

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

Haitao Xu^{*†}, Daiping Liu^{*}, Haining Wang^{*}, Angelos Stavrou[‡]

^{*}University of Delaware
Newark, DE 19716, USA
{hxu,dpliu,hnw}@udel.edu

[†]College of William and Mary
Williamsburg, VA 23187, USA
hxu@cs.wm.edu

[‡]George Mason University
Fairfax, VA 22030, USA
astavrou@gmu.edu

ABSTRACT

In online markets, a store's reputation is closely tied to its profitability. Sellers' desire to quickly achieve high reputation has fueled a profitable underground business, which operates as a specialized crowdsourcing marketplace and accumulates wealth by allowing online sellers to harness human laborers to conduct fake transactions for improving their stores' reputations. We term such an underground market a *seller-reputation-escalation (SRE) market*. In this paper, we investigate the impact of the SRE service on reputation escalation by performing in-depth measurements of the prevalence of the SRE service, the business model and market size of SRE markets, and the characteristics of sellers and offered laborers. To this end, we have infiltrated five SRE markets and studied their operations using daily data collection over a continuous period of two months. We identified more than 11,000 online sellers posting at least 219,165 fake-purchase tasks on the five SRE markets. These transactions earned at least \$46,438 in revenue for the five SRE markets, and the total value of merchandise involved exceeded \$3,452,530. Our study demonstrates that online sellers using SRE service can increase their stores' reputations at least 10 times faster than legitimate ones while only 2.2% of them were detected and penalized. Even worse, we found a newly launched service that can, within a single day, boost a seller's reputation by such a degree that would require a legitimate seller at least a year to accomplish. Finally, armed with our analysis of the operational characteristics of the underground economy, we offer some insights into potential mitigation strategies.

Categories and Subject Descriptors

K.4.4 [Electronic Commerce]: Security; H.3.5 [Online Information Services]: Web-based services

Keywords

E-commerce; reputation manipulation; fake transaction

1. INTRODUCTION

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2015, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3469-3/15/05.
<http://dx.doi.org/10.1145/2736277.2741650>.

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. A large number of online shopping markets including Amazon and eBay have incorporated reputation systems. The primary purpose of these reputation systems is to offer a means for users to give feedback on a product and its seller. Moreover, this process encourages sellers to provide better products because through the scoring process they are rewarded with higher reputation, 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 [34]. 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. Depending on the size and 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 spam and artificial ratings, among others. Several recent works [21,30,31] have aimed to tackle these problems. However, 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 online 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 reputation attracts legitimate customers and at the same time cements the seller's abil-

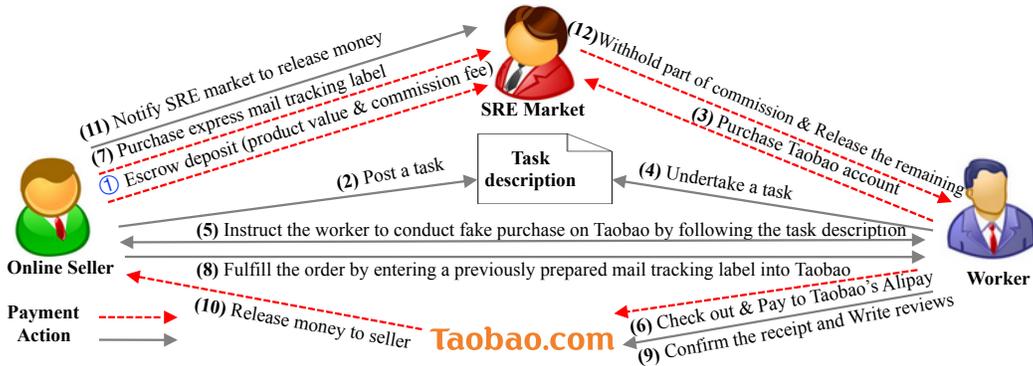


Figure 1: Lifecycle of a fake-purchase task on the SRE market.

