NOKEScam: Understanding and Rectifying Non-Sense Keywords Spear Scam in Search Engines

Mingxuan Liu[§][*], Yunyi Zhang[†][‡][*], Lijie Wu[‡], Baojun Liu[‡][§][✉], Geng Hong[φ], Yiming Zhang[‡], Hui Jiang[‡][¶], Jia Zhang[‡][Φ], Haixin Duan[‡][Φ][✉], Min Zhang[†], Wei Guan[¶], Fan Shi[†], Min Yang[φ]

[§]*Zhongguancun Laboratory*, [‡]*Tsinghua University*, [†]*National University of Defense Technology*, [φ]*Fudan University*, [Φ]*Quan Cheng Laboratory*, [¶]*Baidu Inc*

Abstract

NOKEScam (**NO**n-sense **KE**yword **S**pear **scam**) is an emerging fraud technique. NOKEScam uses uncommon and usually non-sense keywords (NSKeywords) as vectors to lure victims without complex Black Hat SEO techniques. The obscure NSKeywords ensure the top search results as only NOKEScam pages are exactly matched, misleading victims into trusting them. NOKEScam severely impacts victims and search engines, but its uniqueness has hindered prior research and efficient detection methods.

In this paper, we report on joint work with a leading Chinese search engine to combat NOKEScam. Based on an empirical study, we identified three key observations and developed a lightweight detection system. This system can process about 2 billion URLs within one hour. Over seven months, we identified 153,975 NSKeywords across 68,863 domains. Our measurement demonstrated that leveraging search engine trust endorsement, NOKEScam websites attract an average of over 30k page views daily, indicating significant fraudulent profit potential. Driven by this, attackers persist despite search engine crackdowns, employing evasion tactics like using more domain names. Despite these tactics, our detection system remains effective, significantly suppressing the impact of NOKEScam, with a 194-fold reduction in real-world user complaints. Our findings reveal emerging fraud activities and offer valuable governance lessons for the security community.

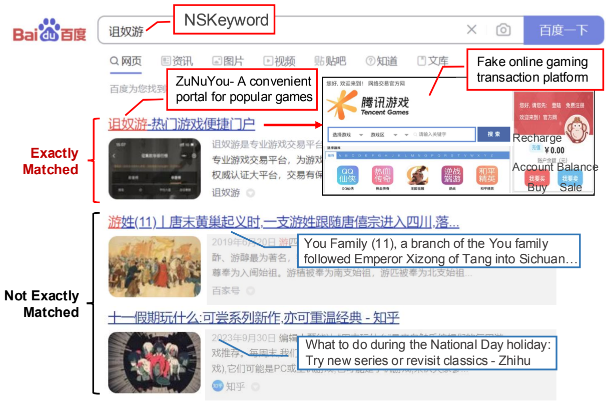

Baidu百度
诅奴游
NSKeyword
百度一下
网页
资讯
图片
视频
贴吧
知道
文库
百度为您找到
ZuNuYou- A convenient portal for popular games
Fake online gaming transaction platform
诅奴游-热门游戏便捷门户
Exactly Matched
腾讯游戏 Tencent Games
Recharge
¥0.00
Account Balance
Buy
Sale
游姓(11) | 唐末黄巢起义时,一支游姓跟随唐僖宗进入四川,落...
You Family (11), a branch of the You family followed Emperor Xizong of Tang into Sichuan…
Not Exactly Matched
十一假期玩什么:可尝系列新作,亦可重温经典 - 知乎
What to do during the National Day holiday: Try new series or revisit classics - Zhihu
知乎

Figure 1: Example of NOKEScam in search engine. Only the top search result exactly matches the NSKeyword, redirecting to a fraudulent online gaming account trading website, while other results do not exactly match.

1 Introduction

Imagine playing an online game when another player offers to buy your account at a high price. But, they insist on using an unfamiliar trading platform. Cautiously, you search for this platform online to verify its legitimacy. As shown in Figure 1, the platform website is listed at the top position in search results as the only exactly matched result. This presentation with a high search ranking eliminates doubts and encourages transaction completion. However, you've unknowingly fallen into a scammer's trap. This scenario illustrates an emerging scam technique called NOKEScam [4].

NOKEScam exploits search engines as trust intermediaries, leveraging "high-search-ranking effects" to induce victims to fraudulent sites. Unlike traditional Black Hat SEO techniques [47, 69], NOKEScam achieves top rankings primarily through NSKeywords, which are uncommon combinations and often *Non-Sense Keywords*. When scammers disseminate NSKeywords instead of links and lure them to search, the NOKEScam page becomes the sole exact match. Furthermore, exact matches consistently rank highest in search engines [28], positioning the NOKEScam page at the top. Unfortunately, victims often perceive top search results as trustworthy [47, 69] and even certified ads [46]. Therefore, by utilizing NSKeywords and search engine credibility, scammers could bypass URL-based detection and enhance their fraud success rates, presenting a significant security threat.

