quickly. The

average time for a new task to be undertaken is less than 300 seconds (5 minutes) on all five markets. It is even faster on the COOL market at less than 2 minutes on average. The minimum time for new tasks to be undertaken is within 2 seconds. All these results indicate that SRE markets serve as popular distributors for fake transactions targeting the Taobao marketplace.

3.3.2 Strategies to Evade Taobao Detection

The high popularity of SRE markets among Taobao sellers benefits from those sellers’ confidence that they will not be detected or penalized by Taobao for fake transactions, or at least the risk is quite low. Indeed, all five markets provide a set of similar guidelines for sellers to follow when posting tasks in order to circumvent Taobao’s detection system. Next we investigate Taobao’s detection mechanism and SRE markets’ tit-for-tat strategies.

The details about the implementation of Taobao’s detection algorithm are not publicly available, but many parameters have been learned from previous penalty from Taobao posed for fake transactions. The detection mechanism is believed to cover all steps in a purchase transaction.

Restriction	SKY	WOOD	EMPIRE	COOL	NET
IVA & Aged	7.48	17.65	21.99	23.22	10.98
Use Frequency	15.26	36.78	38.47	39.97	23.20

Table 3.5: Fraction (%) of tasks with restrictions on workers’ Taobao accounts.

Taobao Buyer Accounts. According to Taobao’s report, more than 90% of registered Taobao buyer accounts are ID-verified accounts (IVAs). Non-IVA accounts and newly registered ones would receive special attention. In addition, Taobao buyer accounts with too many purchases within a short time may have been put in some gray lists by Taobao for close monitoring. A Taobao store with a large

portion of transactions from non-IVA accounts, newly generated accounts, or accounts in gray lists would become a suspect of fake transactions. Correspondingly, SRE markets have restricted the number of tasks each Taobao buyer account can undertake per day to be less than six. Also, Taobao sellers on SRE markets can enforce extra restrictions on Taobao buyer accounts. Table 3.5 lists the fraction of tasks with restrictions on Taobao accounts among the total 215,292 tasks we crawled. Note that on the COOL market, 23.22% of tasks require that the workers' Taobao buyer accounts must be IVA and pre-aged, and about 40% have restrictions on the frequency that a worker's Taobao account can be used to undertake tasks.

Restriction	SKY	WOOD	EMPIRE	COOL	NET
Geographic dist.	0.59	9.89	1.87	3.92	1.44
Designated SA	10.09	19.72	8.03	7.68	12.17

Table 3.6: Fraction (%) of tasks with geographic preference and shipping address (SA) designated.

Geographic Distribution & Shipping Address. Without a diverse pool of IP addresses and shipping addresses, fake transactions could be easily spotted by Taobao's detection system. To avoid detection, workers on SRE markets are required to change IP addresses and clean up browser cookies between two consecutive tasks. In addition, Taobao sellers can also set geographical preferences of the workers or require workers to fill in the shipping addresses of sellers' choice to make fake transactions appear geographically distributed. Table 3.6 shows the fractions of tasks with geographic preference and shipping addresses (SA) designated. It shows that a small portion of tasks have restrictions on geographic distribution or shipping address, which is reasonable considering that workers on the crowdsourcing platform have already been very diverse.

Imitating a Real Purchase on Taobao. In addition to buyers' Taobao ac-

counts, IP, and shipping addresses, their browsing behaviors throughout the purchase are also closely monitored. To imitate a real purchase, the majority of tasks on SRE markets require workers to show some of the following actions. (1) *Search to enter*: Search on Taobao for the designated goods with given keywords and locate the seller’s store; randomly choose three other stores to browse first and then enter the seller’s store. (2) *Browse the store*: Browse several other goods first, then browse the designated goods page; scroll down to the bottom of the page and stay five minutes. (3) *Like the store*: Add the store to favorites. (4) *Start a fake chat*: Talk with the seller via Taobao’s built-in IM tool AliWangWang. Finally check out the goods. Table 3.7 summarizes the fraction of tasks with the required actions. Clearly, a large portion of tasks on SRE markets require workers to show at least one kind of browsing behavior of a real buyer. For instance, more than 73% of tasks on all five markets require workers to present at least one of the four actions. And “search to enter” is the most required actions. More than 47% of tasks on all five markets require the workers to enter the stores by first searching on Taobao with given keywords. Our results indicate that SRE markets have paid significant attention to evading human behavior based detection. Thus, defenders should not rely on only one kind of human behavior and should combine human behaviors with other features for accurate detection.

Requirement	SKY	WOOD	EMPIRE	COOL	NET
Search to enter	66.80	51.81	53.43	68.34	47.38
Browse the store	17.69	17.11	41.14	28.42	54.31
Like the store	42.41	30.25	27.85	27.23	28.21
Fake chat	33.32	43.65	34.06	41.21	41.48
One or more	79.46	73.99	80.33	81.80	82.61

Table 3.7: Fraction (%) of tasks with requirements for each kind of browsing behavior before checking out. “One or more” denotes the tasks with at least one required action.

Payment. When checking out, a worker pays for the ordered products with

either an e-Gift card provided by the seller or her own debit card. Credit card payment is usually not recommended on SRE markets to prevent insincere workers from disputing a credit card refund after completing the task. Tasks with e-Gift payment are often undertaken quickly due to no need of monetary investment. Table 3.8 lists the fractions of tasks with the requirements for e-Gift payment and no credit card payment. It is clear that tasks with e-Gift card payment are quite limited, which is reasonable since too many transactions with e-Gift payment on a store may trigger alarm. Also, as expected, tasks with a declaration of no credit card payment are limited, as well. One possible explanation is that credit card payment is not very popular in China, and sellers try to avoid credit card dispute issues. This observation indicates that most workers pay with their debit cards.

Requirement	SKY	WOOD	EMPIRE	COOL	NET
e-Gift payment	0	0.27	6.12	0	0.59
No credit card	4.73	2.32	0	0	0

Table 3.8: Fraction (%) of tasks with requirements for e-Gift card payment or no credit card payment.

Shipping. Taobao monitors package tracking information as well. For each transaction involving physical goods, a seller needs to submit one package tracking number to Taobao for buyer tracking. To create a fake illusion of shipment, a seller purchases a tracking label from an SRE market and submits the tracking number to Taobao. However, no real product or at most an empty package is delivered by the seller. Table 3.9 shows the fraction of tasks declaring to ship empty packages. We found that most tasks ship nothing to buyers, which means a fake tracking number is enough to avoid detection. This may be because it's difficult for online marketplaces to verify the shipment of a package since goods are usually delivered using third-party shipping services.

Requirement	SKY	WOOD	EMPIRE	COOL	NET
Empty package	6.87	2.02	0	0	0

Table 3.9: Fraction (%) of tasks declaring shipment of empty package.

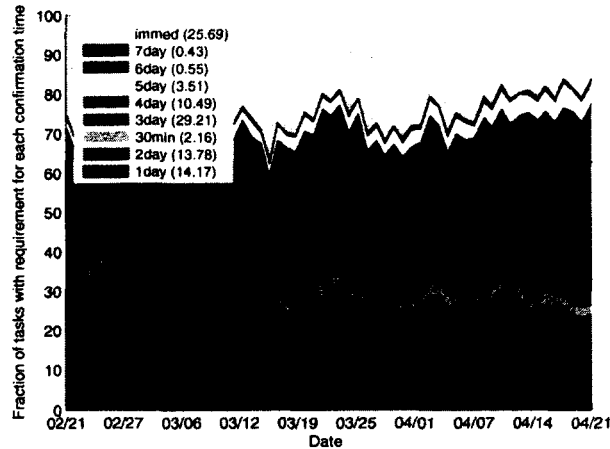


Figure 3.3: A breakdown of when to confirm the receipt. Numbers in parentheses in the legend denote the fraction of the total 219,165 tasks crawled on the five SRE markets.

Receipt Confirmation and Writing Positive Reviews. The last step to complete an online purchase is to confirm the receipt and write reviews. For workers on SRE markets, they must give the highest scores and leave positive reviews. The wait time for receipt confirmation is specified in each task and has nine possible values: immediately, 30 minutes, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, and 7 days. Figure 3.3 shows a breakdown of wait time of the total 219,165 tasks crawled on the five SRE markets. More than 50% of tasks require workers to confirm the receipt two to four days after placing the order with an attempt to match the typical shipping speed. Slightly more than a quarter of tasks require immediate confirmation of receipt, and a close scrutiny indicates that nearly all of them are tasks with virtual goods, which is reasonable due to no need for shipping.

Preparation for Appeal to Taobao against Penalization. Although elab-

Requirement	SKY	WOOD	EMPIRE	COOL	NET
Email & phone	1.16	2.42	0.26	1.56	0.02
Confirm on AliW.	11.65	0	0	0	1.97
Guarantee \$	19.10	5.13	1.26	23.96	4.21

Table 3.10: Fraction (%) of tasks requiring leaving email address and phone number, confirmation on AliWangWang, and guarantee money.

orately conducted, a fake purchase may still be detected, and the sellers involved may be penalized by Taobao. However, sellers have the right to appeal against penalties by presenting evidence of real transactions. Although an express tracking label usually serves as strong evidence, to collect more evidence, some sellers require task workers to leave their phone numbers or email addresses in the placed orders. Some workers are also required to confirm the receipt of goods on other channels like AliWangWang. Moreover, to prevent malicious workers from reporting fake transactions to Taobao after getting rewarded, some tasks require workers to have guarantee money held by SRE markets. Table 3.10 lists the fractions of tasks with each of these requirements. Although only a small portion of tasks have these requirements, they reflect the sophistication of SRE markets against Taobao detection.

3.3.3 Effectiveness of SRE Markets' Evasion Strategies

SRE market operators have developed sophisticated strategies to evade Taobao's fake transaction detection mechanism. Thus, it is quite interesting to check how those evasion strategies are effective against Taobao's hidden detection mechanism. Specifically we evaluate their effectiveness by examining what percentage of those 4,109 Taobao sellers who involved in fake transactions with their Taobao IDs being identified were ever penalized by Taobao in the two-month period, during which we

monitored these sellers' reputation growth on the Taobao marketplace². We first describe Taobao's punitive measures on merchant misconducts and then present our observation of the penalties suffered by SRE sellers for conducting fake transactions.

3.3.3.1 Taobao's Punitive Measures

Taobao takes different punitive measures based on the seriousness of misconduct. Its major penalties are described as follows: (I) remove the transaction volume, reputation score, and customer reviews generated by a fake transaction; (II) demote the involved stores by lowering their rankings in Taobao search results; (III) take the involved products off shelves and do not display them on the Taobao marketplace for specific days; (IV) ban the stores from future advertising and promotion campaigns; (V) deduct penalty points from the involved stores' reputation scores; (VI) zero a store's reputation score; and (VII) (permanently) shut down the involved stores. The penalties VI and VII represent the two most severe penalties. Note that the above penalties may be imposed individually or jointly.

3.3.3.2 Penalties Imposed on SRE Sellers

Although we were able to monitor the daily variation in store reputation for each of the 4,109 Taobao sellers, we cannot observe the effect of the penalties I, II, III, or IV imposed on a Taobao store from outside. Thus, our findings may inevitably underestimate Taobao's crackdown on fake transactions.

We focus on the three most severe and also observable penalties V, VI, and VII imposed on the SRE sellers by the Taobao marketplace. The penalty V is regarded

²Although the manual collection of the 4,109 sellers' Taobao IDs was completely done on March 21, 2014, we began monitoring the already identified sellers' reputation changes on the Taobao marketplace once daily as early as February 21, 2014 and finished monitoring on April 21, 2014.

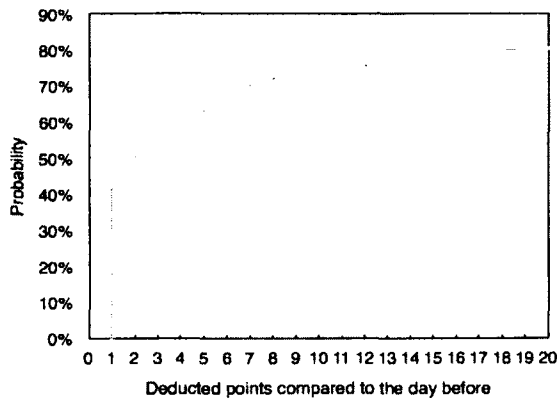


Figure 3.4: CDF of the deducted reputation points of the 932 Taobao sellers suffering from the penalty V.

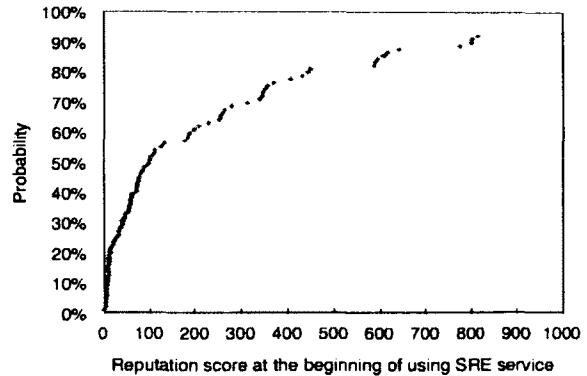


Figure 3.5: CDF of the reputation scores of the 89 heavily penalized sellers when beginning posting fake-transaction tasks on SREs.

to be imposed on a Taobao store if the store’s reputation score decreases compared with the day before. The argument is that a Taobao store’s reputation is supposed to monotonically increase with the time unless the penalty V is imposed. A Taobao store is considered to suffer from the penalty VI if the store’s reputation score is zeroed one day. A Taobao store is undergoing the penalty VII if our search for the store in the Taobao search engine continuously returns information showing that the store does not exist anymore before we ended our monitoring on April 21, 2014.

We examined each SRE store’s reputation changes and its presence to spot any signs of the penalties V, VI, and VII imposed by Taobao. Among the 4,109 identified Taobao sellers, 932 (22.7%) sellers were observed to suffer from the penalty V, i.e., with points deducted from their reputation scores. 89 (2.2%) sellers underwent the two heaviest penalties VI and VII. More specifically, 9 stores had their store reputations zeroed and 80 stores were forcibly shut down.

For the sellers with the penalty V imposed on, we study how much price they pay for fake transactions in terms of penalty points. Figure 3.4 depicts the cumulative distribution function (CDF) of the deducted points of the 932 Taobao sellers.

It shows that about 40% of these sellers had store reputation scores deducted by 1 compared to the day before conducting fake transactions; for about 40% of these sellers, their store reputation score deduction is between 1 and 20; and for the rest of 20% of these sellers, their reputation score deduction is more than 20. Taobao's publicly available penalty rules [41] show that Taobao makes the punishment decisions based on the total number of fake transactions and the frequency.

For those 89 severely penalized Taobao sellers, we further examined their characteristics in terms of their reputation scores at the time of using SRE services, their store ages when penalized, and the possibly unusual reputation growth before penalized.

It is interesting to know the reputation scores of these Taobao sellers when their stores were shut down or reputations were zeroed. Figure 3.5 shows the CDF of reputation scores of the 89 SRE sellers at the beginning of using SRE services. About 65% of those sellers had reputation scores of less than 251, i.e., a diamond grade, when they started fake transaction campaigns on SRE markets, and about 80% with reputation scores of less than 501, i.e., two diamonds grade. The results imply that most of those penalized Taobao stores had low reputation scores when they started to use the SRE services.

We also examined the shop ages of those sellers while being heavily penalized. As shown in Figure 3.6, about 70% of those sellers ran their Taobao stores for less than one year, about 50% of them ran Taobao stores for less than half a year, about 30% ran Taobao stores for only less than three months, and about 7% were penalized at the same month when they started their Taobao stores. The results indicate that the newly opened Taobao stores tend to use SRE services for escalating their reputations, which cause them to be closely monitored by Taobo and their conducted

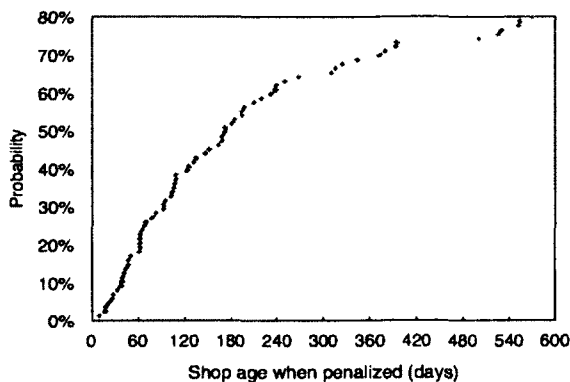


Figure 3.6: CDF of the shop ages of the 89 sellers while being heavily penalized.

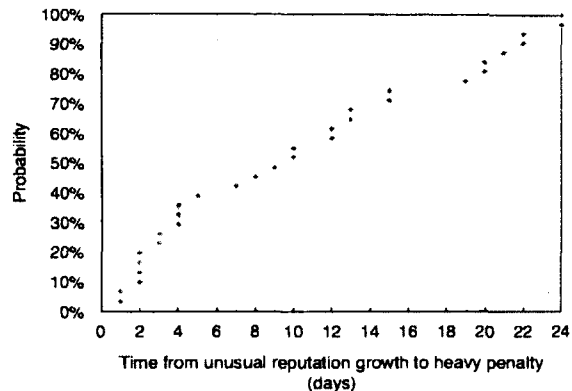


Figure 3.7: CDF of days taken from unusual reputation growth to heavy penalty.

fake transactions to be more easily detected by Taobao.