ity to handle negative reviews. Furthermore, the process is fairly scalable, and the seller may post up to hundreds of such tasks each day, thus 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 paper, we present an empirical analysis of the prevalence of the SRE markets by infiltrating five popular SRE markets, which specialize in providing reputation-escalation services to sellers on the Taobao marketplace [1], the largest consumer-to-consumer (C2C) online marketplace in China. We conducted daily crawls of these five markets for two months and collected 219,165 tasks posted by more than 11,000 online sellers¹ and continuously monitored these sellers' activities on these markets. For approximately 4,000 sellers, we were able to identify their IDs on the Taobao marketplace. To validate our findings, we randomly selected a set of 4,000 legitimate Taobao sellers to act as our control group. We evaluate the effect of SRE services on escalating seller reputations by comparing the two groups' growth curves of their store reputations. We characterize the sellers and workers involved, and we estimate 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 within a single day 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.

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 escalation. We find that online sellers using SRE services can increase their store reputations 10 times faster

than legitimate ones, and only a small portion (2.2%) of them are detected and penalized.

- We discover a newly launched service on one SRE market that can boost a seller's reputation score within one day by such a degree that would require a legitimate seller at least one year to accomplish.

The remainder of the paper is organized as follows. Section 2 provides a brief introduction to Taobao and an overview of the business model of SRE markets. Section 3 describes our data collection. Section 4 presents the results from our infiltration of five SRE markets. Section 5 evaluates the effectiveness of SRE services. Section 6 discusses possible defensive interventions. Section 7 surveys the related work, followed by our conclusions in Section 8.

2. 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 [2]. 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.

2.1 Taobao Overview

Taobao, launched by Alibaba Group [3] 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 [4,5]. Taobao has achieved great success with 60 million daily visitors and 50,000 sales per minute, and is ranked 9th globally by Alexa [6] as of this writing. In 2013, the total gross merchandise traded on Alibaba was more than Amazon and eBay's gross sales combined [7].

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 ma-

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

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 1: Qualification types

for online marketplaces such as Amazon may also suffer from fake transactions conducted through SRE markets.

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

Figure 1 shows a typical lifecycle of a fake-purchase task on a SRE market, including five stages: (1) task creation, (2) task undertaken, (3) conducting fake purchase on Taobao, (4) order fulfillment, and (5) commission realization.

Task Creation. The lifecycle of a fake-purchase task begins with an online seller creating a task on an 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 1 lists the qualification types provided by an SRE market. Each one describes a qualification that 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-aged², 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 the worker to be familiar with the task flow. Q5 highlights the requirement for workers’ geographi-

²Pre-aged Taobao account refers to the account that has been created for some time and ever used for real purchases.

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.
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 predefined waiting time.

Table 2: A typical task description

cal 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 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**).

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 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 Fulfilment. 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

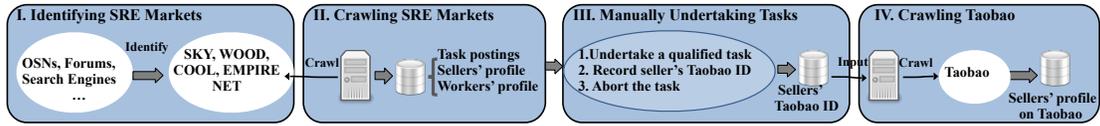


Figure 2: Procedure of data collection.

(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.

In summary, the worker typically needs a verified and aged Taobao account, invests several minutes and tens of US dollars in purchasing goods, and receives 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. DATA COLLECTION

In this section, we describe our crawling mechanism and summarize the dataset.

3.1 Crawling Mechanism

Figure 2 briefly illustrates the procedure of our data collection. First, we identified SRE markets (I). Then web crawlers were 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 undertook tasks for one month (III). For those sellers with their Taobao IDs recognized, we then crawled Taobao for their sales and reputation information to evaluate the impact of their misconducts done through the SRE markets upon the 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 [8], WOOD [9], EMPIRE [10], COOL [11], and NET [12]. All five markets have nearly identical web layout and source code. Whois domain lookup reveals that they have been created 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 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 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 of undertaking 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 had 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’ rating.