[*] These authors contributed equally to this work.
[✉] Corresponding author.

Research gap. NOKEScam presents unique challenges that prior research fails to address effectively. Unlike traditional Black Hat SEO techniques, NOKEScam adopts unpopular keywords, making existing studies on meaningful and popular keyword-based manipulation detections inapplicable [41, 44, 47, 56, 69]. Furthermore, NOKEScam targets specific users with NSKeywords, making detection methods relying on widespread attributes of links ineffective [13, 19, 36, 51, 83, 92]. Last, the vast amount of search engine data also poses challenges for content-based detection approaches [88]. Consequently, effective detection methods are yet to be developed, leading to a lack of systematic understanding of NOKEScam.

Detecting NOKEScam. In this work, we collaborated with Baidu, China's largest search engine, to combat NOKEScam. Our approach leverages Baidu's user complaint data as a ground-truth dataset. From this, we identified three key insights: obscure NSKeywords are often embedded in webpage titles, each NSKeyword exactly corresponds to a unique webpage, and fraudulent semantics are present in webpage titles. The above observations led us to design a three-step detection system characterized by low computational complexity: 1) identifying NSKeywords based on uncommon character combinations; 2) locating NSKeywords exactly linked to a single webpage; and 3) verifying the presence of fraudulent semantics. Implemented at Baidu, our system actively identifies NOKEScam from newly observed indexed webpages (NOIW). Implemented at Baidu, our system actively identifies NOKEScam from newly indexed webpages (NOIW), processing nearly 2 billion URLs within one hour, demonstrating its effectiveness and efficiency.

Understanding NOKEScam. During the detection period from August 17, 2023, to April 2, 2024, we spotted 153,975 NSKeywords across 68,863 domain names, indicating the scale of NOKEScam. Our findings presented that the implementation of our detection system at Baidu has significantly disrupted the operation of NOKEScam activities. The daily count of detected NSKeywrods has drastically decreased to just a few dozen, a marked reduction from the initial, amounting to 12 times less. User complaints, a key metric, demonstrate the system's efficacy. Complaints decreased dramatically by a factor of 194, from approximately 777 to just 4 per week. Additionally, there have been no reports of false positives where legitimate users are incorrectly blocked, underlining the precision of our detection algorithms.

Through a comprehensive analysis of NOKEScam, we unveil sophisticated tactics. First, we identified phishing behavior to increase credibility, with 2,463 NSKeywords incorporating "Tencent", imitating a famous Internet company. Besides, being indexed by search engines is the first step for the NOKEScam threat. Initially, adversaries employed both passive crawling and active submission to index their scam sites. However, due to our detection efforts, they shifted predominantly towards passive crawling to avoid detection. Furthermore, we discovered that NOKEScam engages in three main illegal businesses - online game transaction fraud (OGTF), pornography, and gambling. Notably, OGTF emerged as the predominant fraud type, accounting for 94.36%. More concernedly, combining Baidu's user complaints and click data, we found NOKEScam's victim impact was severe - assessing complaints revealed an average loss of \$2,896 per NOKEScam victim. Therefore, we estimate that Baidu's active users may experience daily financial losses of NOKEScam as high as \$631,328. By clustering scam sites based on title-domain infra overlap, we observed evasion strategies like switching domains and title structures (e.g. embedding longer benign text). Despite evasion attempts, our detection system remains effective. The significant decrease (194-fold) in user complaints corroborates the system's effectiveness in curtailing NOKEScam proliferation.

Despite successful mitigation efforts within the Baidu search engine, NOKEScam remains prevalent across other major search engines such as Google and Bing. We find that these platforms' recommendation features inadvertently contribute to the spread of NOKEScam by automatically suggesting new NSKeywords. In response, we are proactively engaging with these platforms through responsible risk disclosure.

Contributions. Our main contributions include:

• *Disclosure of NOKEScam.* We disclose the NOKEScam for the first time, an emerging spear scam utilizing search engines. It not only poses significant harm to victims but also severely undermines the authority of search engines.

• *Effective detection method for NOKEScam.* We propose a lightweight detection method for NOKEScam and practically deploy it within the Baidu search engine. Our method significantly mitigates the security threat posed by NOKEScam, with a 194-fold reduction in real-world user complaints.