In addition, we wonder if those Taobao stores received any light penalties as advance warnings before being heavily penalized or if any unusually dramatic increase in reputation may trigger an alarm. For the 89 heavily penalized Taobao sellers, we examined the day-to-day variation in reputation scores. We found that 19 (21.3%) sellers underwent decrement of reputation scores at least once before being heavily penalized, which indicates that Taobao had noticed the fake transaction behaviors of those sellers and already imposed the relatively light penalty V as warning before further zeroing their reputation scores or shutting down their stores. One Taobao store was found to have its reputation score decreased for three times within 10 days. Moreover, 16 (18%) of the 89 sellers were found to have stunning growth in their reputation scores within one single day. The unusually rapid growth of reputation score within a short time period likely caused those stores to have store reputations zeroed or be shut down afterwards. For instance, one Taobao store had reputation increased by 2,957 within one day while its daily average reputation increase was usually close to zero, and then was shut down three days later. Next, we scrutinized

how quickly the above 35 sellers got severe penalties after receiving warnings or showing unusual reputation growth. Figure 3.7 shows that all these 35 sellers were severely penalized within one month, 70% sellers suffered a heavy penalty within two weeks, and about 40% within one week.

3.3.3.3 Summary

Our results demonstrate that 22.7% of the identified sellers received penalty points for conducting fake transactions and 2.2% of these sellers had their store reputations zeroed or had their stores forcibly shut down. That is, according to our observation, Taobao penalized about 25% of the 4,109 sellers involved in fake transactions with their Taobao IDs identified. Given that other punitive measures taken by Taobao are not observable from outside, our results could underestimate Taobao's counter-measures against fake transaction campaigns in SRE markets.

3.3.4 Seller Characteristics

Over the two months of measurement, we observed 11,130 Taobao sellers with at least one task posted on the five SRE markets. By manually undertaking tasks for one month from February 21, 2014 to March 20, 2014, we identified 4,109 unique Taobao seller IDs and subsequently crawled their profiles on Taobao. In this section, we first show what kind of Taobao sellers are more likely to use an SRE service based on the profiles of the 4,109 identified sellers. Then we feature how active those sellers are on SRE markets.

Store Start Date. The last day we manually undertook tasks was March 20, 2014. We chose this day as a reference date to calculate the age of those sellers' stores. Figure 3.8 shows the cumulative distribution function (CDF) of the store's

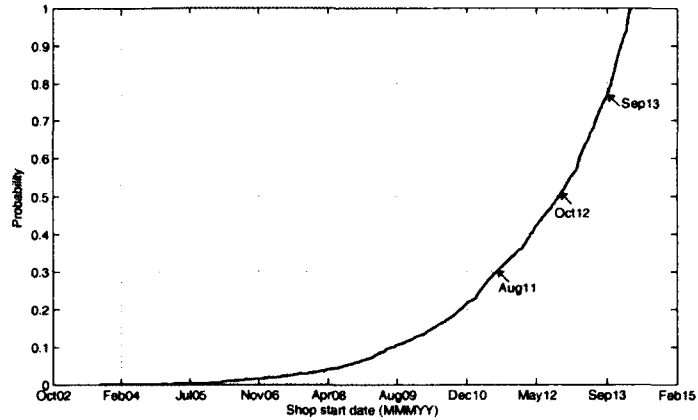


Figure 3.8: CDF of shop start date of the identified Taobao sellers.

start date for the 4,109 identified sellers. The plot shows that about 70% of those sellers opened their stores after August 2011, within 2.5 years; 50% of stores have an age of no more than 1.5 years; 25% of stores were opened for less than half a year. This indicates that SRE markets are more popular among new sellers since they usually have a stronger desire to improve their stores' reputation.

Main Business. Each Taobao store has one main business. The 4,109 identified stores fall into nine categories based on their main businesses. Figure 3.9 shows the top 5 main businesses, which accounts for 69.7% of the 4,109 stores. Specifically, nearly 40% of stores sell clothing and accessories, which conforms to the fact that apparel is the most popular buying category on Taobao marketplace. About 10% of stores sell game and phone cards.

Store Reputations. We studied the distribution of the store reputations of the 4,109 identified Taobao sellers when they were observed to use SRE services at the first time. In Table 3.11, reputation scores are divided into specific ranges corresponding to various grades specified by the Taobao marketplace. For instance, a Taobao seller with a reputation score between 251 and 500 has a grade of one diamond. A reputation score between 501 and 1,000 corresponds to a grade of two

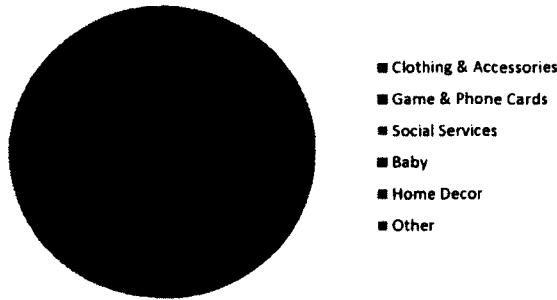


Figure 3.9: Top 5 main businesses run by the 4,109 Taobao sellers.

diamonds. More details about the Taobao reputation grades are described in Table 3.14 in Section 3.4. Table 3.11 shows that most of the 4,109 sellers have small reputation scores. About 50% of those sellers have reputation scores less than 251, and about 80% have reputation scores less than 1,001. The result is reasonable since it is pursuing a high reputation score that formulates the motivation to use SRE services for most Taobao sellers.

Reputation Range	0-250	251-500	501-1,000	1,001-2,000	2,001-
Percentage	51.1%	15.5%	12.9%	9.0%	11.6%

Table 3.11: Distribution of the reputation scores of the 4,109 identified Taobao sellers at the inception of using SRE services.

Active Duration. We crawled the 11,130 sellers’ profiles on SRE markets. However, only the COOL market provides the consistent and correct information, while the other SRE markets have shown strange variation in sellers’ profiles over time. For instance, the total number of tasks posted by a seller on those markets does not monotonically increase over time but fluctuates irregularly. Thus, we only consider the profile dataset crawled from the COOL market. A seller is considered to be active throughout the period from her first posted task to the last one. The length of this time period is counted as the seller’s active duration on the market.

Figure 3.10 shows the CDF of active durations for the 2,419 sellers on the COOL

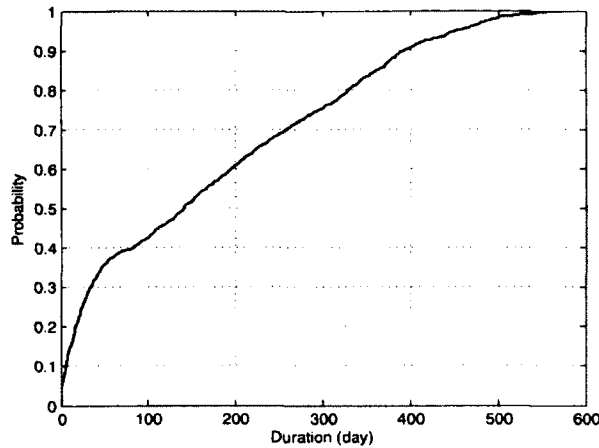


Figure 3.10: CDF of active duration of sellers on the COOL market.

market. About 57% of sellers stay active on the COOL market for more than 100 days, 40% for more than 200 days, and about 2% for more than 500 days. Based on the articles on the COOL market and Whois query results, we conjecture that the COOL market was founded on August 2012. Our results imply that most sellers on the COOL market may post tasks for several months or years, and a small portion have remained active since shortly after the market was formed.

Daily Tasks Posted per Seller. We also investigate how many tasks a seller posts daily. Figure 3.11 shows the number of tasks posted daily per active seller on the five SRE markets during the two months we crawled. It is obvious that the sellers on the SKY market are most active, with 4.2 tasks posted daily on average per seller, while the sellers on the NET market are least active, with an average of 2.2 tasks posted daily per seller. The parameter—the number of tasks posted daily per active seller—shows low variation over time for all markets except that it is quite low for COOL and NET markets at the end of February and at the beginning of March. The reason is that the authentication cookies of these two markets expire so quickly that we didn't realize the problem at first and thus missed many newly

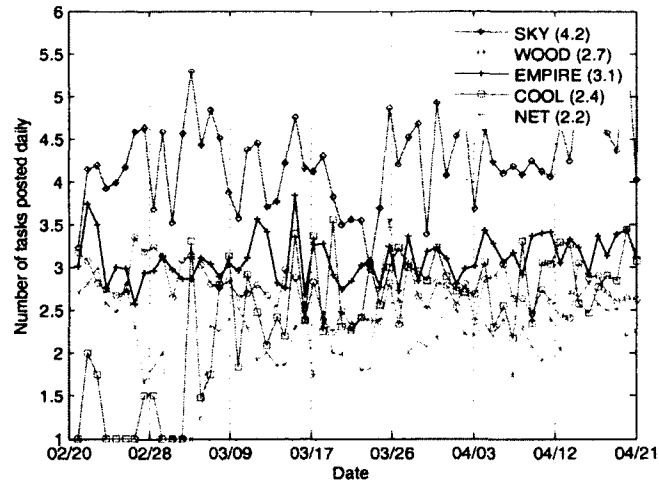


Figure 3.11: Tasks posted daily per active seller on SRE markets over time. Numbers in parentheses in the legend denote the mean values of the number of tasks posted daily per active seller on each SRE market during our crawl interval.

posted tasks.

3.3.5 Worker Characteristics

The only publicly accessible data about workers on each market is a list of the top 10 workers. In addition, only the COOL market provides consistently reasonable profile information. For instance, all the SRE markets but COOL set the first time of sellers to post tasks to be January 1, 2014. Thus, we use the 55 unique workers appearing in the top 10 worker list on the COOL market for analysis.

We examine their active durations, average tasks undertaken daily, and average daily earnings. We compute average daily earnings by multiplying average tasks undertaken daily by 80% of average commission fee per task (because 20% of commission fee is withheld by the market). Figure 3.12(a) plots the CDF of the active durations for the 55 workers on the COOL market. The active duration of a worker is the length of the time period during which the worker undertakes tasks. The

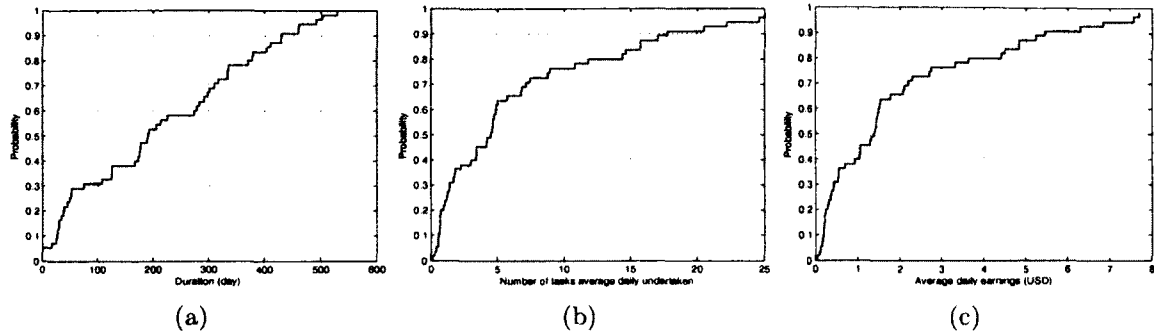


Figure 3.12: (a) CDF of active duration of the top workers on the COOL market. (b) CDF of average tasks undertaken daily by the top workers on the COOL market. (c) CDF of average daily earnings of the top workers on the COOL market.

CDF plot shows that about 70% of workers have been active for more than 100 days, about 30% of workers for more than 300 days, and about 2% for more than 500 days, which demonstrates that the top workers could remain active on the markets for several years. Figure 3.12(b) shows the CDF of average tasks undertaken daily by the top workers. About 40% of workers undertake more than 5 tasks daily, about 25% undertake more than 10 tasks daily, and about 10% undertake more than 20 tasks daily. It takes about 5 minutes to undertake a task as revealed by the chat contents on SRE markets. Thus, nearly all workers only spend less than 2 hours in taking tasks on SRE markets, which implies that most workers may take tasks only in their spare time. We do not have statistics about the demographics of workers, but close monitoring of the IM chat groups on each market reveals that most workers are college students, housewives, and freelancers. Figure 3.12(c) shows the CDF of average daily earnings of the 55 workers on the COOL market. More than one third can earn more than \$2 daily, and about 5% earn more than \$7 daily. The daily earnings seems quite low, but it is still attractive considering that the completion of a task only costs about 5 minutes, and the per capita daily income for a Chinese person is about \$16 according to the World Bank statistics [39].

3.3.6 Estimating Revenue and Fake-Transaction Volume

We estimate how much revenue these five SRE markets generated and how large of a transaction volume they handled during the two-month period we monitored. The revenue generated by each market consists of the withheld commission fee from task postings on that market and the earnings from express tracking label sales. For each task, SRE markets withhold 20% of the associated commission fee. And for each physical goods, the corresponding seller needs to purchase from the SRE market one express tracking label to complete the transaction. Each express tracking label is charged at a price of \$0.4-0.7 depending on the shipment company. SRE markets cooperate with shipment companies to provide tracking labels for sale. We do not know how they split the revenue from each sold tracking label and assume a 50/50 basis. Thus, the revenue generated by a market could be calculated based on the formula: $20\% \times \text{CPT} \times (\text{total task number}) + 50\% \times \text{label price} \times (\text{total physical tasks})$, where CPT denotes the average commission fee per task, ranging from \$0.28 to \$0.38 depending on the five markets. In addition, we calculate the fake-transaction volume handled by an SRE market by adding together the goods' value in each task we crawled over the course of two months. This metric reflects the total value of fake transactions conducted through SRE markets during our observation time period.

	SKY	WOOD	EMPIRE	COOL	NET
Revenue(\$)	11,805	12,369	9,938	9,189	3,137
Trans. Vol.(\$)	815,130	1,121,700	674,300	714,730	126,670
CPT(\$)	0.32	0.32	0.28	0.38	0.30
# tasks	63,343	54,824	48,120	39,823	13,055
# phys. tasks	38,753	44,301	36,216	30,811	11,768
Label price(\$)	0.4	0.4	0.4	0.4	0.4

Table 3.12: Estimated revenue and transaction volume

Based on the two-month crawled data, Table 3.12 lists our estimation of the

revenue generated and the fake transaction volume handled by each market during the two months, along with the parameters involved in the formula for calculating revenue. We estimate that WOOD generated a revenue of at least \$12,369 during the two months. SKY, EMPIRE, and COOL all generated more than \$9,000 revenue. The revenue generated by NET is slightly more than \$3,000. One main reason for the relatively low revenue of NET is that our crawler missed a large portion of task postings due to the quick cookie expiration. Based on the statistics for the COOL market, we estimate that its annual revenue will be more than \$74,000. Note that we did not catch all the task postings on SRE markets, and all five SRE markets simultaneously profit from a variety of other services like selling Taobao accounts and the TRUSTEE service (we will introduce this in Section 3.4). Therefore, the estimated results likely represent only a lower bound of their overall revenues.

We estimate that the fake transaction volume handled by each market is enormous. For instance, the COOL market handled at least \$1,121,700 during the two months, and it is estimated to handle the annual transaction volume of more than \$6,700,000. The operators of SRE markets accumulate such a large amount of wealth in a short time, and they may make off with money that sellers and workers deposit into the markets. Actually, at least two SRE markets have been reported to make off with millions of dollars, and the involved sellers and workers suffer heavy financial losses [43, 44].

3.4 Effectiveness of SRE services

In this section, we first evaluate the effectiveness of a typical SRE service requiring sellers to post tasks on SRE markets. Then, we present a more worrisome service newly launched by one SRE market and evaluate its effectiveness.

3.4.1 Effect of Posting Tasks on SRE Market

By manually undertaking tasks for one month, we identified 4,109 unique Taobao sellers, denoted as “EVIL” sellers. One interesting question is whether posting tasks on SRE markets could indeed improve sellers’ reputations on Taobao. Or, can a seller using SRE services increase her store reputation remarkably faster than a fellow Taobao seller who has the same store age, sells the same categories of goods, but does not use SRE services? To address this question, we randomly selected 4,000 legitimate Taobao sellers who follow the same distribution of store ages and main businesses as those 4,109 SRE sellers. We denote these random sellers as “BENI”, standing for benign sellers, which is arguably a fine assumption since the possibility that a randomly selected Taobao seller performs fake transactions on the SRE markets is extremely small considering the order of magnitude of active sellers on Taobao. We compare the two groups’ growth curves in their store reputations to evaluate the effectiveness of SRE services.