We took a similar recipe as in [18] to make sure our continuous crawl unnoticed 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 Data Summary

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: 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.

Table 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 successfully 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. 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.

4. 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.

4.1 SRE Market Popularity

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 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.

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

4.2 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.

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 re-

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

stricted 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 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.

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 the workers to fill in the shipping addresses of sellers’ choice to make fake transactions appear geographically distributed. Table 6 lists 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.

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

Imitating a Real Purchase on Taobao. In addition to buyers’ Taobao accounts, 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 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 action. 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 evade 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.

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

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 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.

the task. Tasks with e-Gift payment are often undertaken quickly due to no need of monetary investment. Table 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 8: Fraction (%) of tasks with requirements for e-Gift card payment or no credit card payment.

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

Table 9: Fraction (%) of tasks declaring shipment of empty package.

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 9 lists 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 is difficult for online marketplaces to verify the shipment of a package since goods are usually delivered using third-party shipping services.

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 3a 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 elaborately conducted, a fake purchase may still be detected, and the sellers involved may be penalized by Taobao. However, sellers have the right to appeal

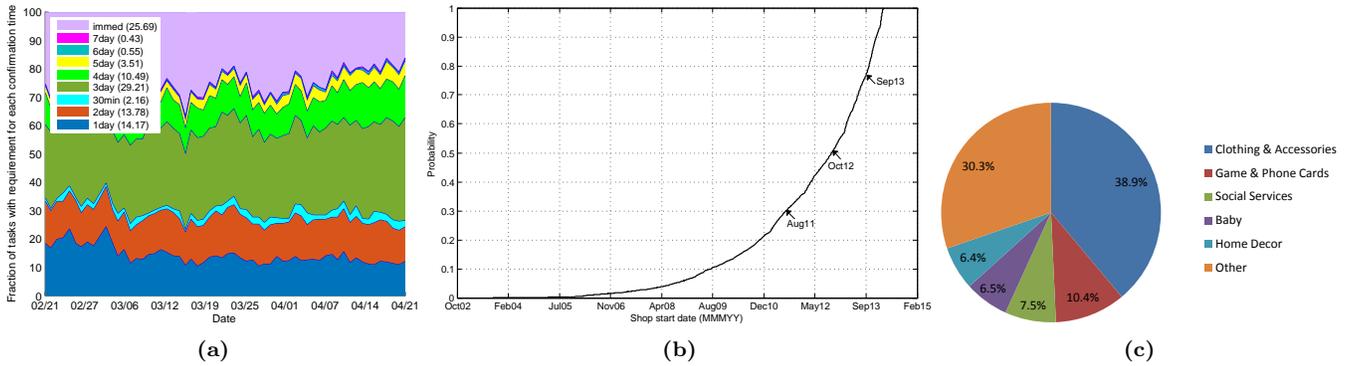


Figure 3: (a) A breakdown of when to confirm the receipt. Numbers in parentheses in the legend denote the fraction of the total 219,165 tasks crawled on the five SRE markets. (b) CDF of shop start date of the identified Taobao sellers. (c) Top 5 main businesses run by the 4,109 Taobao sellers.