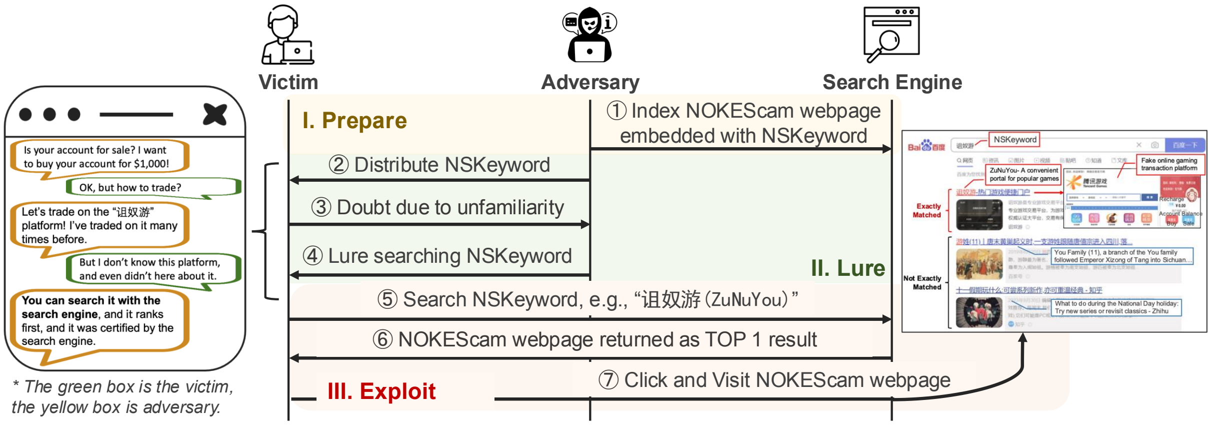
• *Understanding NOKEScam.* We conduct the first comprehensive analysis of the NOKEScam ecosystem, aiding the security community in gaining a deeper understanding of and defending against NOKEScam.

2 Background and Threat Model

In this section, starting from presenting the operational framework of search engines, we introduce the workflow and threat model of NOKEScam.

2.1 Preliminaries of Search Engines

Indexing in Search Engines. Figure 3 illustrates the search engine indexing process. Initially, new website *URLs* must be indexed (❶) to be searchable. This can be done passively ($M1$) by linking to indexed sites and waiting for crawlers, or actively ($M2$) by submitting URLs directly to search engines with an account, expediting crawler discovery [7, 25]. Crawlers parse and extract information (❷) like titles and


Victim
Adversary
Search Engine
I. Prepare
① Index NOKEScam webpage embedded with NSKeyword
② Distribute NSKeyword
③ Doubt due to unfamiliarity
④ Lure searching NSKeyword
II. Lure
⑤ Search NSKeyword, e.g., "诅奴游 (ZuNuYou)"
⑥ NOKEScam webpage returned as TOP 1 result
III. Exploit
⑦ Click and Visit NOKEScam webpage
NSKeyword
ZuNuYou- A convenient portal for popular games
Fake online gaming transaction platform
Exactly Matched
Not Exactly Matched
Recharge
Account Balance
Buy
Sale
You Family (11), a branch of the You family followed Emperor Xizong of Tang into Sichuan...
What to do during the National Day holiday: Try new series or revisit classics - Zhihu

Is your account for sale? I want to buy your account for $1,000!

OK, but how to trade?

Let's trade on the "诅奴游" platform! I've traded on it many times before.

But I don't know this platform, and even didn't here about it.

You can search it with the search engine, and it ranks first, and it was certified by the search engine.

诅奴游

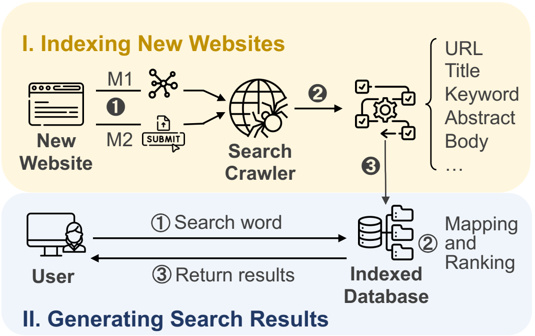
诅奴游-热门游戏便捷门户

游姓(11) | 唐末黄巢起义时,一支游姓跟随唐僖宗进入四川,落...

十一假期玩什么:可尝系列新作,亦可重温经典 - 知乎

* *The green box is the victim, the yellow box is adversary.*

Figure 2: Workflow of NOKEScam.



M1: Crawl passively M2: Submission actively

Figure 3: Indexing process of search engine for a new website and returning search results for users.

content to build indexes, while only limited content is indexed for performance [27]. *Titles* and *Abstracts* (defined by $< description >$ tag [89]) are always indexed, while other content, such as webpage *Bodies*, is selectively indexed (e.g., only indexing content in $< a >$ tags). Low-quality content, including spam, may be excluded [27]. Search engines create *Keyword* lists for each page to build inverted indexes, favoring prominent words from *Titles* and *Abstracts*. In Chinese, additional word segmentation is needed for keyword extraction. The extracted information is stored in an indexed database (❸) as $< URL, Title, Keyword, Abstract, Body, \ldots >$.