Set	1	2	3	4	5	6
Store age	<1m	<2m	<3m	<6m	<1y	<2y

Table 3.13: The store ages based on which we partition EVIL and BENI sellers. “<1m” denotes a store age of less than 1 month while “<1y” denotes less than 1 year.

We argue that only the comparison in reputation growth between the two stores with similar ages and selling the same category of goods makes sense. Thus, for the two groups of sellers, EVIL and BENI, we only consider those selling clothing and accessories (the most popular business run by the Taobao sellers identified to be using SRE services, see Figure 3.9). We further partition each group into six sets based on store age. Table 3.13 enumerates the store ages used for partition. We focus on the sellers with store ages not exceeding 2 years because they represent a

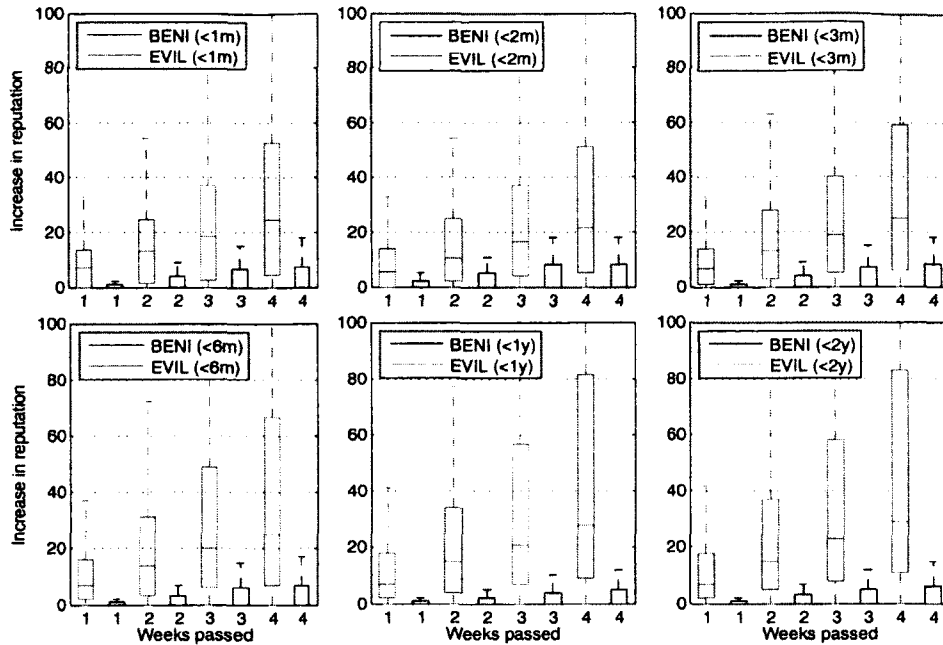


Figure 3.13: Comparison of the reputation growth distribution between BENI and EVIL stores with varying store ages over the course of one month.

majority of Taobao sellers on SRE markets (see Figure 3.8).

In Figure 3.13, we use boxplots to compare the distribution of reputation growth across one month between EVIL stores and BENI stores with varying store ages. Note that for each box, its bottom corresponds to the reputation increase of the Taobao seller on the 25th percentile, its top corresponds to that of the seller on the 75th percentile, and the line across the box corresponds to that of the seller in the median.

These boxplots clearly show that the reputations of EVIL stores increase much faster than those of BENI stores, regardless of the store age and time interval. Within one week, EVIL stores increase their reputation scores by a median value (represented by the median line of each magenta box) of 6 to 8, and by a median value of 22 to 30 within four weeks. The top 25% of EVIL stores (represented by the

whiskers on top of magenta boxes) increase their reputation scores by 14 to 18 at least (depending on store ages) within one week, and by 50 to 82 at least within one month. In contrast, the reputation scores of BENI stores with different store ages increase at a much slower rate. The median line of each blue box representing BENI stores overlaps with the x axis, implying that the median increase in reputation scores of BENI stores with different store ages is zero. It indicates that about 50% of BENI stores have not completed any transactions during the entire month. In addition, the top 25% of BENI stores (represented by the whiskers on top of the blue boxes) increase their reputation scores by 1 at least within one week and by 5 to 8 at least within one month, which is merely one-tenth of the reputation increase of the top 25% EVIL stores within one month. For all six kinds of store ages, the reputation increase of EVIL stores within one week (represented by the first magenta box in each subplot) is much larger than that of BENI stores within one month (represented by the last blue box in each subplot).

In summary, Taobao stores can remarkably increase their reputation scores by posting tasks on SRE markets, achieving higher ratings up to 10 times faster than legitimate stores. Actually, the effectiveness of SRE services is directly related to the Taobao's reputation computation method, which could be boiled down to one sentence that a Taobao seller earns one reputation score for each completed transaction with good ratings. And the three main inputs to the Taobao's reputation computation method—transaction volumes, product ratings, and customer reviews—are exactly what an unscrupulous seller gains from fake-purchase activities conducted through SRE markets.

3.4.2 An Emerging Service and Its Effectiveness

A new service was launched by the EMPIRE market on March 7, 2014. Different from the typical SRE service discussed above, this service does not require sellers to post tasks but demands full control of sellers' Taobao stores during the service time. The market operator does not disclose how this service is implemented but guarantees to increase sellers' reputation scores by up to 10,000 within several days. This service is quite attractive, since a legitimate seller may need several years to achieve the same reputation level, and even an insincere seller must tediously post about 10,000 tasks on SRE markets. We refer to this service as the *TRUSTEE service*.

Desired Grade	Reputation Δ	Fee (USD)	Days
1 diamond	251	96	3
2 diamonds	501	192	5
3 diamonds	1,001	384	7
4 diamonds	2,001	720	9
5 diamonds	5,001	1,200	11
1 crown	10,001	2,080	13

Table 3.14: TRUSTEE service expense standard: list of desired Taobao grade, corresponding reputation score increase, charged fees, and days needed to complete.

Table 3.14 presents the expense standard of the TRUSTEE service. The charged fees vary with desired Taobao grades³. According to the expense standard, a seller can obtain a diamond grade (i.e., increasing reputation by 251) within 3 days at the cost of \$96 while gaining a crown grade (i.e., increasing reputation by 10,001) requires \$2,080 and 13 days.

We crawled the EMPIRE market once daily for the list of customers who purchased this service from March 9, 2014 to April 21, 2014 and collected 108 Taobao sellers using this service. However, their Taobao IDs are not included in the crawled

³Taobao sellers have twenty grades going from one to five hearts, then one to five diamonds, then one to five crowns, and lastly one to five golden crowns. Taobao sellers need a specific number of transactions completed with positive reviews to progress to higher grades.

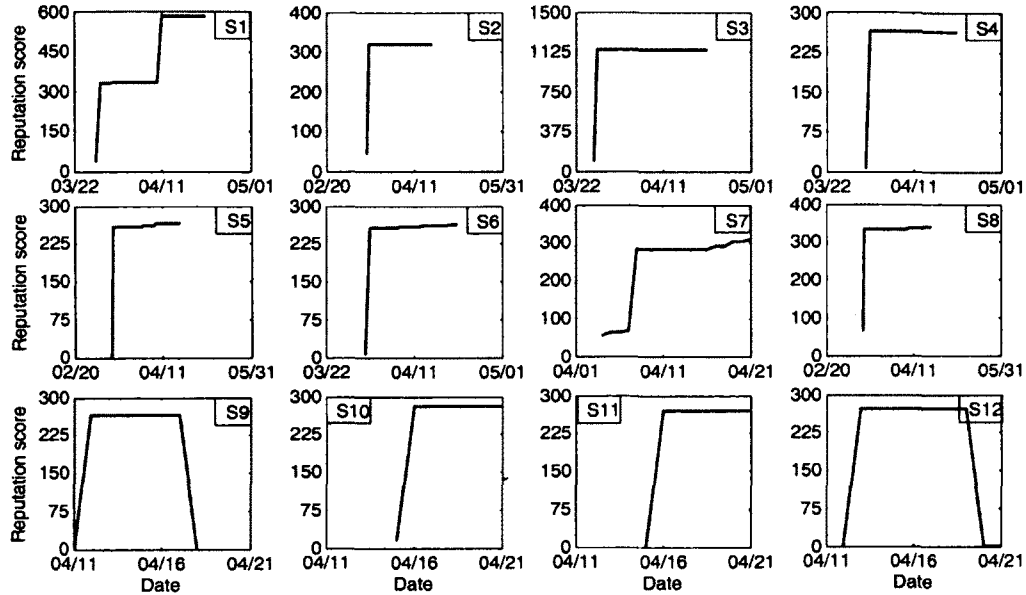


Figure 3.14: Reputation changes over time for the 12 Taobao sellers identified to use TRUSTEE service.

data and cannot be recognized by manually undertaking tasks, due to no task postings from those sellers. To reveal their Taobao IDs, we leverage one observation that some Taobao sellers use their Taobao IDs as their SRE account names. Thus, we crawled the Taobao marketplace and examined whether a Taobao store whose ID matches an existing SRE account. In this way, we successfully identified the Taobao IDs for 12 sellers. Subsequently, we performed daily crawling of these stores on Taobao to monitor their reputation changes.

Figure 3.14 shows the dynamics of the reputation scores of these 12 sellers, denoted as S1-S12, during and after their use of the TRUSTEE service. Each subplot depicts one seller’s reputation change. In each subplot, a steep increase corresponds to one use of the TRUSTEE service. We make several observations from this figure. First, the use of the TRUSTEE service can significantly increase sellers’ store reputation by a desired amount. We observed that each of these 12

sellers increased their reputation scores after receiving the TRUSTEE service by 251 or 1,001, corresponding to one diamond and three diamonds in Table 3.14, respectively. Second, some sellers may use the TRUSTEE service more than once. For instance, seller S1 used the TRUSTEE service twice within two weeks and requested an increase of 251 each time. Third, each request for the TRUSTEE service can be fulfilled within one day. Note that seller S3 requested an increase of 1,001 in reputation and was also satisfied within one day. Forth, it seems that the TRUSTEE service cannot guarantee a continuous increase in reputation nor an instant increase in sales. For each seller, the curve remains flat in the following 10 to 15 days after using the service. Fifth, only two sellers were observed to be penalized by Taobao. The reputation scores of two sellers S9 and S12 were reduced to zero in 7 to 10 days after their use of the TRUSTEE service. We conjecture that the two sellers were penalized by Taobao for their reputation manipulation. Lastly, four of those 12 Taobao stores using the service were newly opened within the past six months. Especially, two stores began using the service just a few days after their opening. It seems that this service is quite popular among new Taobao stores.

3.5 Potential Mitigation Strategies and Limitation

Although reputation manipulation is known for a long time in e-commerce, there is no open literature studying this specific SRE problem and we are the first to term the SRE markets and investigate them. Existing attacks against reputation systems include self-promoting, whitewashing, slandering, orchestrated attack, and DoS attack [45]. Compared to those known attacks, the newly emerging SRE problem is

much more sophisticated in three aspects: (1) much more organized (in the form of crowdsourcing), (2) much more severe (about tens of thousands of sellers and workers are involved during a short window of two months), and (3) much harder to detect (SRE market operators have formulated elaborate strategies for each step involved in a purchase transaction). In addition, our study shows that only 2.2% of the fraudulent sellers were detected, indicating that the Taobao's existing proprietary fake-transaction detection mechanisms fail in the face of the SRE problem.

Moreover, we believe that most existing detection mechanisms are vulnerable to SRE reputation escalation because the existing detection mechanisms do not consider the factor of shipping. In fact, whether the ordered products are delivered or not is the only difference between fake purchases conducted through SRE markets and real purchases.

We realize that the booming SRE business depends on four components: a highly available website to connect Taobao sellers and workers; express mail tracking labels sold to sellers; Taobao buyer accounts sold to task workers; and an escrow mechanism to resolve disputes between sellers and workers. Accordingly, defenders could develop a set of intervention approaches.

Domain Registrar and Web Hosting. If registrars were to suspend SRE markets' domains and web hosting service providers were to take down SRE market sites, the business of SRE markets would be interrupted immediately. We note that even temporary unavailability of SRE sites causes panic among involved sellers and workers greatly since they worry about their deposits on the SRE markets.

Shipping. Taobao could identify the shipment companies colluding with SRE markets and pressure them to terminate the cooperation. In addition, Taobao could collaborate with shipment companies to identify fake tracking numbers.

Taobao Accounts. To undertake tasks, a professional worker may need tens of Taobao accounts. Taobao accounts on SRE markets serve as the essential tools to conduct fake transactions, which highlights the need for Taobao account abuse detection at registration time.

Escrow Service Provided by SRE Markets. Escrow services used by SRE markets are all based on Alipay, the escrow system provided by Taobao. Actually, each SRE market operator has a publicly visible Alipay account to accept the money from a seller before posting tasks and to release money to task workers after the completion of tasks. Thus, suspending Alipay accounts used by SRE markets would dramatically demonetize the underlying enterprise.

In addition, by targeting SRE markets' evasion strategies revealed in this work, Taobao can further improve its current detection mechanism. A high risk of being penalized for fake transactions would cause Taobao sellers to abandon SRE services. In addition, the especially high popularity of SRE markets with new Taobao stores indicates that some measures should be taken by online marketplaces to help new stores to promote without hurting the fairness to established stores.

Limitations. Here we clarify the limitations of our study. First, we only investigated the SRE markets catering to online sellers from the Taobao marketplace in our study, although there indeed exist several other SRE markets targeting other Chinese marketplaces such as JD.com [40]. Taobao is much larger than the rest of the online marketplaces in China and even larger than Amazon and eBay. We believe that the SRE markets we studied are representative of the current SRE industry. In addition, our study sounds an alarm for other major online marketplaces, indicating that they may also suffer the same problem of fake transactions.

Second, there exist limitations in our dataset. Our crawler failed to catch all

task postings during the two months we monitored, due to unexpected network connection failures or expired authentication cookies. Thus, our evaluation results only represent a lower bound. Moreover, we cannot collect much more information about workers due to a limited amount of data available. Specifically, a list of the top 10 workers periodically updated is all the data related to workers and publicly available on an SRE market. Furthermore, some crawled data turned out to be unreasonable and was discarded. For instance, all the SRE markets but COOL set the first time of sellers to post tasks to be January 1, 2014. Only the COOL market provides consistently reasonable seller profile information. Thus, we only use the profile dataset crawled on the COOL market for analysis.

Finally, we did not implement a robust fake-purchase detection mechanism for evaluating our proposed intervention approaches, which requires the deployment cooperation from Taobao. We leave the defense evaluation for our future work.

3.6 Revisit the SRE Ecosystem One Year Later

Since the spring 2014, the problem of fake transactions in the e-commerce ecosystem has attracted wide attention from the public, industry, and research communities. Thus, we revisit the SRE markets infiltrated one year ago, including the involved Taobao sellers, to examine the possible changes in the SRE ecosystem.

Among the five SRE markets we infiltrated, EMPIRE is found to have been shut down by some law enforcement agencies. Another SRE market, NET, cannot be accessed by us due to the account/password loss. Thus, during a time period of 40 days from April 16, 2015 to May 26, 2015, we performed continuous crawling of the remaining three SRE markets — SKY, WOOD, and COOL — for newly posted tasks, and crawled the Taobao marketplace once daily for the reputation

scores of the 3,221 SRE sellers who posted tasks on the three SRE markets and had their Taobao IDs identified by us in 2014. For the convenience of illustration, we continue to use *EVIL* to denote those 3,221 SRE stores as we did for the 4,109 ones in Section 3.4. For the *BENI* stores mentioned in Section 3.4, we examined their current statuses (e.g., shut down or still open) and compared the reputation growth across one year between the *BENI* stores and the *EVIL* ones.

With the collected data set, we attempt to answer the following questions: (1) are the SRE markets still so dynamic as they were a year ago? (2) are the 11,130 SRE sellers observed to post tasks last year still posting tasks on the SRE markets? (3) how do the reputation scores of the 3,221 SRE stores grow in the past one year? and (4) is the reputation growth curve of the *EVIL* stores different from that of the *BENI* ones, and how?

3.6.1 Current Statuses of SRE Markets

By continuously crawling the three SRE markets for a time period of 40 days from April 16, 2015 to May 26, 2015, we evaluate the current dynamism of the SRE markets and attempt to spot any changes in their popularity across one year.

Table 3.15 summarizes the dataset we crawled on the three SRE markets. Specifically, we collected 31,997 fake-transaction tasks in total, contributed by 2,687 unique Taobao sellers. Each day, up to 794 new tasks were posted on an SRE market, and up to 297 sellers were observed to post tasks on a market.