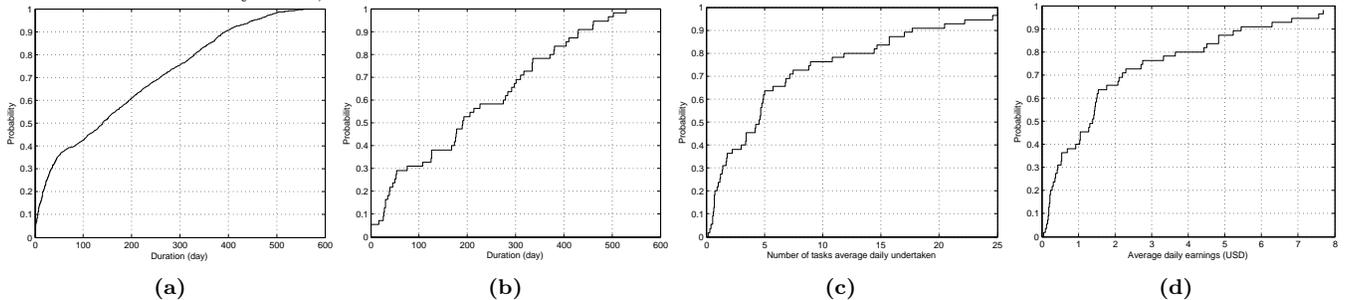


Figure 4: (a) CDF of active duration of sellers on the COOL market. (b) CDF of active duration of the top workers on the COOL market. (c) CDF of average tasks undertaken daily by the top workers on the COOL market. (d) CDF of average daily earnings of the top workers on the COOL market.

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

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 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.

SRE markets’ strategies have turned out to be quite effective in evading Taobao’s detection of fake transactions. Among the 4,109 identified Taobao sellers involved in fake transactions, only 89 (2.2%) sellers were detected and penalized by Taobao, in the form of either having store reputation reduced to zero or having the store forcibly shut down.

4.3 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.

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 3b shows the cumulative distribution function (CDF) of the store’s 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 3c shows the top 5 main businesses, which account 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 the Taobao marketplace. About 10% of stores sell game and phone cards.

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

4a shows the CDF of active durations for the 2,419 sellers on the COOL 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. Table 11 lists the average 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 the most active, with 4.2 tasks posted daily on average per seller, while the sellers on the NET market are the least active, with an average of 2.2 tasks posted daily per seller.

	SKY	WOOD	EMPIRE	COOL	NET
Avg. task #	4.2	2.7	3.1	2.4	2.2

Table 11: Average number of tasks posted daily per seller on each SRE market.

4.4 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 4b 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 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 4c 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 4d 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 [13].

4.5 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 12: Estimated revenue and transaction volume

Based on the two-month crawled data, Table 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 (see Section 5). 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 its estimated annual transaction volume is 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,

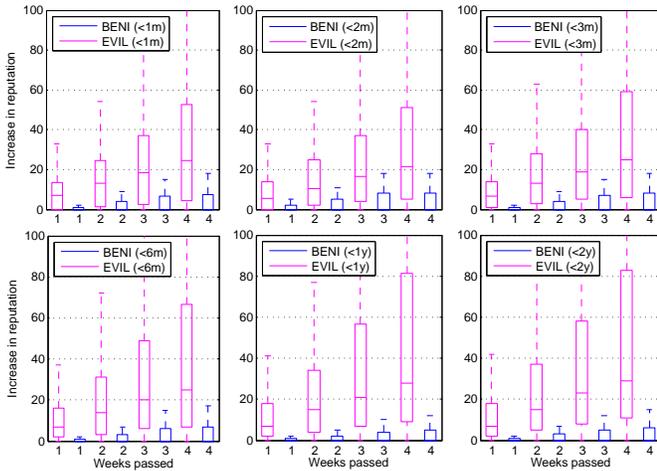


Figure 5: Comparison of the reputation growth distribution between BENI and EVIL stores with varying store ages over the course of one month.

at least two SRE markets have been reported to make off with millions of dollars, and the involved sellers and workers suffered heavy financial losses [14, 15].

5. EFFECTIVENESS OF SRE SERVICES

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

5.1 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 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 for using SRE services, see Figure 3c). We further partition each group into six sets based on store age. Table 13 enumerates the store ages used for partition. We focus on the sellers with store ages not exceeding 2 years because they represent a majority of Taobao sellers on SRE markets (see Figure 3b).

In Figure 5, 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.