Generating search results. As shown in Figure 3, firstly, users enter search words (①), and search engines locate and rank relevant webpages in the indexed database (②). The mapping of relevant webpages begins by checking for exact *Keyword* matches in the indexed database. If no match is found, webpages are fuzzily matched based on page content relevance. Relevant pages are then ranked using algorithms like PageRank [11] and presented to users. While multiple factors influence search results, the exact match between a webpage's *Keywords* and the user's search item is crucial for top ranking. Some search engines also offer an exact search mode that displays only exact matches. Notably, search engine advertisements typically appear as the topmost results [54]. This positioning creates a preconceived notion among victims that the first result is likely a verified advertisement. Finally, ranked search results are returned to the user (③).

2.2 NOKEScam

NOKEScam VS Other Scam. Previously studied scams primarily focused on URL-based promotion methods, such as spear phishing emails [30, 31]. In contrast, NOKEScam utilizes NSKeywords as a promotion medium rather than links, mainly for two reasons. First, to circumvent URL-based detection and potential blocking by chat platforms [80], diverse non-URL promotional media have emerged alongside NSKeywords, including QR codes and images [91]. Second, exploiting search engine credibility was crucial in using NSKeywords for promotion. NOKEScam achieves top search results using NSKeywords without complex Black Hat SEO techniques. Specifically, website names (e.g., yahoo) facilitate quick discovery via search engines, with NSKeywords often serving as site names. These obscure NSKeywords rarely appear elsewhere, typically ranking high as the sole exact match, enhancing perceived credibility. Analysis of user complaints (Section 3) indicates that victims mistakenly view the top-ranked unique result as trustworthy, similar to certified advertising or proprietary content from search engines.

Threat model. The primary goal of NOKEScam is to leverage search engine credibility (high-search-ranking effects) to successfully lure victims to scam pages. This strategy relies on two key conditions. First, miscreants construct uncommon NSKeywords to achieve top search results, misleading victims into perceiving NOKEScam pages as search engine-certified advertisements. Second, NOKEScam pages containing NSKeywords must be indexed by search engines.

Workflow. Figure 2 shows the workflow of NOKEScam, including three steps. First, in the *"Prepare" step*, miscreants create an NSKeyword and embed it in a fraudulent website. To exploit search engine trust, they use diverse ways for indexing (Section 2.1). Notably, the NSKeyword is uncommon, which ensures the fraudulent site is the only exact match. For example, meaningless character combinations (e.g., "诅奴游 (ZuNuYou)") or non-existent words (e.g., "navigrake"). Then, it shifts to the *"Lure" step*, similar to spear-phishing [12, 33, 39, 48], they entice victims using persuasive language, such as offering high prices for game accounts (as left of Figure 2). They suggest conducting transactions on a platform called NSKeyword. When victims express doubt, attackers encourage them to search for it. Finally, in the *"Exploit" step*, once victims search for the NSKeyword, they are exposed to NOKEScam websites. Due to the obscurity, the search engine returns the only exactly matched fraudulent site as the top result. Unfortunately, the high-search-ranking effects mislead victims into believing, visiting, and even trading on the site.

3 Dataset and Empirical Study

As an emerging spear scam, NOKEScam is poorly understood, let alone the publicly available ground-truth dataset. Thus, we cooperated with the Baidu search engine, the most famous Chinese search engine, using its vast search database to gain insights for the detection of NOKEScam. Based on this dataset, we empirically conclude three key observations.

3.1 Real-world Search Engine Dataset

User complaint dataset. To understand NOKEScam for detection, we created a ground-truth dataset from user complaints submitted to Baidu, averaging 1,721 complaints daily, reflecting active user engagement. A dedicated team manually reviews each complaint, with at least two specialists involved before escalation to the relevant security departments. An example of a "NOKEScam-related" complaint is included in Appendix A. Ethical considerations were addressed by anonymizing personal data using MD5 hashing (discussed in Ethics9). Notably, keywords, domains, and links mentioned in manually confirmed user complaints are promptly removed from the search index to prevent ongoing user impact.