We made a comparison between the dynamism of the three SRE markets in 2015 and that of the same SRE markets about a year ago. Table 3.16 details the comparison results, showing that the three SRE markets are not as active as they were a year ago. Specifically, compared to the results in 2014, we observed evident

Market	Total tasks	Total sellers	Daily new tasks (max)	Daily active sellers (max)
SKY	9,035	683	457	162
WOOD	11,449	904	669	262
COOL	11,513	1,100	794	297
Total	31,997	2,687	-	-

Table 3.15: Statistics of total task postings, total active sellers, maximum daily new tasks, and maximum daily active sellers on the three SRE markets during 40 days.

decrements in the three metrics — daily new tasks, daily active sellers, and daily tasks posted per seller. On the three SRE markets, the average number of daily new tasks falls by 42.1% to 68.3%; the average population of daily active sellers falls by 36.1% to 54.9%; and the average number of daily tasks per seller falls by 12.5% to 40.5%. We speculate that the widespread concern over the e-commerce reputation manipulation problem and much more effective countermeasures taken by Taobao since last year have contributed to the dramatic decline in activities on the SRE markets.

Market	Daily new tasks (avg.)			Daily active sellers (avg.)			Daily tasks per seller (avg.)		
	2015	2014	rate	2015	2014	rate	2015	2014	rate
SKY	301	951	↓68.3%	101	224	↓54.9%	2.5	4.2	↓40.5%
WOOD	382	816	↓53.2%	138	297	↓53.5%	2.3	2.7	↓14.8%
COOL	384	663	↓42.1%	149	233	↓36.1%	2.1	2.4	↓12.5%

Table 3.16: Variation of the dynamism of the three SRE markets in the past one year.

3.6.2 Current Activities of SRE Sellers on SRE Markets

During the two months from February 2014 to April 2014, we observed 8,473 SRE sellers ever posting tasks on the SKY, WOOD, and COOL markets. Note that there are 3,221 out of the 8,473 sellers, for which we were able to identify their Taobao IDs by manually undertaking tasks.

We revisited the three SRE markets to examine how many of the 8,473 sellers previously captured are still active on the SRE markets and then monitor their task posting activities. Table 3.17 lists the number of sellers still active on each SRE market, the percentage of them in the 8,473 sellers, and the percentage they represent in these sellers currently active in 2015. It shows that in total 559 SRE sellers observed last year were still posting tasks on the SRE markets during the 40 days we monitored this year. They represent 6.6% of the 8,473 sellers and 20.8% of the 2,687 SRE sellers currently active on the three SRE markets.

Market	# of 8,473 sellers still active	% of 8,473 sellers still active	occupy % in currently active sellers
SKY	89	3.2% (89/2,789)	13.0% (89/683)
WOOD	221	7.4% (221/2,968)	24.4% (221/904)
COOL	249	9.2% (249/2,716)	22.6% (249/1,100)
Total	559	6.6% (559/8,473)	20.8% (559/2,687)

Table 3.17: Current status of the 8,473 sellers on the SRE markets.

We then analyzed how active these 559 sellers were on the SRE markets in 2015 in terms of the number of task postings per active seller per day. Each seller was found to post 2.2 tasks each day on average and the median was 1.75. Then we compared these results in 2015 to those in 2014 for the same set of 559 active sellers. Figure 3.15 shows the comparison, indicating that the number of task postings per active seller per day in 2015 is similar to that in 2014, with just a slight decrease.

3.6.3 Current Statuses of the EVIL Taobao Stores on the Taobao Marketplace

Based on the information returned by Taobao upon our crawling for the profiles of the 3,221 *EVIL* stores, we disclosed those stores' current statuses. Specifically, a Taobao store is regarded to be inaccessible when Taobao continuously returns in-

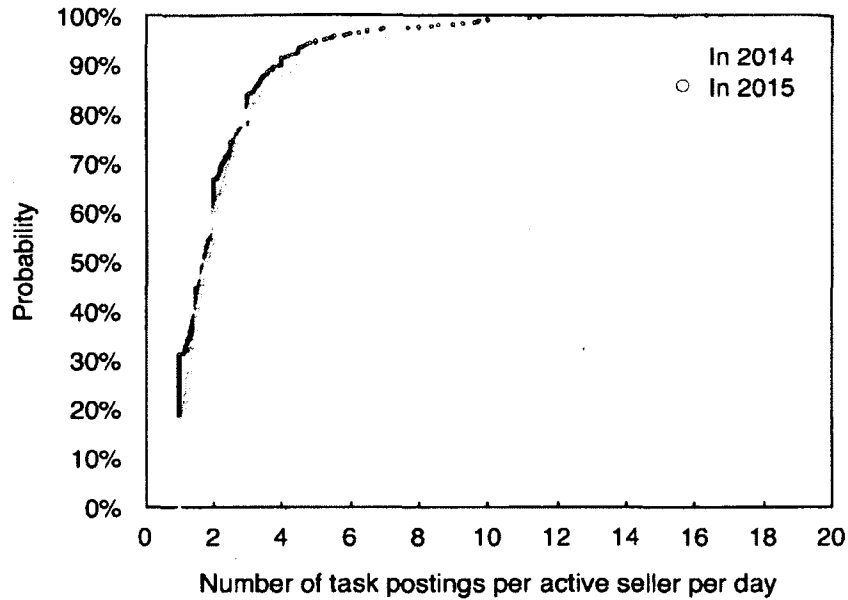


Figure 3.15: CDF of task postings per active seller per day in 2015 v.s. that in 2014

formation showing that the store does not exist during the 40 days we monitored; otherwise, a store is considered active on the Taobao marketplace if its profile information could be normally retrieved. In addition, for the purpose of comparison, we performed the same operations for the 4,000 *BENI* stores and unveiled their current statuses.

Table 3.18 lists the statistics about the current statuses of both *EVIL* and *BENI* Taobao stores. It shows that currently 83.3% of the *EVIL* stores are active while 16.7% are inaccessible, and for the *BENI* stores, 87.8% are active and 12.2% are inaccessible. The inaccessible ratio of the *EVIL* stores is slightly higher than that of the *BENI* ones. We also paid special attention to the 89 Taobao stores that suffered from heavy penalties (reputation zeroed or forcibly shut down) in 2014, and found that 47 of them are still inaccessible, which may indicate that more than a half of Taobao stores will not be reopened once reputation-zeroed or shut down.

3.6.3.1 Reputation Growth of the Active *EVIL* Stores Across One Year

Since we stopped crawling the SRE markets during the time period from April 22, 2014 to April 15, 2015, we cannot know whether the *EVIL* sellers continued posting tasks on the SRE markets during that period of time. However, it is still interesting to examine the difference between the 2,682 active *EVIL* stores and 3,512 active *BENI* stores in their reputation growth in the past one year.

Similar to what we did in Section 3.4, we only considered those selling clothing and accessories in the two groups of *EVIL* and *BENI* stores. A store’s current reputation score does affect its future reputation growth over a long time period, like one year. Thus, based on a store’s reputation score on April 21, 2014 and the reputation grades defined by Taobao in Table 3.14, we partitioned each group of stores into four sets — “ $x < 501$,” “ $501 \leq x < 1,001$,” “ $1,001 \leq x < 2,001$,” and “ $x \geq 2,001$.” We then compared the reputation growth curve between the two group of stores with similar reputation scores in 2014 and selling the same category of goods. Specifically, we compute the reputation increase as the difference between the reputation scores of 2015 and 2014, and then divide the reputation increase by the reputation score of 2014 to derive the reputation growth rate for each seller.

In Figures 3.16 and 3.17, we use boxplots to compare the distribution of reputation growth across one year, in terms of the reputation increase and the reputation growth rate, between *EVIL* stores and *BENI* stores with varying reputation grades.

Market	% active	% inaccessible
SKY	83.3% (1,109/1,332)	16.7% (223/1,332)
WOOD	83.4% (1,027/1,232)	16.6% (205/1,232)
COOL	83.1% (546/657)	16.9% (111/657)
Total	83.3% (2,682/3,221)	16.7% (539/3,221)
BENI	87.8% (3,512/4,000)	12.2% (488/4,000)

Table 3.18: Current statuses of *EVIL* and *BENI* stores.

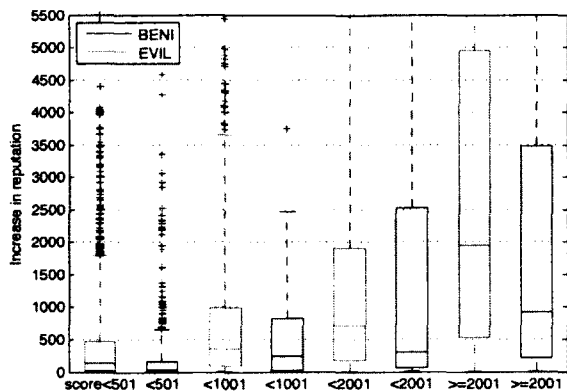


Figure 3.16: Comparison between currently active *EVIL* and *BENI* sellers in terms of reputation increase across one year.

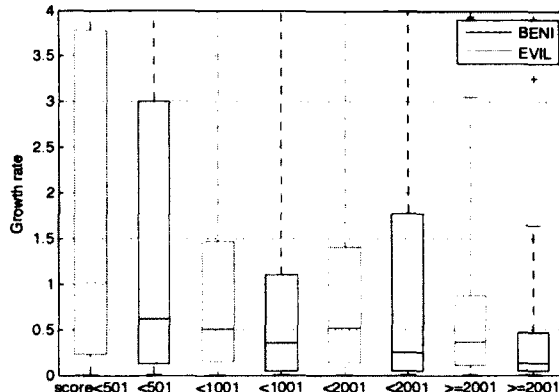


Figure 3.17: Comparison between currently active *EVIL* and *BENI* sellers in terms of reputation growth rate across one year.

The two figures clearly show that the reputations of *EVIL* stores increase faster than those of *BENI* stores in both metrics — reputation increase and reputation growth rate — for the stores with all varying reputation grades except the grade set of “ $1,001 \leq x < 2,001$.” Take the stores with the reputation grade set of “ < 501 ” as an example, within one year, *EVIL* stores increase their reputation scores by a median value of 129 and at a median rate of 100%, while *BENI* stores only increase store reputations by a median value of 31 and at a median rate of 61%.

3.6.3.2 Reasoning Why *EVIL* Stores Become Inaccessible Now

For the 539 currently inaccessible *EVIL* stores that were active at the end of our monitoring in April 2014, it is hard to tell whether each store was forcibly shut down by Taobao for conducting fake transactions or it just died a “natural death.” This is because we have not continuously monitored their reputation changes in the past one year. However, we attempt to uncover the hidden reasons why those stores become inaccessible by checking their profile information on Taobao across

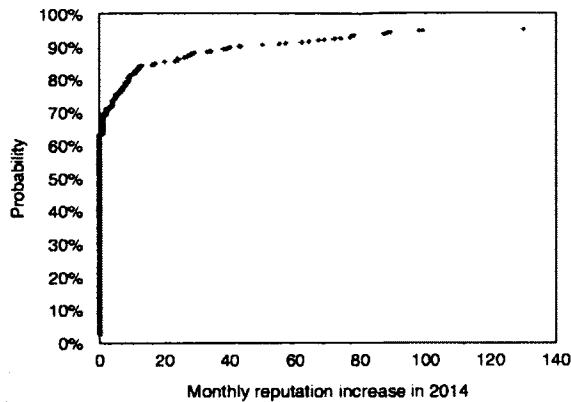


Figure 3.18: CDF of monthly reputation increase of the currently inaccessible *BENI* stores between 3/21/2014 and 4/21/2014.

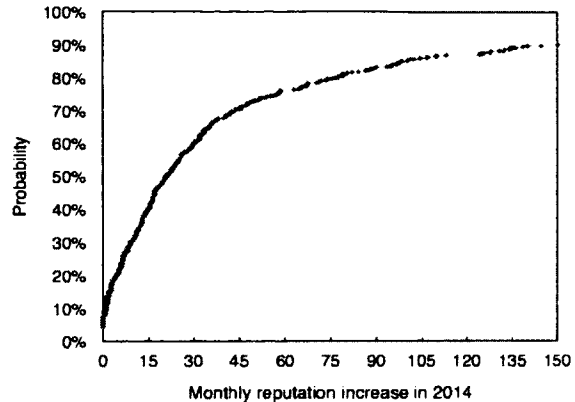


Figure 3.19: CDF of monthly reputation increase of the currently inaccessible *EVIL* stores between 2/21/2014 and 4/21/2014.

the two months we monitored last year from February 2014 to April 2014. To make insightful findings, we took the 488 currently inaccessible *BENI* stores as a control group with the assumption that each *BENI* store underwent a “natural death.” The assumption is reasonable since there are millions of stores on Taobao and it is quite normal for a small proportion of stores to die naturally each year.

Specifically, for each of the 539 inaccessible *EVIL* stores, we examined its monthly reputation increase, monthly reputation growth rate, and monthly completed transaction volumes from February 2014 to April 2014. We also did the same examination on the 488 currently inaccessible *BENI* stores. We then compared the distributions of those three metrics between *EVIL* stores and *BENI* stores to spot the possible reasons for those *EVIL* stores to become inaccessible now.

Figure 3.18 depicts the monthly reputation increase of the currently inaccessible *BENI* stores during one month we monitored in 2014. It shows that more than 60% *BENI* stores did not increase their reputations at all within a whole month; and only 20% stores increased their reputations by at least 15. According to our

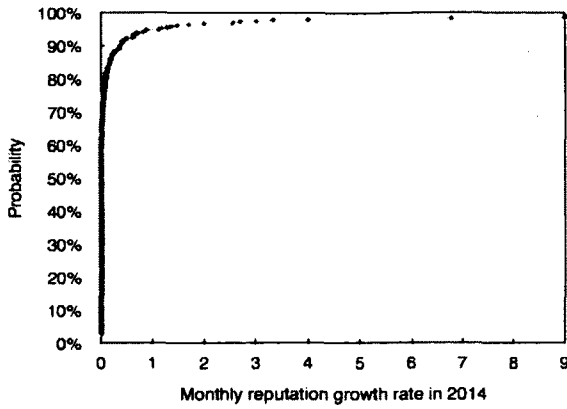


Figure 3.20: CDF of monthly reputation growth rate of the currently inaccessible *BENI* stores between 3/21/2014 and 4/21/2014.

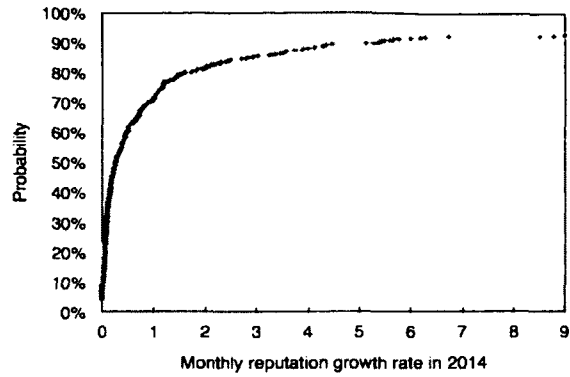


Figure 3.21: CDF of monthly reputation growth rate of the currently inaccessible *EVIL* stores between 2/21/2014 and 4/21/2014.

assumption, the distribution depicted in Figure 3.18 could be regarded to represent the distribution of monthly reputation increase for typical Taobao stores which are to be “dying naturally.” More specifically, the figure indicates that most dying Taobao stores have nearly stopped or at least have difficulty in gaining business before a “natural death.” Figure 3.19 depicts the monthly reputation increase of the currently inaccessible *EVIL* stores during the two months last year. It shows that about 95% of *EVIL* stores achieved increase in their reputations within a month; 60% increased their reputations by at least 15; and 30% by at least 45. These results demonstrate that those *EVIL* stores showed great ability in gaining business in 2014, quite different from the performance that those *BENI* stores presented. This observation indicates that the currently inaccessible *EVIL* stores were probably forcibly shut down by Taobao for fake transactions.

We also examined the monthly growth rate of reputation score in 2014 for both the currently inaccessible *EVIL* stores and *BENI* stores. In Figure 3.20, it shows that over 70% *BENI* stores had a monthly reputation growth rate of zero. In contrast,

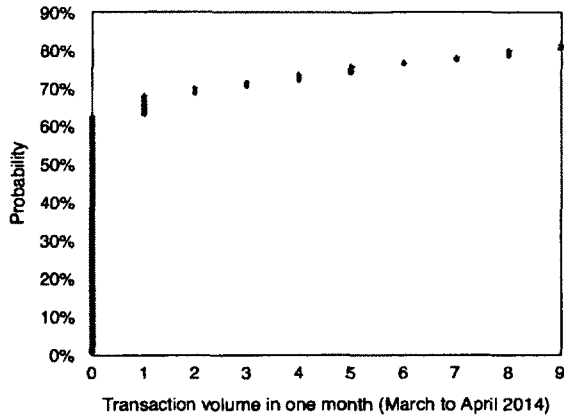


Figure 3.22: CDF of the monthly transaction volumes completed by the currently inaccessible *BENI* sellers in 2014.