5.2 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

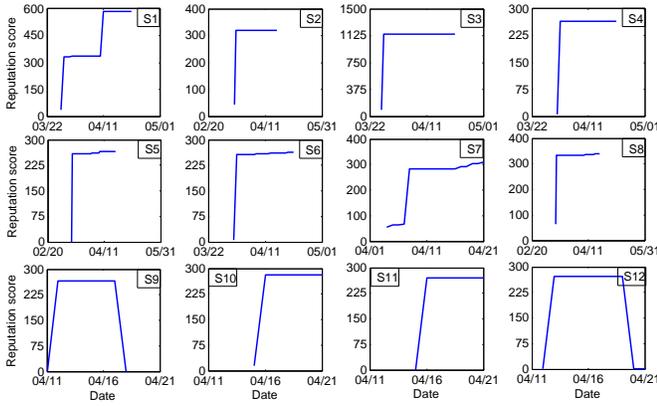


Figure 6: Reputation changes over time for the 12 Taobao sellers identified to use TRUSTEE service.

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

Table 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 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 6 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

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.

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 reputations 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 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.

6. POTENTIAL MITIGATION STRATEGIES

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 denial-of-service attack [16]. 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.

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, including markets for selling fraudulent accounts [35, 36] and Twitter followers [33]. McCoy et al. examined the role of payment processing in the underground economy [26]. Motoyama et al. studied the social dynamics of underground forums [29] and investigated the market for CAPTCHA-solving services [28]. Franklin et al. [19] measured the commoditization of fraudulent activities on an underground market. Christin [18] performed a similar measurement on Silk Road, an anonymous online marketplace. Caballero et al. [17] and Grier et al. [25] studied the pay-per-install market and exploit-as-a-service model for malware distribution. In [20, 27], the authors studied the markets for online pharmaceutical sales. Park et al. [32] leveraged magnetic honeypot ads to study Nigerian scams on Craigslist.

Reputation systems are important to the e-commerce ecosystem. Several works [22–24, 31] investigated the online review manipulation, and some others [21, 30, 34] proposed methods to detect deceptive opinion spam, i.e. fake reviews written to deliberately mislead readers. 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.

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 formulated by SRE markets for evading the online marketplace defenders' detection mechanism of fake transactions. Those tactics are so sophisticated that only 2.2% of illegitimate online sellers were detected and 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. Finally, we 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.

9. REFERENCES

- [1] <http://en.wikipedia.org/wiki/Taobao>.
- [2] <http://blog.ebay.com/ebay-marketplaces-introduces-new-logo/>.
- [3] http://en.wikipedia.org/wiki/Alibaba_Group.
- [4] <http://blogs.wsj.com/chinarealtime/2015/03/03/cat-and-mouse-game-alibaba-exec-on-fake-transactions/>
- [5] http://www.jamaicaobserver.com/business/Alibaba--China-s-Internet-behemoth_17598819.
- [6] <http://www.alexa.com/topsites/global>.
- [7] <http://venturebeat.com/2014/03/31/alibaba-plunks-down-692m-to-push-into-offline-retail/>.
- [8] <http://www2.88sxy.com/>.
- [9] <http://www.ntyjy.com/>.
- [10] <http://www.shuazuanshuaxinyu.com/>.
- [11] <http://www.kus.cc/Index.html>.
- [12] <http://www.shuakewang.com/>.
- [13] <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
- [14] <http://bbs.tianya.cn/post-law-438844-1.shtml>.
- [15] http://weitao.taobao.com/tzh/feed/feed_detail_display.htm?feedId=100505423&wsnsUid=2054965595.
- [16] K. Hoffman, D. Zage, and C. Nita-Rotaru. A survey of attack and defense techniques for reputation systems. In *ACM Computing Surveys (CSUR)*, 2009.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] C. Nosko, and S. Tadelis. The limits of reputation in platform markets: An empirical analysis and field experiment. *Working paper*, 2014
- [24] 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
- [25] 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.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] 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.
- [36] 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.