NOIW dataset. From the NOKEScam workflow, search engines are optimal detection points, as indexing NOKEScam pages is essential. We collaborate with Baidu to detect NOKEScam pages with their extent index data. Early detection is crucial due to the harmful, time-sensitive nature of NOKEScam. We focus on Newly Observed Indexed Webpages (NOIW)-webpages newly indexed daily. Screening NOIWs allows for preemptive identification and mitigation of malicious sites, reducing potential harm. The NOIW dataset is structured as quadruples: $<URL, Title, Abstract, Body>$, with this information retrieved by Baidu's crawlers. The dataset is vast, averaging 2 billion URLs per hour, and our analysis covers 209 days from August 17, 2023, to April 2, 2024.[1]

3.2 Empirical Study and Key Observation

Ground-truth dataset. For efficient NOKEScam detection, we conducted an empirical analysis. We randomly sampled from a week's worth of user complaints related to NOKEScam (August 17-24, 2023), resulting in a dataset of 2,000 verified NOKEScam webpages containing NSKeywords. User complaints are valuable for two reasons: 1) they reflect real victim experiences, providing an authentic scam nature; 2) the randomness of complaints offers a broad problem overview. Two researchers analyzed this ground-truth dataset, identifying three key observations to aid NOKEScam detection.

- **Observation I: Uncommon NSKeywords are embedded in NOKEScam webpage titles to ensure exposure via search engines.** As noted in Section 2.2, NSKeywords serve as gateways leading victims to NOKEScam pages through search engine queries. To ensure searches for NSKeywords yield NOKEScam pages, NSKeyword must be embedded on these pages. In our ground-truth dataset, 100% of NOKEScam pages embed NSKeywords in their titles. This is likely for two reasons: 1) Titles are always indexed by search engines (Section 2.1), guaranteeing NSKeyword inclusion in search engine indices; 2) Titles are prominently displayed in search results (see Figure 1), ensuring maximum visibility to victims. Additionally, NSKeywords are typically created uncommonly enough to avoid appearing on other sites. Adversaries use rare character combinations to create meaningless strings and combine common words into unusual arrangements. To further obscure recognition, they introduce interfering symbols, like random digits, in fabricated words.
- **Observation II: Search results for NSKeyword contain only one exact match, i.e., NOKEScam page, ranking at the top.** Merely finding the NOKEScam page by searching NSKeywords is insufficient; adversaries exploit high-search-ranking effects strategically. Unlike complex Black Hat SEO techniques [44, 46, 47, 51, 69], attackers in NOKEScam achieve top rankings by exploiting the rarity of NSKeywords. Specifically, NSKeywords are created obscure enough to avoid appearing on other sites. In the ground-truth dataset, we found that NSKeywords yield only one precisely matched result, i.e., NOKEScam page. Despite apparent inefficiency, victim feedback reveals this as a deliberate construct by attackers. Specifically, the unique NOKEScam page matching each NSKeyword typically appears as the top result, mimicking search engine-verified advertisements. This carefully crafted unique top placement easily misleads victims into

[1] Due to maintenance on Baidu's internal servers, there are 3 gaps in data storage, specifically from August 25 to September 7, from September 24 to September 26, and from October 1 to October 6.

perceiving NOKEScam pages as verified ads, significantly enhancing trust and increasing fraud success rates.

- **Observation III: Unique semantic feature.** We found that NOKEScam sites often include explanatory text to inform users of the site's purpose, appearing in two components: 1) Site titles, typically formatted as"NSKeyword-Explanatory", e.g., "诅奴游 (ZuNuYou) - 热门游戏便捷门户 (Popular game trading platform)" (Figure 1); 2) Site abstracts, which appear as brief content overviews. Further analysis of user complaints confirms that explanatory language is crucial for successful deception, as the obscure nature of NSKeywords impedes victim comprehension. Three fraud types were manually identified: online game transaction fraud (first disclosed), gambling, and pornography. Game transaction fraud lures victims via high-priced game account offers in chat windows (exampled in Appendix D). The other two are typical fraud types, while spreading NSKeywords mitigates risks associated with the detection of URL-based promotion.

4 Detection Methodology

Building on three key observations, we develop an efficient and effective detection methodology. Below, we provide details of our detection system designed to identify NOKEScam.

4.1 Detection Goal and Challenges

Detection Goals. To mitigate NOKEScam, intercepting NOKEScam webpage indexing is the optimal vantage point, disrupting its initial phase. Our objective is to exclude fraudulent pages from search engine index databases, preventing scammers from exploiting search engines to mislead victims. We focus detection efforts on the indexing process, identifying and removing NOKEScam URLs from NOIW databases.

Detection Challenges. We face two challenges in detection. First, the sheer volume of data (2 billion webpages per hour) necessitates timely detection, making computationally intensive models impractical. For example, DMOS [88] can process 50 million webpages in five months, which is insufficient for daily billions. Second, NOKEScam's webpages typically appear as normal webpages and exhibit minimal evidence of traditional phishing activities, like mimicking well-known websites or trademarks. Therefore, this prevents NOKEScam pages from being identified through existing analysis of malicious behavior, like phishing detection [31, 49, 75].