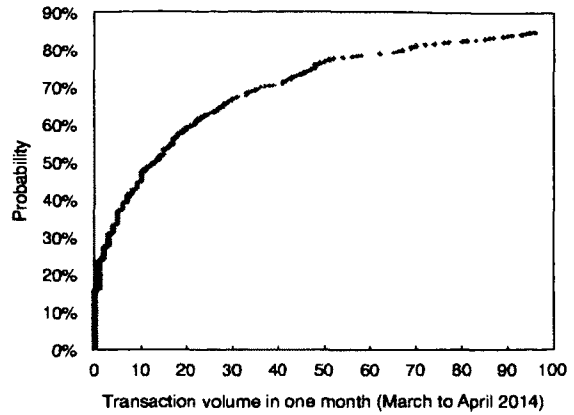


Figure 3.23: CDF of the monthly transaction volumes completed by the currently inaccessible *EVIL* sellers in 2014.

Figure 3.21 shows that only less than 20% of the currently inaccessible *EVIL* stores increased their reputations at a rate of zero, and more than 30% of those *EVIL* stores at least doubled their reputations within a month (i.e., at a growth rate of 1). Thus, it is unlikely that those *EVIL* stores died naturally when they were able to increase their reputations at such a large rate.

Lastly, we examined the monthly transaction volumes completed by both the currently inaccessible *EVIL* stores and *BENI* stores in 2014. In Figure 3.22, it shows that more than 60% of the currently inaccessible *BENI* stores did not complete any transaction volume within a month and about 20% completed the transaction volumes of between 1 and 9. On the contrary, Figure 3.23 shows that only about 15% of the currently inaccessible *EVIL* stores did not complete any transaction volume, and about 55% completed transaction volumes of more than 10. Again, those *EVIL* sellers seemed quite active in doing business on Taobao and unlikely closed their stores voluntarily in the following one year.

3.6.3.3 Summary

The above analysis of the currently inaccessible *EVIL* stores, in terms of their reputation changes and the completed transaction volumes during the two months we monitored last year, implies that it is probably the heavy penalties imposed by Taobao for fake transactions that causes those *EVIL* stores to become inaccessible currently, rather than that they shut down the stores at their own choices.

3.6.4 Summary

By revisiting the SRE markets and the involved sellers one year later since our previous study, we first evaluated the current dynamism of the SRE markets and found that the SRE markets are not so active as they were one year ago. Then we examined the current activities of the involved sellers on the SRE markets and observed that only 6.6% of those sellers are still posting fake-transaction tasks. Finally, we scrutinized the current statuses of these ID-identified SRE stores on the Taobao marketplace and found that about 83% are active and 17% become inaccessible. With the randomly selected Taobao stores as a control group, we found that those active SRE stores increased their reputations faster than the random Taobao stores in the past one year. For the currently inaccessible SRE stores, we speculate that they probably suffered heavy penalties imposed by Taobao for their fake transactions during the past one year, which caused those stores to close down.

3.7 Related Work

Over the past few years, many researchers have focused their studies on underground markets. Several works studied the underground economy related to Twitter, in-

cluding markets for selling fraudulent accounts [70, 71] and Twitter followers [65]. McCoy et al. examined the role of payment processing in the underground economy [57]. Motoyama et al. studied the social dynamics of underground forums [60] and investigated the market for CAPTCHA-solving services [59]. Franklin et al. [48] measured the commoditization of fraudulent activities on an underground market. Christin performed a similar measurement on Silk Road, an anonymous online marketplace [47]. Caballero et al. [46] and Grier et al. [49] studied the pay-per-install market and exploit-as-a-service model for malware distribution. In [52, 58], the authors studied the markets for online pharmaceutical sales. Park et al. [64] leveraged magnetic honeypot ads to study Nigerian scams on Craigslist.

Our work is also related to previous studies on crowdsourcing marketplaces, which enable people (known as *requesters*) to coordinate the use of human intelligence to perform tasks that computers are currently unable to do [42]. Since workers on crowdsourcing marketplaces can intentionally deliver low-quality work, Ipeirotis et al. presented an algorithm to estimate the quality of workers [51]. In addition, requesters may maliciously deny payment to workers. To address this problem, Ho et al. proposed the social-norms-based incentive mechanisms to augment crowdsourcing systems [50]. Crowdsourcing markets can also be employed for web service abuse. Miscreants could easily recruit a large group of workers to solve CAPTCHAs [59], register a multitude of fraudulent OSN accounts, and send email spam [61], etc. Similarly, SRE markets also operate in the crowdsourcing model. However, SRE markets target special requesters, i.e., online sellers, and provide special services, i.e., fake purchases on online shopping marketplaces.

Reputation systems are quite important to the e-commerce ecosystem. Several work [63, 54, 55, 56] investigated the online review manipulation and some others

[62, 53, 66] proposed methods to detect deceptive opinion spam, i.e. fake reviews written to deliberately mislead readers. Ott et al. explored the prevalence of deception in several popular online review communities [63]. Mayzlin et al. provided an empirical investigation of online review manipulation on two travel websites [54]. Chen et al. [69] conducted a real case study of opinion spams in a web forum. Nosko et al. discussed the limits of reputation mechanisms used in e-commerce markets [55]. Mukherjee et al. aimed to spot fake reviewers using behavior footprints [62]. Li et al. proposed a three-layer graph model to identify manipulated offerings on review portals [53]. Swamynathan et al. [66] proposed a reliable reputation system to address the attacks targeting reputation systems. Akoglu et al. proposed approaches that utilize network effect [67] or consider both the metadata and network factors [68] for spotting fraudsters and fake reviews. In this study, we examined a newly emerging underground industry in which a potentially unbounded number of inexpensive human laborers are hired to conduct fake purchases for reputation inflation. This way of tainting the reputation system is more advanced and beyond the attacks known previously.

3.8 Conclusion

We have conducted the first systematic study of a seller-reputation-escalation (SRE) ecosystem by infiltrating five SRE markets. These markets specialize in accommodating online marketplace sellers to post fake-purchase tasks for escalating their business reputations. We performed daily crawls for two months and observed that more than 11,000 online sellers posted nearly 220,000 tasks on the five SRE markets. Each new task could be undertaken within seconds. SRE markets turn out to be quite popular with online sellers. In addition, we examined the tactics for-

mulated by SRE markets for evading the online marketplace defenders' detection mechanism of fake transactions. Those tactics are so sophisticated that only about 25% of illegitimate online sellers were visibly penalized for fake transaction. Moreover, we characterized the online sellers involved in fake transactions and discovered that most of them run new stores and mainly sell clothing or game cards on the online marketplace. Furthermore, we evaluated the effectiveness of SRE services and revealed that the illegitimate sellers using SRE services can increase their reputations 10 times faster than legitimate ones. In addition, we investigated a newly launched SRE service and found that the service can increase sellers' reputations by up to thousands within one day. We estimated that an SRE market can generate annual revenue of over \$70,000 and handle annual fake-transaction volume of over \$6,700,000. We also discussed possible intervention approaches and proposed that the joint interventions at the domain, web hosting, shipping, account registration, and payment tiers are probably the most viable defense strategy. Finally, we presented the findings from our revisiting of the SRE ecosystem one year later. We found that the SRE markets are not so active as they were and that about 17% of the 4,109 identified Taobao stores are inaccessible probably due to the heavy penalties imposed by Taobao for fake transactions.

Chapter 4

Assessing Privacy Risks on Online Photos

Online photo privacy, which we are trying to address, has become a great concern nowadays. In this work, we first conduct a large amount of measurement to examine the prevalence of metadata and study the photo handling policies adopted by hundreds of top media sites. Then we demonstrate an important attack vector not exploited before and further propose an efficient re-identification attack.

To obtain a representative dataset for our study, we collected nearly 200,000 photos in total in various ways including soliciting freshly taken photos through crowdsourcing, downloading original sized, intact photos from a major photo sharing site, and crawling “wild” photos from Google Images and over 600 top ranked websites. We examined the metadata information embedded in these photos and found that metadata was prevalent among photos at each of the three stages. We paid special attention to the metadata fields that may give rise to great privacy concerns. We found that about 10% of “fresh” photos were tagged with GPS coordinates while 27%-37% of “intact” photos and only about 1% of “wild” photos contained GPS

information. We also measured the percentages of photos containing other sensitive metadata information including a photographer’s name and modification history.

To understand how a photo is processed after being shared online, we also investigated online sites’ policies on handling photos based on 97,664 photos crawled from 679 unique top sites in seven categories—“social networking,” “news,” “we-blog,” “college,” “government,” “shopping,” and “classified”¹ sites. We found that photo handling policies adopted by online sites vary with different categories. The “college” and “government” sites hardly resize the photos they host or remove the embedded metadata information. However, the sites in the other categories are more likely to resize the photos and remove the metadata information. Finally, we proposed that the metadata field *camera serial number* could be used as an attack vector. For 62.6% of unique photographers, we were able to uncover their both online and real-world identities with just one photo they ever took and posted online.

The remainder of the chapter is organized as follows. We provide background knowledge in Section 4.1. We describe data collection methods for “fresh” photos and characterize them in Section 4.2. We examine “intact” photos in Section 4.3. We characterize “wild” photos and investigate online sites’ photo handling policies in Section 4.4. We demonstrate the re-identification attack in Section 4.5. We discuss the limitation of this work and propose our future work in Section 4.6. We survey the related work in Section 4.7 and conclude the chapter in Section 4.8.

¹“Classified” refers to the classified advertisements sites such as Craigslist.

4.1 Background

In this section, we first give an overview of the metadata information typically contained in a digital photo, then discuss the potential privacy concerns, and finally illustrate the three stages we define for digital photos.

4.1.1 Metadata Information in a Photo

There are three most commonly used metadata standards for photos: EXIF, XMP, and IPTC. They often coexist in a photo and constitute the main part of the photo metadata. Table 4.1 lists the metadata fields typically included in a photo grouped by category.

Category	Information	Fields
When	Date Time	create time, modify time
Where	Location	GPS coordinates, city/state/country
How	Device Info.	camera make, model, serial number, light source, exposure mode, flash, aperture settings, ISO setting, shutter speed, focal length, color information
Who	People	artist's name
What	Description	title, headline, caption, by-line, keywords, copyright, special instructions
Modification	Modification History	create tool, xmp toolkit, history action, history when, history software agent, history parameters

Table 4.1: List of metadata information typically included in a digital photo.

A digital photo typically contains ample metadata information. When a shot is taken, the camera automatically embeds into the photo all the information it knows about the camera itself and the photo. In addition, users can add their own descriptive information with image processing software. Specifically, typical metadata information can be summarized as follows: (1) *when* – when the photo is created and modified if applicable, (2) *where* – the exact location (GPS coordinates and altitude) at which the photo is captured if a GPS receiver is equipped and enabled, or coarse-grained location information such as city/state/country, (3) *how* – the

camera device used, its make, model, serial number, light circumstances (sunny or cloudy, flash on or off), exposure (auto or manual), and all other parameters used, (4) *who* – the photographer and the people depicted in the photo if manually added during post processing, (5) *what* – title, headline, caption, keywords, copyright restriction, and other detailed descriptions added for logging, organization or copyright protection, and (6) *modification* – if the photo is modified, on what date and time, by what software on what computer, and the specific actions done to the photo.

4.1.2 Potential Privacy Concerns Arising from Photo Metadata

Most metadata fields may look innocent and trivial. However, some could raise serious privacy concerns. We highlight several sensitive metadata fields below.

Geolocation. Contemporary cameras and smartphones are typically equipped with GPS functions. When taking photos with these GPS-enabled devices, geolocation information is automatically saved into the metadata. For a photo posted online, anybody able to access it could check the metadata information and may get the geolocation where the photo was taken. This definitely violates the privacy of the photographer and the people depicted. For instance, the time and location embedded in an online photo indicated that a public figure had been at an embarrassing location and not where he claimed to have been [76]. Moreover, a geo-located photo obviously taken at home and depicting high-value goods may give burglars incentives. In addition, young parents usually like to post many photos of their kids online, which may raise great concerns because the photos tagged with GPS coordinates could disclose the exact locations of where their kids live, play, or study.

Photographer's/Owner's information. Some photos explicitly contain in the metadata the photographers' information, among which the name information is most commonly seen. No matter whether such information is embedded with or without the photographers' awareness, disclosing such information may cause identity leakage, especially given the availability of geolocation information in the metadata.

Modification History. When post processing a digital photo, an image processing software like Adobe Photoshop and Apple iPhoto often automatically embeds into the photo the detailed modification information, represented by three metadata fields: History When, History Software, and History Parameters. Table 4.2 presents an example of the embedded modification information in a photo. For the convenience of illustration, we add the photo's shot time in the table. It clearly shows that the photo has been processed twice in less than one month since it was taken on July 16, 2014. And two versions of Adobe Photoshop on one or two Macintosh computers were ever used for format conversion and save actions.

Create Date	History When	History Software	History Parameters
2014:07:16 15:13:56	2014:07:19 01:30:03, 2014:08:08 21:17:25	Adobe Photoshop Lightroom 5.4 (Macintosh), Adobe Photoshop Lightroom 5.6 (Macintosh)	converted from image/x-nikon-nef to image/dng, saved to new location, converted from image/dng to image/jpeg, saved to new location

Table 4.2: An example of modification information contained in a photo's metadata.

A photographer may not want to disclose such modification information, especially when such information may undermine what the photographer tries to convey through the photo. For instance, the contained modification information may cast doubt on the legitimacy of a photo used as digital photographic evidence in court. In addition, celebrities may not like the public to know the photos they were depicted

in are actually photoshopped.

4.1.3 Three Stages of Digital Photos

Based on their propagation process, contemporary digital photos fall into three stages: “fresh,” “intact,” and “wild.” In the “fresh” stage, a photo is freshly taken, free from any post-processing manipulations and still stored in the local camera device. All the metadata information contained in a “fresh” photo is automatically embedded by the camera device, instead of being subsequently introduced by a post processing. In the “intact” stage, a photo has been uploaded online, but remains intact and has not yet been compressed or resized by the hosting media site. For a photo in the “wild” stage, it may have undergone resizing, cropping, and other editing actions conducted by the hosting site, which could change the hidden metadata too. By characterizing digital photos in these three different stages, we aim to depict the status of contemporary digital photos.

4.2 Fresh Photos

The photos in the “fresh” stage are just freshly created. We examine the metadata information, especially sensitive information, embedded in those freshly taken photos. In this section, we first describe the method used for collecting “fresh” photos and then characterize the collected photos.

4.2.1 Data Collection

The collection of “fresh” photos is not easy due to their inherent characteristics. We found that it is an effective way to solicit “fresh” photos through crowdsourcing.

We posted tasks on a crowdsourcing platform. In each task, the required actions for a worker to take are two-fold: (1) pick up her smartphone, take a photo, and then send the photo to us directly via the instrumented email client application, and (2) take a short survey asking for her demographics information. In addition, to guarantee the unique origin of each photo, each worker is allowed to take our task only once.

For each received photo, we employed various methods to check if it is freshly taken with a smartphone rather than a photo randomly grabbed from the Internet. In addition, according to our tests, sending a photo via email does not affect its embedded metadata. Thus, our task requirements guarantee that the collected photos are freshly created and intact from any post-processing manipulation. The data collection lasts for two months and we collected 782 photos in total. We filtered out 170 photos that are either post-processed or created by other tools. We use the set of the remaining 612 photos for our study.

4.2.2 Characterizing “Fresh” Photos

Demographics. The 612 photos were collected from 612 unique workers from 76 countries. Table 4.3 lists the demographic statistics of the worker participants: (1) 71.7% of workers were male and the rest were female, (2) 45.5% of workers were from the top five countries, including India, United States, Serbia, Nepal, and Macedonia, (3) 82.1% of workers were between the ages of 18-34 and 10.8% between 35-44, (4) 47% of workers received the bachelor’s degree, 33.3% with high school degree, and 17.7% with graduate degree, and (5) 72.8% of photos were taken with Android phones and 18.2% with iOS phones.

(Sensitive) Metadata Prevalence. Although Table 4.1 lists quite a few meta-

Gender	Percent	Country	Percent	Age	Percent	Education	Percent	MobileOS	Percent
Male	71.7%	India	14.4%	<=17	2.3%	Graduate	17.7%	Android	72.8%
Female	28.3%	USA	13.7%	18–24	45.8%	Bachelor	47.0%	iOS	18.2%
NA	NA	Serbia	7.8%	25–34	36.3%	High Sch.	33.3%	WindowsP	5.2%
NA	NA	Nepal	5.3%	35–44	10.8%	Middle Sch.	1.7%	Blackberry	1.8%
NA	NA	Macedonia	4.4%	>=45	4.7%	Elementary	0.4%	Other	2.0%

Table 4.3: Demographic statistics of worker participants

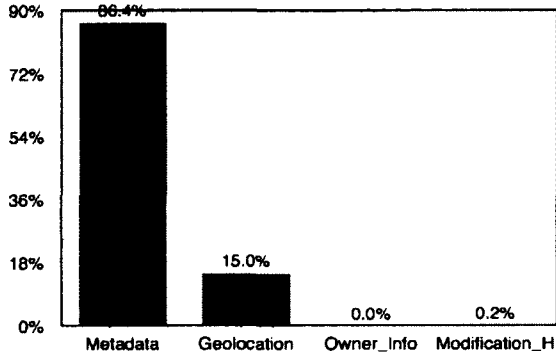


Figure 4.1: Percentage of “fresh” photos containing metadata information.