4.2 Overview

Considering detection goals and challenges, we propose an efficient and lightweight detection system, as shown in Figure 4.[2] Due to Baidu's vast data volume, a single-step method is impractical. Therefore, our three-step detection method works like a funnel, progressively reducing the number of processed URLs to filter out potential false positives in subsequent steps. First, *NSKeyword Identifier* examines webpage titles for obscure words. It employs contextual perplexity to detect unusual sequences of characters or words, which are then labeled as NSKeywords if identified. Notably, since Baidu primarily serves Chinese users, this study uses Chinese word segmentation to calculate perplexity. However, the concept of detecting uncommon character combinations is applicable to other languages (discussed in Section 7), allowing for broader implementation. This filtering reduces URLs from 2 billion to tens of thousands. While benign sites may use obscure terms, these are essential for constructing NOKEScam. The subsequent steps further filter out false positives. Subsequently, the *Match Relationship Verifier* determines if the association between the NSKeyword and the webpage is unique, i.e., only one webpage (the intentionally crafted scam site) corresponds to the NSKeyword, since benign sites typically have multiple pages. Finally, to mitigate the impact of rare keywords in legitimate websites (e.g., uncommon names or custom site titles), the *Semantic Checker* evaluates fraudulent characteristics based on the semantic content of page titles and abstracts.

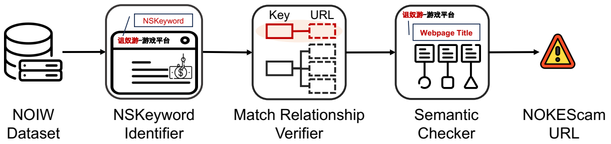


Figure 4: Workflow of our detection system for NOKEScam.

4.3 NSKeyword Identifier

Based on *Observation I*, uncommon NSKeywords are typically embedded in fraudulent website titles. To ensure the rarity of NSKeywords, they are constructed from uncommon character combinations, often resulting in meaningless outputs. Consequently, we find that perplexity in natural language processing intuitively assesses character combination rarity [43]. However, its computational complexity is unsuitable for the scale of billions of URLs per hour.

Given that Baidu mainly serves Chinese users, we employ low-cost traditional word segmentation to estimate word combination likelihood based on perplexity [35], as detailed in Algorithm 1. Segmentation uses prefix dictionaries for efficient graph scanning to build a directed acyclic graph (DAG) of all possible Chinese word combinations in a sentence. Dynamic programming then finds the maximum probability path for the optimal word split based on word frequencies. Thus, segmentation tends to yield highly probable combinations, like "苹果 (apple)". In contrast, isolated Chinese characters

[2] The detection scripts are published on http://nokescam.com, including Chinese and English implementations.

Algorithm 1 : The detail of identifying NSKeywords.

```
Input: Title of webpage ($title$), NSKeyword identification
    threshold ($\theta$).
Output: NSKeyword List ($nskeyword\_list$) in $title$.
$title\_list = segment(title)$    ▷ Segment Webpage Title
$title\_list = Replace\_digit\_english(title\_list)$
$nskeyword\_list = []$
$nskeyword_{tmp} = []$
for each $i \in [0, len(title\_list) - 1]$ do
    if $title\_list[i] == DIGIT$ then
        $title\_list[i] = D$
    else if $title\_list[i] == ENGLISH$ then
        $title\_list[i] = E$
    end if
    $L_i = len(title\_list[i])$
    if $L_i == 1$ then
        $nskeyword_{tmp}.append(L_i)$
    else if $L_i \geq 2$ and $(L_{i+1} == 1$ or $L_{i-1} == 1)$ then
        $nskeyword_{tmp}.append(L_i)$
    else
        $nskeyword_{str} = \text{``''}.join(nskeyword_{tmp})$
        if $len(nskeyword_{str}) \geq \theta$ then
            $nskeyword\_list.append(nskeyword_{tmp})$
            $nskeyword_{tmp} = []$
        end if
    end if
end for
return $nskeyword\_list$;
```

likely cannot combine with neighbors to form common words, indicating obscurity. We thus check the longest consecutive isolated Chinese characters post-segmentation. Considering Chinese language conventions [37] and ground-truth data observations, we set the identification threshold as more than 3 single characters. Thus, if the longest consecutive isolated Chinese characters post-segmentation exceeds 3, we identify obscure words in the webpage title. For example, "诅奴游" is identified as an NSKeyword with three isolated characters, i.e., ["诅 (Zu)", "奴 (Nu)", "游 (You)"].

Additionally, we notice that some NSKeywords contain common Chinese words that may be accidentally combined by scammers when generating them. For example, "狡舒畅 (Pinyin: JiaoShuChang)" where "舒畅 (comfortable)" is a common Chinese word, which is not totally meaningless. Thus, to improve the accuracy of our approach, we determine if a character combination is obscure by checking if the characters flanking common words are single characters. Besides, in ground-truth data, we find that numbers and English characters are sometimes embedded in NSKeywords for obfuscation, e.g. "诅奴2753游". Thus we pre-process all numbers and English characters in titles by replacing them with specific tokens, turning numbers (like "2753") into "DIGIT" and English letters (like "game") into "ENGLISH". For example, the segmentation sequence becomes ["诅", "奴", "DIGIT", "游"]. Treating these specific tokens as an individual character also allows for correctly identifying these cases as NSKeywords.