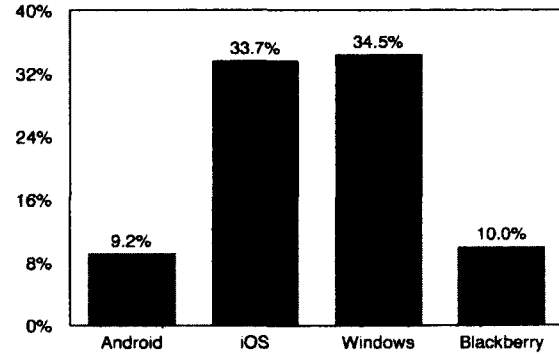


Figure 4.2: Percentage of “fresh” photos tagged with GPS for smartphone OS.

data fields typically embedded in a photo, a specific photo often has a large portion of its metadata information missing. According to our measurement results, we found that two metadata fields, camera make and model, are the most fundamental metadata information. That is, if they are missing in a photo, most other metadata fields are missing too. Thus, we decide whether a photo contains metadata information based on these two fields. A photo is regarded as containing metadata if either of the two fields has a non-empty value.

With the help of a third-party library [73], we examined the prevalence of metadata information among 612 “fresh” photos. We also examined if “fresh” photos contain any sensitive metadata fields, including geolocation, owner’s information, and modification history, as mentioned in Section 4.1. Figure 4.1 shows the percentages of photos containing metadata and sensitive metadata fields. As high as 86.4% of “fresh” photos contain metadata, which demonstrates the prevalence of meta-

data information among freshly taken digital photos. As of the sensitive metadata fields, 15% of fresh photos are tagged with geolocation information. The results show that although nearly all smartphones are now GPS-equipped, only some of them are GPS-enabled. The percentage is expected to be even lower if more people are aware that smartphones may automatically embed geolocation into photos and then choose to turn the GPS functionality off. None or hardly any of “fresh” photos contain photographers’ information or modification history in their metadata. We speculate that it is due to (1) our strict task requirements and (2) the possibility that these two kinds of sensitive metadata fields may not be automatically embedded at the time of a photo shot.

Impact of Smartphone OS on Geolocation metadata. It is interesting to examine which kind of smartphone OSes are more likely to automatically embed the sensitive geolocation information into photos. Figure 4.2 shows that about one third of iOS and Windows phones automatically embed geolocation into photos while only about 10% of Android and Blackberry phones do this.

4.3 Intact Photos

In the “intact” stage, photos have been posted online while retaining intact metadata information. From this perspective, “intact photos” could reflect the status of metadata in digital photos at the time of being shared online. In this section, we describe our data collection method for “intact” photos and examine the embedded metadata information in them.

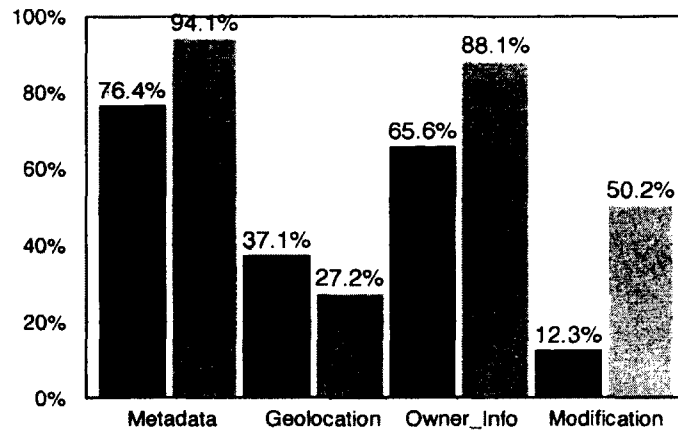


Figure 4.3: Percentage of “intact” photos containing metadata information. In each of four pairs of columns, the left black column represents *Flickr_p* while the right gray *Flickr_6*.

4.3.1 Data Collection

To collect such photos, we crawled photos from Flickr, a large photo-sharing website, using its API with the download option of “original size,” which guarantees that the photos remain original and intact from the site. More specifically, we collected two sets of “intact” photos from Flickr. The first set denoted by *Flickr_p* contains 18,404 photos exclusively taken with smartphones. Those photos were crawled from the Flickr group “Smartphone Photography” where all photos were taken with smartphones. The other set denoted by *Flickr_6* contains 43,704 photos uploaded within six months from July 1, 2014 to December 31, 2014. Our further examination shows that 94.3% of the photos in *Flickr_6* were taken with digital cameras.

4.3.2 Metadata Information Embedded

Similarly, we examined the percentage of “intact” photos containing metadata information, especially sensitive metadata fields. Figure 4.3 shows the percentages of “intact” photos containing metadata and sensitive metadata fields.

It shows that intact photos in *Flickr_p* and *Flickr_6* have quite high percentages containing metadata information, 76.4% and 94.1%, respectively. The results indicate that most digital photos taken with either digital cameras or smartphones contain metadata when being uploaded online. In addition, 37.1% *Flickr_p* and 27.2% *Flickr_6* photos contain GPS information. Considering 15% of “fresh” photos tagged with geolocation, we speculate that some photo owners may embed GPS information into photos during post processing to better show their photographic works on Flickr. Moreover, up to 65.6% and 88.1% *Flickr_p* and *Flickr_6* photos contain the photographer information, which could pose a great risk of identity leakage to photo owners. Additionally, about a half of *Flickr_6* photos contain modification information. Most photos in the set are taken with professional digital cameras and photo owners often show intense interest in refining their works with image processing software. By contrast, a much lower percentage of *Flickr_p* photos taken with smartphones are modified.

4.4 Wild Photos

In the “wild” stage, most online photos have lingered on the Internet for a while and may have experienced multiple modifications by the hosting sites. In this section, we attempt to figure out the metadata information remaining in the “wild” photos and explore how the top media sites handle the photos hosted on them.

4.4.1 Data Collection

We employed two methods to collect “wild” photos. The first method is to randomly collect photos by Google Images Search. In the custom search control panel, we set

the image type as photo, file type as JPG/JPEG files, image size as larger than 400*300, and the date range from January 1, 2012 until January 1, 2015. Nearly all digital photos are in JPEG format. The specified image size can filter out most of graphs, drawings, and other non-photo images. In addition, we only focus on the photos posted online in the past three years. We totally collected 38,140 photos in this way and denoted them by *GoogleImage*.

Secondly, to investigate top media sites' policies on handling photos, we need to obtain a representative set of media sites. Alexa categorizes millions of sites and defines a list of site categories [75], from which we selected seven categories, which are "social networking," "weblog," "news," "college," "government," "classified," and "shopping". The reason why we chose them is that presumably the sites in these categories usually host large amounts of photos. Alexa provides for each category a list of the top 500 sites. We selected the top 100 sites for each category and thus we had 700 unique top ranked sites in total as our subject representative of online media sites.

Not every photo appearing on a site is hosted by the site. A photo is considered being hosted on a site only if its image URL has the same domain as the site URL. Only the photos hosted on a site are eligible to be used for studying the site's policies. During our photo collection from each site, we only crawled the photos hosted on that site. Specifically, for each of the 700 sites, we attempted to crawl 1,000 photos that appeared online after January 1, 2012. Those photos are expected to reflect the photo policy used by the hosting site under an assumption that the site has not made significant changes to its photo handling policy in the recent years. Due to unexpected factors including network connection failure and access permission denied, we were able to crawl 97,664 photos from 679 unique sites. To ensure the

representativeness of these photos, we filtered out the sites from which less than 10 photos were collected. Finally, we had 97,403 photos for 611 unique sites as our dataset for the study, about 160 photos per site on average. This set of photos are denoted as *TopSitesPhoto*.

4.4.2 Ethical Consideration

In our study, we leveraged several methods to collect photos, including: (1) soliciting “fresh” photos from crowdsourcing workers, (2) crawling photos from Flickr using its API, (3) random Google Image Search, and (4) crawling top websites for limited amounts of photos. Note that our crowdsourcing study has been vetted and approved by the Institutional Review Board (IRB) at our institution. During our photo collection, we did not receive any concerns or get warnings from those involved sites and did not interfere with their normal operations. In addition, with the collected photos, we anonymized the metadata information embedded before using them for study. We strictly abide by the copyright licenses if present.

Figure 4.4 depicts the number of photos crawled from each site. It shows that about 80% of sites have over 60 photos crawled, about 35% of sites have over 120 photos crawled, and about 20% have over 300 photos crawled. We crawled a maximum number of 1,026 photos for one site².

4.4.3 Metadata Information Embedded

Figure 4.5 shows the percentages of “wild” photos containing metadata, especially those sensitive metadata fields. It shows that the percentages of “wild” photos containing metadata information in the sets *GoogleImage* and *TopSitesPhoto* are 41.5%

²We crawled the site twice and collected over 1,000 photos.

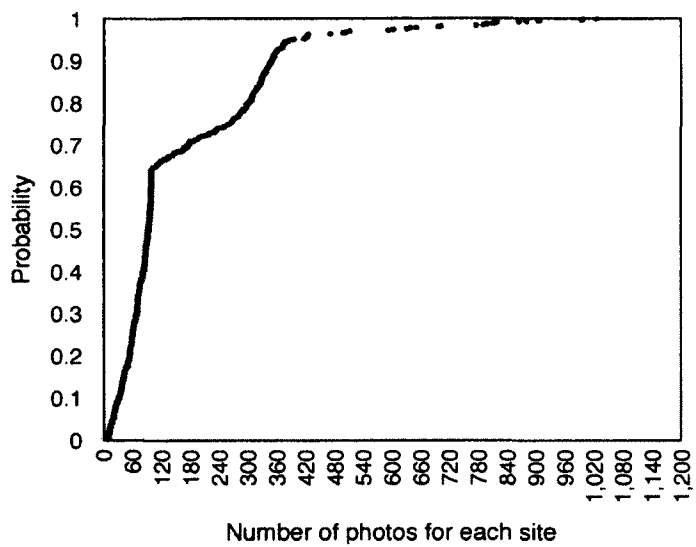


Figure 4.4: CDF of number of photos crawled from each site.

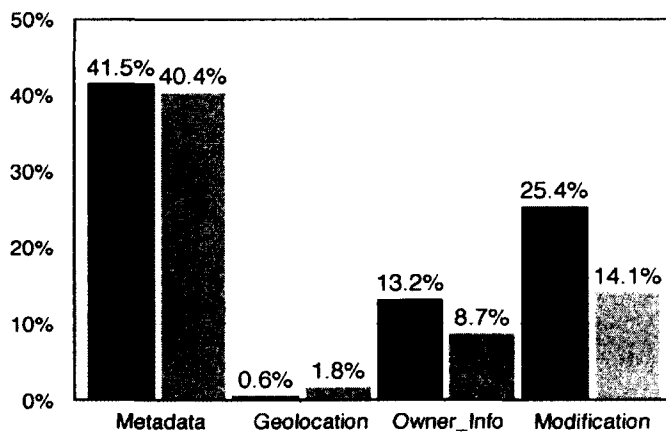


Figure 4.5: Percentage of “wild” photos containing metadata information. In each of four pairs of columns, the left black column represents *GoogleImage* while the right gray *TopSitesPhoto*.

and 40.4%, respectively, which are much smaller than that of “intact” photos (up to 94.1%). In addition, very few “wild” photos are tagged with GPS coordinates. In *GoogleImage* and *TopSitesPhoto*, the percentages are 0.6% and 1.8%, respectively, smaller than those of “fresh” and “intact” photos. Moreover, only 13.2% of *GoogleImage* photos and 8.7% of *TopSitesPhoto* photos contain photographers’ identification information. About 25.4% of *GoogleImage* photos and 14.1% of *TopSitesPhoto* photos contain modification history information. These results imply that compared to “fresh” and “intact” photos, a considerable proportion of “wild” photos have their embedded metadata stripped away.

4.4.4 Inferring Online Sites’ Photo Handling Policies

Based on *TopSitesPhoto*, we have built a set of photos for each of the 611 unique sites. We attempt to infer a site’s photo handling policy by characterizing the photos collected from the site. Specifically, we aim to answer two questions about a site’s photo handling policy. One is whether the site resizes the photos it hosts, and the other is whether the site removes the metadata information embedded in those photos.

Whether a site resizes its hosted photos? After upload, a photo is typically compressed and resized by the hosting site in several sizes. For instance, Instagram uses an image size of 640 pixels in width and 640 pixels in height for nearly all its hosted photos. More commonly, an online site confines a photo’s longest side length to a small set of values. Flickr resizes its photos in the following sizes: 100 pixels (on the longest side), 240 pixels, 800 pixels, 1600 pixels and so on [81]. Therefore, if the majority of photos hosted by a site have their longest side (width or height) lengths falling into a small set of numbers, then we speculate that the site does resize the

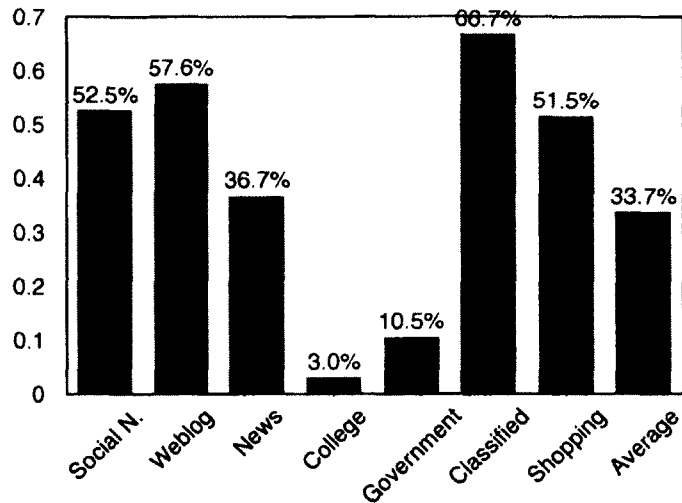


Figure 4.6: Percentage of sites estimated to resize their photos across the seven categories.

photos it hosts.

For each photo in our dataset, we retrieved its longest side length from its file information. About 2% of photos had no image size information available and were ruled out. Suppose “*DDDD*” is the longest side length value that is observed most frequently on a site. We calculated the proportion of the photos on the site with their longest side length of the value “*DDDD*”. We then leveraged the proportion number to decide whether the site resizes its photos or not. If over 50% of photos on the site have the longest side length of “*DDDD*”, the site is considered to resize its photos. The argument is based on our observation that among more than 40,000 photos downloaded from Flickr with “original size” option, only 3.47% have their longest side length of 1,600 pixels, while this length value occurs much more frequently for the photos that have been resized.

Figure 4.6 shows what percentage of sites that are regarded to resize the photos on their sites across the 7 categories. It is not surprising to see that only 3.0% of “College” sites and 10.5% “Government” sites have resized their photos, since

colleges and governments usually have sufficient hosting resources to store high-resolution photos. About 36.7% of “News” sites are estimated to resize the photos they host. A close examination reveals that news sites often resize their photos to many different sizes, which thereby lowers the percentage of photos with a unique longest side length size. In reality, there are probably much more news sites that resize their photos. In each of the other four categories, “Social networking,” “Weblog,” “Classified,” and “Shopping,” over 50% of sites have resized the photos they host. The sites in those categories often contain large amounts of photos and resizing photos is an effective means to save valuable storage space. Irrespective of categories, at least one third of all sites in our dataset are regarded to resize the photos they host. Note that our results represent a lower bound of the percentage of sites that resize their photos.

Whether a site strips out the metadata information embedded in the photos it hosts? There is another issue people may be concerned about when they upload photos online. As mentioned before, we use two fields in the metadata—camera make and model—to determine if the metadata information exists or not. For each site in our dataset, we calculated the percentage of its photos containing metadata information. Note that a photo may have its metadata information erased by its owner before posted online. Thus, our estimated percentage of online sites that strip out the metadata information of the photos they host represents an upper bound.

Figure 4.7 shows the CDF of the percentage of photos containing metadata information on each of the 611 sites in the seven categories. About 16% of sites have no photos containing metadata information. It is highly probable that those sites remove the metadata information from all hosted photos. About 45% of total

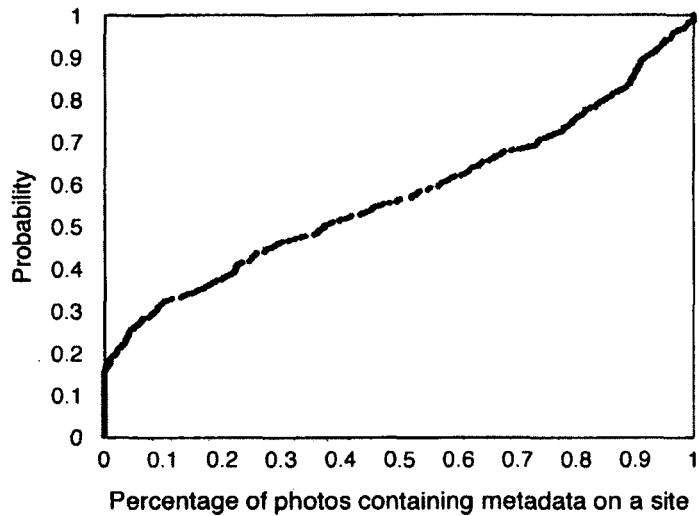


Figure 4.7: CDF of the percentage of photos containing metadata information on each site.

sites have at least half of their hosted photos containing metadata information. We determine that a site adopts a policy of removing photo metadata information if no photos hosted by the site contain metadata information; otherwise, the site is considered to preserve the metadata information of photos it hosts.

Figure 4.8 shows the percentage of sites in each category which are estimated to preserve the metadata information of photos they host. Again we found that the two categories “College” and “Government” present quite different statistical characteristics in preserving the photo metadata than the rest five categories. Specifically, 98% of college sites and 93.7% of government sites are estimated to preserve the photo metadata information. Combined with the above estimation results on a site’s photo resizing policy, we draw the conclusion that college and government sites seldom resize the photos they host or remove the embedded photo metadata information. In each of the other five categories, the proportions of the sites that preserve the photo metadata information are between 40% and 60%, much lower than those of college and government sites. On average, up to 68.4% of the top sites

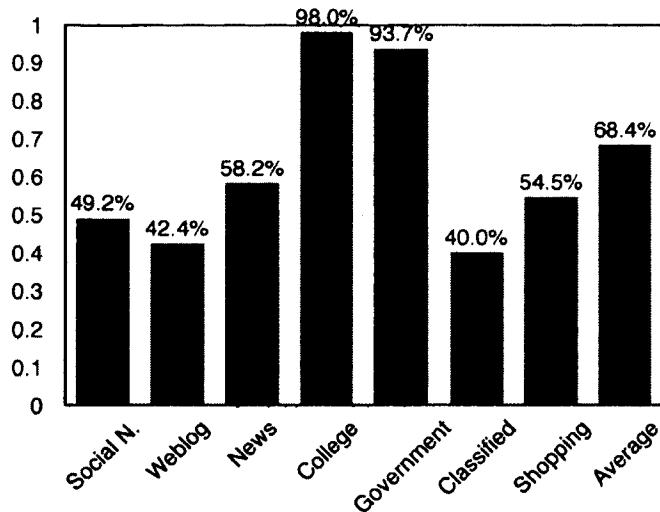


Figure 4.8: Percentage of sites estimated to preserve the photo metadata information across the seven categories.

in the seven categories preserve the photo metadata information, which suggests that a number of online photos may still have their metadata information open to public access for years.

4.5 Re-Identification Attack

Except the sensitive metadata fields including geolocation, owner’s information, and modification history, other metadata fields may appear relatively innocent. However, in this section, we demonstrate the feasibility of exploiting a trivial looking metadata field for re-identification attack.

Even without the photographer information explicitly included, a photographer can still be identified based on even only one photo she ever took. This can happen through a new attack vector—the camera serial number field in the photo metadata. A camera serial number can uniquely identify a camera most of the time.³ All photos

³A serial number is unique within a camera brand. Combined with camera make and model, a

taken with a same digital camera are supposed to have the same serial number if provided.⁴ In theory, a single photo with a camera serial number embedded could be used to trace other online photos taken with the same camera. Those photos together facilitate identifying the photographer.

We figured out that a public online database *stolencamerafinder* [74] could be leveraged to search for online photos tagged with a given camera serial number, although the online service was established to help find stolen cameras. For each given serial number, *stolencamerafinder* returns a list of online photos taken with the same camera, and for each photo provides the page URL where the photo is posted and the image URL linking to the photo.

Next, we do experiments to prove it quite easy to identify a photo owner with only one photo she ever shared online in the case that the photo has a camera serial number embedded. About 12% of the “wild” photos in the two sets *GoogleImage* and *TopSitesPhoto* were found to contain the serial number information. We randomly selected 2,000 unique serial numbers from them, then manually searched each serial number in the *stolencamerafinder*, and finally got back search results for 1,037 serial numbers in total. Note that not every camera serial number could get search results back. For those 1,037 serial numbers, by following the image URLs returned, we collected 38,140 photos that were posted on 4,712 unique websites. The photos collected for a specific serial number only represent a subset of all photos available online and tagged with the same serial number, due to the impossibility of finding all online photos with a given serial number.

Figure 4.9 shows the cumulative distribution function (CDF) of the number of photos that a single serial number links to. About 30% of serial numbers link to serial number can uniquely identify a camera.

⁴Smartphones typically do not store their serial numbers in their photos.

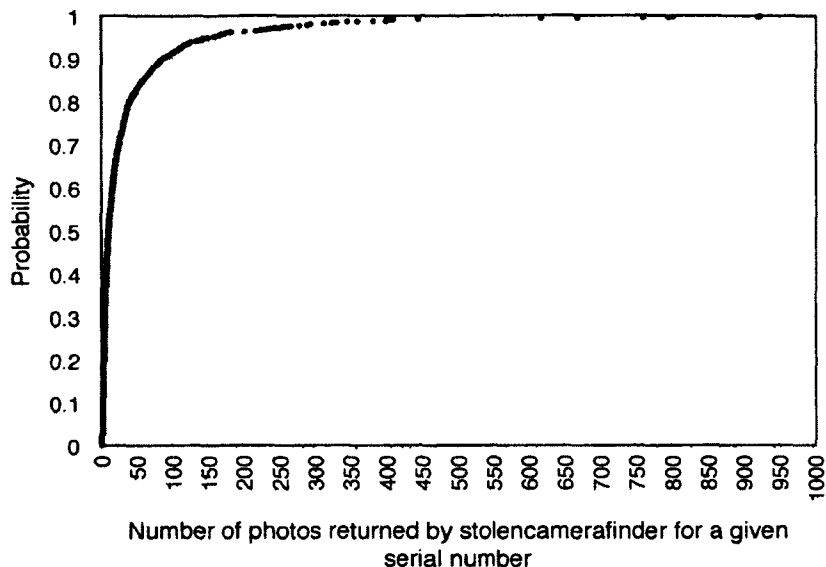


Figure 4.9: CDF of the number of photos returned by *stolencamerafinder* for a given serial number.

over 25 photos and about 10% link to over 100 photos. The average number of photos linked to a same serial number is 36.8, the median is 10, and the maximum is 923. With the considerable number of photos tagged with a same camera serial number, together with the page URLs where the photos are posted, and the photos already existing in the photo sets *GoogleImage* and *TopSitesPhoto*, we were able to set up a knowledge base for each serial number (tentatively a digital camera). The rich information available can evidently disclose much more privacy information about the camera owner than a single serial number itself. This demonstrates the potential of a camera serial number as an attractive attack vector for mounting privacy attacks.

Identifying a Photographer. The page URL and the page where a photo is posted can provide important clues to reveal a photographer’s online identity. For instance, the URL <https://plus.google.com/XYZ/photos> suggests that the photographer should have a Google+ [79] account with the ID of “XYZ”. Following the

URL allows us to retrieve more information about the photographer, such as her real name, college attended, current employer, and photos posted on her account page. We have observed a great many such URL strings in our dataset with photographers' online social networks (OSNs) account IDs embedded. The involved OSNs include but not limited to Flickr, Facebook [77], Twitter [78], Google+, and 500px [80]. A photographer may have her multiple OSN accounts disclosed in this way. Table 4.4 lists the information typically contained in an account profile of the five social networks mentioned above. It shows that an account profile typically contains demographics and other sensitive information including age, gender, education, occupation, living city, other OSN accounts, and much more. Once one OSN account is identified, the true identity of the user in the real world can be readily disclosed.

OSN	Account Profile Information
Flickr	Name, Occupation, Living City, Hometown, Gender, Personal Website(s), Email, Joined Time, Biography, Age, Religion
500px	Name, Biography, Living City, Contact, other OSN accounts
Google+	Name, Gender, Living City, Colleges Attended, Current Employer, Work Experience
Twitter	Name, Occupation, Living City, Telephone, Email, Personal Website(s), Joined Time, Photos and Videos, Tweets, Followings, Followers and Favorites
Facebook	Name, Living City, Gender, Education, Telephone, other OSN accounts, Life Events

Table 4.4: List of the information typically contained in an account profile in each of the five OSNs. Note that the listed information represents the maximum amount of information available with public permissions of an OSN account.

Figure 4.10 shows the percentage of serial numbers from which we are able to identify the corresponding camera owners' IDs in one or more OSNs by scrutinizing the page URLs where the photos were posted. Among the 1,037 unique serial numbers in our dataset, 51.4% (533) of the serial numbers have the camera owners' OSN accounts identified, and 9.0% (93) have account IDs in two or more OSNs identified. And for one serial number we even identified the camera owner's four account IDs in four OSNs respectively.

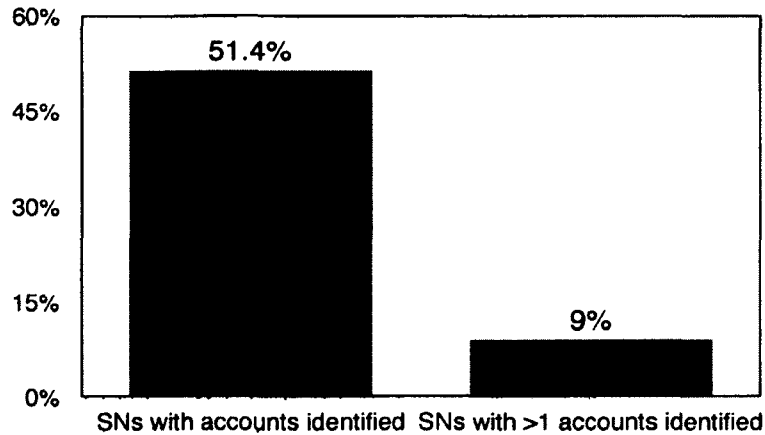


Figure 4.10: Percentage of camera serial numbers (SNs) with camera owners' OSN accounts identified.

As mentioned before, we were able to retrieve about 37 online photos on average for a given serial number. Those photos tagged with the same serial number may contain metadata information that could help identify the photographer. We closely examined the metadata information embedded in the related photos for each of the remaining 504 serial numbers without any OSN accounts identified in the previous step. Among them, we successfully identified the photographers for 116 serial numbers. Compared to the photographers with their OSN accounts identified, the available information on those 116 photographers are restricted to the photo metadata embedded, mainly including their names, the processing softwares, and OSes used. However, more information could be collected online once a person's name is identified. Overall, 62.6% (649) of serial numbers have had their photographers identified.

4.6 Discussion

One goal of this work is to track the propagation of the sensitive metadata information embedded in the digital photos at different stages. One ideal way is to monitor the process of creation, modification, and elimination of the metadata information contained in a same set of photos that sequentially experience three stages—“fresh,” “intact,” and “wild.” However, it is very hard to obtain such an ideal photo set in large-scale. Instead, we employed different data collection methods and obtained three kinds of photo sets to represent the digital photos at the corresponding three stages.

We collected 612 valid “fresh” photos through crowdsourcing in a period of two months. Each photo collected was taken by a unique participant with a unique device, and participants from 76 countries contributed to this dataset. In addition, those photos were solicited directly from smartphones and no photos taken with digital cameras were collected in order to avoid data contamination. Therefore, although the dataset size of “fresh” photos is not comparable to those of “intact” and “wild” photos, its representativeness is high enough for this study.

To infer online media sites’ policies on handling metadata information in the photos they host, we adopt a passive approach, that is, by examining the metadata information of the photos collected from the sites. Actually, we once considered to take an active approach to detect media sites’ policies, by submitting (uploading) different types of photos to the sites, then re-downloading them, and comparing metadata fields. However, we had to abandon this approach because most of the 611 sites in the seven categories have specific user groups and are not open to public registration, not to mention photo uploading.

Although it is known that a camera serial number can uniquely identify a camera

to some extent, we are not aware of any previous research work revealing potential threats arising from this attribute in an empirical and systematic manner. We demonstrated the feasibility of re-identification attack by exploiting camera serial number. We were able to identify over 60% of photo owners based on their camera serial numbers available in a public online database.

When a user shares a digital photo online, two questions about privacy issues are readily raised. One is whether sensitive hidden metadata information is embedded in the photo. The other concerning question is what the media site will do with the photo. According to our experiment results, a considerable proportion of digital photos contain sensitive metadata information, and many sites resize the photos they host or remove the embedded photo metadata information. In our future work, we will develop a browser extension to give users direct answers to these two questions.

Sensitive Metadata	Potential Threats	Website's Policy
Geolocation	Location disclosure, house robbery	Metadata removing
Photographer's Name	Identity disclosure	Photo resizing
Modification History	Undermining photo's authenticity	NA
Camera Serial Number	Re-identification attack	NA

Table 4.5: Main functions of the browser extension prototype

The major functions that the tool should have are illustrated in Table 4.5. Specifically, once the sensitive metadata information in a photo being uploaded is detected, the browser extension should issue an alarm by popping up a window on the screen and provide customized alert information, including the sensitive metadata information embedded, the corresponding privacy risks, and the current visiting site's policy on photo handling. Note that the browser extension should display the alert information only when the privacy-related metadata information is detected, and thus it should not often interfere with normal photo upload workflows. Although there are already browser extensions for photo metadata visualization, we will focus

on informing users of the sensitive metadata contained and customized privacy risks. Moreover, we will ensure users' right to know the actions that the hosting media sites will perform on their photos.

4.7 Related Work

Several previous works conduct user studies to understand users' privacy decisions during the photo sharing process and their privacy concerns on others' photo-sharing activities. Clark et al. [82] revealed the problem of unintended photo storage without users' awareness, which is mainly caused by the automatic features of cloud-based photo backup services. Ahern et al. [83] found that mobile users' decisions to post photos privately or publicly were determined more by identity or impression concerns than security concerns. Besmer et al. [84] made similar findings. They studied users' perception of being tagged in undesired photos uploaded by others. They found that a user's privacy concerns on that domain were mainly related to identity and impression management within her existing social circles. Henne et al. [85] showed in their survey results that among the information potentially disclosed by the tagged photos, personal references and location data raised most privacy concerns.

More related to our work, several researchers examined the privacy threat posed by the textual metadata information contained in online photos. Friedland and Sommer [86] focused on the privacy threats posed by the geolocation information available online. They showed that the geolocation data could be exploited to mount privacy attacks using three scenarios on Craigslist, Twitter, and YouTube, respectively. Pesce et al. [91] demonstrated that photo tagging on Facebook could be exploited to enhance prediction of users' information like gender, city, and country.

Another work from Mahmood and Desmedt [90] discussed possible privacy violations from Google+'s policy that any users who access a photo can see its metadata online. While the above three works addressed the privacy issues with photos, we investigated the privacy issues with online photos on a much larger scale. We assessed the privacy risks arising from leakage of all possible sensitive metadata information rather than just geolocation data. Moreover, our study is not restricted to one media site. Instead, we collect our photo dataset from hundreds of top-ranked websites and through crowdsourcing platforms. Those photos cover various stages, i.e., "fresh," "intact," and "wild." In addition, we introduce a new attack vector and show its unexpected power in conducting a re-identification attack. We also performed a large-scale measurement of photo handling policies adopted by various categories of media sites.

Another large body of previous work has attempted to enhance people's privacy when sharing photos online. Besmer et al. [96] designed a privacy enhancement tool to improve the photo tagging process on Facebook. The tool allows tagged users to negotiate online with the photo uploaders about the permission settings on the photo. Fang and LeFevre [92] built a machine learning model for OSN users to configure privacy settings automatically with a limited number of rules provided. Zerr et al. [97] developed privacy classification models for users to search for private photos about themselves posted by others at an early stage. Henne et al. [95] proposed a watchdog service that allows users to keep track of potentially harmful photos uploaded by others at the expense of sharing their location data with the service. Ra et al. [93] presented a selective encryption algorithm that enables a photo to hide its "secret" part from the host photo-sharing site and the unauthorized viewers and only expose its "public" part. Ilia et al. [94] refined the access control

mechanism currently used by OSNs on photo sharing. The new mechanism allows the depicted users in a photo to decide the exposure of their own face, and could present photos with the restricted faces blurred out to a visitor. Complementary to those works attempting to enhance privacy on the web server side, this study assesses the privacy risks arising from sensitive photo metadata and provides some guidelines for developing client-side privacy leakage prevention tools, which should be able to alert online users of potential privacy risks posed by uploading photos and also inform them of the photo handling policies adopted by the currently visiting website.

To the best of our knowledge, we have conducted the first large-scale empirical measurement study of the status of contemporary digital photos at the three different stages. In addition to examining the sensitive metadata information embedded, we inferred the photo handling policies used by hundreds of top-ranked sites, and proposed to exploit the camera identification number as an attack vector for re-identification attack. We are not aware of any previous work studying these topics.