Finally, this step enables the rapid filtering of suspicious websites and significantly reduces the volume of URLs requiring processing from billions to tens of thousands.

4.4 Match Relationship Verifier

Per *Observation 2*, NSKeywords are crafted to be extremely obscure, ensuring sparse associations with NOKEScam webpages. In ground-truth data, this manifests as unique associations. This uniqueness guarantees that exact NSKeyword matches yield only fraudulent sites, typically ranking first in search results, enhancing victim trust. Consequently, for NSKeywords identified in Step I, we verify if only one precise match exists in the search engine database.

To assess the matching relationship of the indexed websites corresponding to a keyword, the most straightforward approach is to search for the keyword and see how many indexed websites can be precisely matched. Since search engines often employ complex anti-crawler mechanisms [26], excessive requesting frequency can lead to identification by the anti-scraping system, resulting in the blocking of scraping access. Our search engine partner provides us with an internal search result interface to address this issue. Leveraging this interface, we execute queries and retrieve results, bypassing the parsing time required for web rendering, thus offering high query efficiency. The query interface (API) has an average latency of less than 1s and can withstand a concurrency of over 2,500 queries per second (QPS). Specifically, the keyword is sent through the query interface, and the returned fields include the title, abstract, and website information of result entries on the first page of search results for that keyword.

We define association degree as the number of websites where a keyword exactly matches in search results. If keyword k appears in n site titles, its association is n. As precise matches likely appear on the first page (illustrated in Section 2.1), we only assess first-page results. Keywords with an association degree less than or equal to 1 are determined to be NSKeywords ($association_k \leq 1$), and the corresponding websites are filtered as NOKEScam website candidates.

4.5 Semantic Checker

Based on *Observation 3*, NOKEScam pages exhibit distinct semantic characteristics in titles and abstracts (defined by $<$*description*$>$ tag). To avoid misclassifying legitimate cases, we employ semantic differentiation. After reducing candidate URLs by 3 orders of magnitude in previous steps, we can efficiently use a more complex machine learning model. We developed a BERT-based semantic detection model to distinguish normal and NOKEScam pages based on title and abstract semantics [16]. To enable fine-grained ecosystem analysis, we designed a multi-classifier with four categories: normal, online game transaction fraud, gambling, and pornography, based on three fraud types identified in ground-truth data. This classifier both detects NOKEScam pages and facilitates subsequent fraud category analysis.

4.6 Implementation and Evaluation

Implementation. Our detection system uses Python, processing about 2 billion URLs within one hour. It runs on a Baidu internal Linux server, which has 48 cores and 128GB of memory, but no GPU. The *NSKeyword Identifier* employs Jieba[3] for word segmentation. The *Semantic Detector* employs the bert-base-chinese pre-trained model with its default BertTokenizer [34] for embedding, given Baidu's Chinese user focus. To train the multi-classification model, we constructed a comprehensive dataset. To ensure comprehensive coverage, we sourced benign webpages from three categories: 1) Popular websites: selected the top 1k domains with titles from Secrank [86] and Tranco [61] lists separately, yielding 2,000 URLs; 2) Websites from Portal platforms, such as hao123[4], contributing 23,000 URLs; 3) Prevalent news websites, like sina[5] and covering 10 common categories (including finance, stocks, education, etc), providing 10k URLs. Subsequently, all domains were screened using VirusTotal [29], resulting in a final dataset of 35,000 legitimate websites. For malicious webpages, we randomly sampled 45k suspected NOKEScam pages from the *Match Relationship Verifier* results, which were manually categorized by three security researchers. After removing instances whose text length is shorter than 5, we obtained 19k online game transaction fraud websites, 17k pornographic-related websites, and 6k gambling-related websites. The dataset was then randomly split into training, validation, and test sets in an 8:1:1 ratio. The multi-classification model achieved an accuracy of 95.93% and 95.78% on the test and validation sets, respectively. Besides, detected NOKEScam URLs are reported for mitigation actions, namely removing them from Baidu's database, preventing user exposure to NOKEScam sites via NSKeyword searches (details in Section 6.1).

Evaluation. To assess our detection system's performance, we randomly sampled 100 NOKEScam URLs daily across 7 dates from August 17 to December 8, with 12-day intervals, totaling 700 URLs. This ensured that our evaluation spanned the main detection cycle. We comprehensively assess various factors to manually determine if a website is fraudulent. These factors include: 1) Anomalous page content (e.g., mimicking legitimate sites or containing illegal content), observed in 673 pages; 2) Falsified filing information, detected in 84 cases; 3) Anomalous customer service practices, identified in 26 instances; 4) Suspicious payment methods, found in 17 cases.