4.8 Conclusion

In this chapter, we performed a data-driven assessment of privacy risks on contemporary digital photos. We first collected from the Web nearly 200,000 digital photos at three different stages as our dataset. Then for photos at each stage, we measured the prevalence of metadata and assessed the privacy risks posed by metadata leakage. We found that metadata is quite prevalent among digital photos at each stage. In particular, 15% of “fresh” photos, about 30% “intact” photos, and about 1% “wild” photos were tagged with GPS coordinates. The percentage of “wild” photos containing other sensitive metadata information is also much lower

than that of “intact” photos. A possible reason is that online sites often remove the metadata information of the photos they host. Our speculation was confirmed by our investigation of photo handling policies based on nearly 100,000 photos crawled from 679 top sites in seven categories. We further found that photo policies used by a site vary with the category that the site belongs to. Finally, we proposed to use the camera serial number as a new attack vector towards privacy inference and demonstrated its power in deriving both online and real-world identities of a photographer with just one photo she ever took. In our future work, we will build a browser extension prototype to prevent users’ photo privacy leakage and increase their knowledge of the online services’ policies on photo handling.

Chapter 5

Conclusion and Future Work

Internet business is playing a more and more important role in the global economy. In this dissertation, we mainly focus on uncovering and detecting the fraudulent activities in two kinds of important Internet businesses, online advertising system and e-commerce system. In our first project, we proposed a novel challenge-response based detection mechanism for advertisers to independently detect fraudulent clicks on their advertisements without the help of either publishers or ad networks. We verified the effectiveness of the detection mechanism by deploying it in the real world. In our second project, we introduced a newly emerging threat to the e-commerce reputation system. By infiltrating several underground markets providing the SRE (seller-reputation-escalation) service, we performed a deep analysis of the operational characteristics of those markets. We also investigated the effectiveness of SRE services and found that online sellers using SRE services can increase their store reputation 10 times faster than legitimate ones. In addition, e-commerce sites represent one kind of sites that are hit hardest by trash web traffic generated by bad bots. To assess the potential privacy risks arising from online photo sharing, in our third project, we performed a data-driven assessment of privacy risks on contempo-

rary digital photos. We found that privacy sensitive metadata information is quite prevalent among digital photos at each stage. We also proposed to use the camera serial number as a new attack vector towards privacy inference and demonstrated its power in deriving both online and real-world identities of a photographer with just one photo she ever took.

The e-commerce business is increasingly thriving in the recent several years. In our future work, we will still focus our attention on the fake transaction problem that are greatly affecting the continuous development of e-commerce business. Our goal is to develop a practical fake transaction detection system for e-commerce marketplaces to identify a fake transaction before it is finished. Our previous study on the SRE service has shown that to finish a fake transaction, the hired human laborers have to follow the operational steps predefined by the SRE market operators and it usually takes them several days to complete a fake transaction. Thus we could first research on the operational strategies that were specially designed for conducting fake transactions and then leverage the inferred behavioral patterns of human laborers on the e-commerce sites as the signature to identify fake transactions in time. We plan to cooperate with one large e-commerce marketplace to implement and deploy the detection system.

Bibliography

- [1] Global cost of online fraud. <http://www.theguardian.com/technology/2013/oct/30/online-fraud-costs-more-than-100-billion-dollars>
- [2] Online advertising. https://en.wikipedia.org/wiki/Online_advertising.
- [3] Chameleon Botnet. <http://www.spider.io/blog/2013/03/chameleon-botnet/>.
- [4] How browsers work. <http://taligarsiel.com/Projects/howbrowserswork1.htm>.
- [5] How many users have JavaScript disabled? <https://developer.yahoo.com/blogs/ydnfourblog/many-users-javascript-disabled-14121.html>.
- [6] Top 5 browsers. <http://gs.statcounter.com/>.
- [7] MaxMind. http://www.maxmind.com/en/web_services.
- [8] Usage share of web browsers. http://en.wikipedia.org/wiki/Usage_share_of_web_browsers.
- [9] Blacklistalert. <http://www.blacklistalert.org/>.

- [10] Weka. <http://www.cs.waikato.ac.nz/ml/weka/>.
- [11] Sar. [http://en.wikipedia.org/wiki/Sar_\(Unix\)](http://en.wikipedia.org/wiki/Sar_(Unix)).
- [12] S. A. Alrwais, C. W. Dun, M. Gupta, A. Gerber, O. Spatscheck, and E. Osterweil. Dissecting ghost clicks: Ad fraud via misdirected human clicks. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)*, 2012.
- [13] N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder. Online advertising fraud. In *Crimeware: Understanding New Attacks and Defenses*. Addison-Wesley Professional, 2008.
- [14] N. Daswani and M. Stoppelman. The anatomy of clickbot.a. In *Proceedings of the Workshop on Hot Topics in Understanding Botnets (HotBots)*, 2007.
- [15] V. Dave, S. Guha, and Y. Zhang. Measuring and fingerprinting click-spam in ad networks. In *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, 2012.
- [16] V. Dave, S. Guha, and Y. Zhang. Viceroi: Catching click-spam in search ad networks. In *Proceedings of ACM Conference on Computer and Communications Security (CCS)*, 2013.
- [17] P. Eckersley. How unique is your web browser? In *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2010.
- [18] H. Haddadi. Fighting online click-fraud using bluff ads. In *ACM SIGCOMM Computer Communication Review (CCR)*, 2010.

- [19] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang. Knowing your enemy: Understanding and detecting malicious web advertising. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, 2012.
- [20] A. Metwally. Detectives: Detecting coalition hit inflation attacks in advertising networks streams. In *Proceedings of the International Conference on World Wide Web (WWW)*, 2007.
- [21] A. Metwally. Sleuth: Single-publisher attack detection using correlation hunting. In *Proceedings of the International Conference on Very Large Data Bases (VLDB)*, 2008.
- [22] A. Metwally, D. Agrawal, and A. E. Abbadi. Duplicate detection in click streams. In *Proceedings of the International Conference on World Wide Web (WWW)*, 2005.
- [23] B. Miller, P. Pearce, C. Grier, C. Kreibich, and V. Paxson. What's clicking what? techniques and innovations of today's clickbots. In *Proceedings of the International Conference on Detection of Intrusions, Malware & Vulnerability Assessment (DIMVA)*, 2011.
- [24] J. Quinlan. C4.5: Programs for machine learning. Morgan Kaufmann Publishers, 1993.
- [25] B. Schulte, H. Andrianakis, K. Sun, and A. Stavrou. Netgator: Malware detection using program interactive challenges. In *Proceedings of the International Conference on Detection of Intrusions, Malware & Vulnerability Assessment (DIMVA)*, 2012.

- [26] T. Yen, X. Huang, F. Monrose, and M. Reiter. Browser fingerprinting from coarse traffic summaries: Techniques and implications. In *Proceedings of the International Conference on Detection of Intrusions, Malware & Vulnerability Assessment (DIMVA)*, 2009.
- [27] Taobao marketplace. <http://en.wikipedia.org/wiki/Taobao>.
- [28] eBay marketplace. <http://blog.ebay.com/ebay-marketplaces-introduces-new-logo/>.
- [29] Alibaba Group. http://en.wikipedia.org/wiki/Alibaba_Group.
- [30] Alibaba exec on fake transactions. <http://blogs.wsj.com/chinarealtime/2015/03/03/cat-and-mouse-game-alibaba-exec-on-fake-transactions/>
- [31] Alibaba, China's Internet behemoth. http://www.jamaicaobserver.com/business/Alibaba--China-s-Internet-behemoth_17598819.
- [32] Alexa top sites. <http://www.alexa.com/topsites/global>.
- [33] Alibaba plunks down \$692M to push into offline retail. <http://venturebeat.com/2014/03/31/alibaba-plunks-down-692m-to-push-into-offline-retail/>.
- [34] SKY brushing marketplace. <http://www2.88sxy.com/>.
- [35] WOOD brushing marketplace. <http://www.ntyjy.com/>.
- [36] EMPIRE brushing marketplace. <http://www.shuazuanshuaxinyu.com/>.
- [37] COOL brushing marketplace. <http://www.kus.cc/Index.html>.

- [38] NET brushing marketplace. <http://www.shuakewang.com/>.
- [39] GDP per capita. <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
- [40] Jingdong marketplace. 2015. <http://www.jd.com/>.
- [41] Taobao's penalty rules (in Chinese). 2015. <https://rule.taobao.com/detail-113.htm>.
- [42] Amazon Mechanical Turk. http://en.wikipedia.org/wiki/Amazon_Mechanical_Turk.
- [43] A SRE market makes off with users' money. <http://bbs.tianya.cn/post-law-438844-1.shtml>.
- [44] A SRE market makes off with million US dollars. http://weita.taobao.com/tzh/feed/feed_detail_display.htm?feedId=100505423&wsnsUid=2054965595.
- [45] K. Hoffman, D. Zage, and C. Nita-Rotaru. A Survey of Attack and Defense Techniques for Reputation Systems. In *ACM Computing Surveys (CSUR)*, 2009.
- [46] J. Caballero, C. Grier, C. Kreibich, and V. Paxson. Measuring pay-per-install: The commoditization of malware distribution. In *Proceedings of the 20th USENIX Security Symposium*, 2011.
- [47] N. Christin. Traveling the silk road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd International Conference on World Wide Web (WWW)*, 2013.

- [48] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. In *Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS)*, 2007.
- [49] C. Grier, L. Ballard, J. Caballero, N. Chachra, C. J. Dietrich, K. Levchenko, P. Mavrommatis, D. McCoy, A. Nappa, A. Pitsillidis, N. Provos, M. Z. Rafique, M. A. Rajab, C. Rossow, K. Thomas, V. Paxson, S. Savage, and G. M. Voelker. Manufacturing compromise: The emergence of exploit-as-a-service. In *Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS)*, 2012.
- [50] C. Ho, Y. Zhang, J. W. Vaughan, and M. van der Schaar. Towards social norm design for crowdsourcing markets. In *Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP)*, 2012.
- [51] P. G. Ipeirotis, F. Provost, and J. Wang. Quality management on amazon mechanical turk. In *Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP)*, 2010.
- [52] K. Levchenko, A. Pitsillidis, N. Chachra, B. Enright, M. Felegyhazi, C. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu, D. McCoy, N. Weaver, V. Paxson, G. M. Voelker, and S. Savage. Click trajectories: End-to-end analysis of the spam value chain. In *Proceedings of the 32nd IEEE Symposium on Security and Privacy*, 2011.
- [53] J. Li, M. Ott, and C. Cardie. Identifying manipulated offerings on review portals. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2013.

- [54] D. Mayzlin, Y. Dover, and J. Chevalier. Promotional reviews: An empirical investigation of online review manipulation. in *National Bureau of Economic Research, No. w18340*, 2012.
- [55] C. Nosko, and S. Tadelis. The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment. *Working paper*, 2014
- [56] C. Dellarocas, and C. A. Wood. The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias. *Management Science*, 54(3):460-476, 2008
- [57] D. McCoy, H. Dharmdasani, C. Kreibich, G. M. Voelker, and S. Savage. Priceless: The role of payments in abuse-advertised goods. In *Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS)*, 2012.
- [58] D. McCoy, A. Pitsillidis, G. Jordan, N. Weaver, C. Kreibich, B. Krebs, G. M. Voelker, S. Savage, and K. Levchenko. Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs. In *Proceedings of the 21st USENIX Security Symposium*, 2012.
- [59] M. Motoyama, K. Levchenko, C. Kanich, D. McCoy, G. M. Voelker, and S. Savage. Re: Captchas - understanding captcha-solving services in an economic context. In *Proceedings of the 19th USENIX Security Symposium*, 2010.
- [60] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. An analysis of underground forums. In *Proceedings of ACM SIGCOMM conference on Internet Measurement Conference (IMC)*, 2011.

- [61] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. Dirty jobs: The role of freelance labor in web service abuse. In *Proceedings of the 20th USENIX Security Symposium*, 2011.
- [62] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh. Spotting opinion spammers using behavioral footprints. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2013.
- [63] M. Ott, C. Cardie, and J. Hancock. Estimating the prevalence of deception in online review communities. In *Proceedings of the 21st International Conference on World Wide Web (WWW)*, 2012.
- [64] Y. Park, J. Jones, D. McCoy, E. Shi, and M. Jakobsson. Scambaiter: Understanding targeted nigerian scams on craigslist. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2014.
- [65] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, and B. Zhao. Follow the green: Growth and dynamics in twitter follower markets. In *Proceedings of ACM SIGCOMM conference on Internet Measurement Conference (IMC)*, 2013.
- [66] G. Swamynathan, K. C. Almeroth, and B. Y. Zhao. The design of a reliable reputation system. In *Electronic Commerce Research*, 10.3-4 (2010): 239-270.
- [67] L. Akoglu, R. Chandy, and C. Faloutsos. 2013. Opinion Fraud Detection in Online Reviews by Network Effects. In *Proceedings of the 7th international AAAI Conference on Weblogs and Social Media (ICWSM)*.

- [68] S. Rayana, and L. Akoglu. 2015. Collective Opinion Spam Detection: Bridging Review Networks and Metadata. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*.
- [69] Y. Chen, and H. Chen. 2015. Opinion Spam Detection in Web Forum: A Real Case Study. In *Proceedings of the 24th International Conference on World Wide Web (WWW)*.
- [70] K. Thomas, C. Grier, V. Paxson, and D. Song. Suspended accounts in retrospect: An analysis of twitter spam. In *Proceedings of ACM SIGCOMM Conference on Internet Measurement Conference (IMC)*, 2011.
- [71] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson. Trafficking fraudulent accounts: The role of the underground market in twitter spam and abuse. In *Proceedings of the 22nd USENIX Security Symposium*, 2013.
- [72] Number of photos uploaded to Flickr. <https://www.flickr.com/photos/franckmichel/6855169886/>.
- [73] ExifTool library. <http://www.sno.phy.queensu.ca/~phil/exiftool/>.
- [74] Site stolencamerafinder: Find your camera. <http://www.stolencamerafinder.com/>
- [75] Alexa top sites by category. <http://www.alexa.com/topsites/category/Top>
- [76] McAfee's location is leaked with photo metadata. <http://www.wired.co.uk/news/archive/2012-12/04/vice-give-away-mcafee-location>
- [77] Facebook. <https://www.facebook.com/>

- [78] Twitter. <https://twitter.com/>
- [79] Google+. <https://plus.google.com/>
- [80] 500px. <https://500px.com/>
- [81] Flickr file size limits. <https://www.flickr.com/help/photos/>
- [82] J. W. Clark, P. Snyder, D. McCoy, and C. Kanich. I Saw Images I Didn't Even Know I Had: Understanding User Perceptions of Cloud Storage Privacy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2015.
- [83] S. Ahern, D. Eckles, N. Good, S. King, M. Naaman, and R. Nair. Overexposed? privacy patterns and considerations in online and mobile photo sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2007.
- [84] A. Besmer, and H. R. Lipford. Poster: Privacy Perceptions of Photo Sharing in Facebook. In *Proceedings of the 4th Symposium on Usable Privacy and Security (SOUPS)*, 2008.
- [85] B. Henne, and M. Smith. Awareness about photos on the web and how privacy-privacy-tradeoffs could help. In *Proceedings of the International Conference on Financial Cryptography and Data Security (FC)*, 2013.
- [86] G. Friedland, and R. Sommer. Cybercasing the joint: on the privacy implications of geo-tagging. In *Proceedings of the 5th USENIX Conference on Hot Topics in Security (HotSec)*, 2010.

- [87] G. Zhong, I. Goldberg, and U. Hengartner. Louis, lester and pierre: three protocols for location privacy. In *7th Privacy Enhancing Technologies Symposium (PETS)*, 2007.
- [88] J. Krumm. A survey of computational location privacy. In *Personal and Ubiquitous Computing*, 2009.
- [89] F. Olumofin, P. Tysowski, I. Goldberg, and U. Hengartner. Achieving efficient query privacy for location based services. In *10th Privacy Enhancing Technologies Symposium (PETS)*, 2010.
- [90] S. Mahmood, and Y. Desmedt. Poster: Preliminary analysis of Google+'s privacy. In *Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS)*, 2011.
- [91] J. P. Pesce, D. L. Casas, G. Rauber, and V. Almeida. Privacy attacks in social media using photo tagging networks: a case study with Facebook. In *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media (PSOSM)*, 2012.
- [92] L. Fang, and K. LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th International Conference on World Wide Web (WWW)*, 2010.
- [93] M. Ra, R. Govindan, and A. Ortega. P3: toward privacy-preserving photo sharing. In *Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2013.
- [94] P. Ilia, I. Polakis, E. Athanasopoulos, F. Maggi, and S. Ioannidis. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *Proceedings of*

the 22nd ACM Conference on Computer and Communications Security (CCS), 2015.

- [95] B. Henne, C. Szongott, and M. Smith. SnapMe if you can: privacy threats of other peoples' geo-tagged media and what we can do about it. In *Proceedings of the 6th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*, 2013.
- [96] A. Besmer, and H. R. Lipford. Moving beyond untagging: photo privacy in a tagged world. In *Proceedings of the 28th SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2010.
- [97] S. Zerr, S. Siersdorfer, J. Hare, and E. Demidova. Privacy-aware image classification and search. In *Proceedings of the 35th International ACM Conference on Research and Development in Information Retrieval (SIGIR)*, 2012.