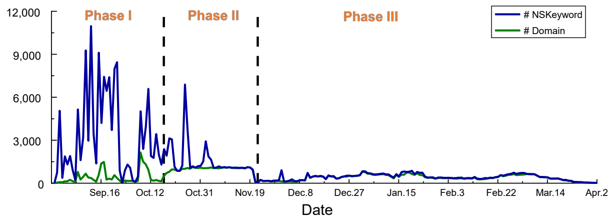


Figure 5: Distribution of the number of detected NSKeywords and related domain names daily.

Our method achieved 99.10% accuracy. Manual examination reveals that False Positives (FPs) predominantly arise from mis-classification in *Semantic Checker* by the multi-classifier, with a 0.96% FP rate. Mis-classifications likely stem from missing abstracts or short titles, limiting learnable semantic content. Despite false positives, the error rate is only 0.9%. Given the average daily detection of 737, the cost of manual verification for FPs remains acceptable. Besides, Baidu's governance requires human verification before NOKEScam link removal, preventing impacts on legitimate websites. Section 6.1 details this process. In real-world detection, false negative rates are usually challenging to assess directly [30, 31]. However, user complaints indirectly reflect detection recall, as undetected websites are reported by users. Section 6 details our analysis of real-world user complaints, revealing a significant decrease (194-fold reduction) post-system deployment, suggesting minimal impact from false negatives in our detection method.

5 Characteristics of NOKEScam

We implemented our detection method within Baidu and conducted daily monitoring of NOIW data from August 17, 2023, to April 2, 2024. Through a comprehensive analysis of detected NOKEScam websites, we uncovered the overall ecosystem of NOKEScam, aiding the security community in enhancing their understanding and defensive capabilities against NOKEScam.

5.1 Overview of NOKEScam

During the seven-month detection period, our method detected **153,975 NSKeywords under 68,863 domain names**, distributed in Figure 5. We identified 226,138 NOKEScam URLs, more than the number of NSKeywords. However, this does not invalidate our Observation II; instead, it reveals an index-retry strategy with reused NSKeywords of scammers (see details in Section 5.3). The seven-month detection period provided clear insights into NOKEScam's activity and tactical evolution. The most significant changes were in the number of domains used and pages deployed per domain. Consequently, we analyze these developments in three phases.

[3] https://github.com/fxsjy/jieba
[4] https://www.hao123.com/
[5] https://www.sina.com.cn/

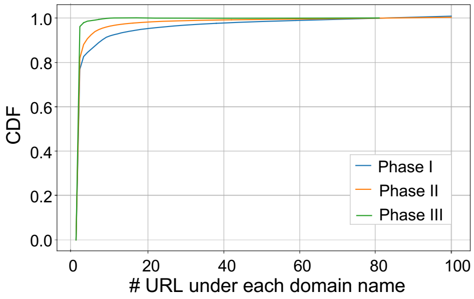


Figure 6: CDF distribution of NOKEScam URLs under each domain.

Phase I: Unawareness stage (August 17 - October 17, 2023). In the initial phase (lasting 43 days), detection had just taken effect and adversaries did not show an obvious response yet. Thus, this period revealed NOKEScam's most primitive strategies with minimal countermeasures - conducting fraud with low domain registration cost, manifested as using a small number of domains but embedding a large number of NSKeywords under each domain. In this phase, we discovered a total of 122,316 NSKeywords, each deployed on a unique webpage across only 13,851 domains. As shown in Figure 6, we found that while only the minimum domains were used in this stage, each domain had the maximum number of configured URLs among 3 phases, with 9 NOKEScam URLs under each domain on average. We found that 2,253 domains contain more than 5 NOKEScam URLs, with the most reaching 4,423 URLs within a single domain name, i.e., *av.ny12306.site*, which hosts primarily pornographic content.

Phase II: Adversarial stage (October 18, 2023-November 22, 2023). In this phase, following our detection, scammers realized something was wrong (their websites were not being indexed) and erroneously guessed the domains used might be problematic. They started switching the domains used in an attempt to circumvent the blocking. This stage lasted 36 days during which we detected 57,506 NSKeywords covering 31,614 domains. Despite not fully grasping our detection logic, adversaries try to evade detection by using more domain names. Specifically, the number of domains used by miscreants was significantly higher than the first stage, at 2.28 times, while the number of NSKeywords was only 0.47 of the first stage's NSKeywords. At this stage, on average only 1.82 NOKEScam URLs were configured under each domain, which is 0.2 times that of Phase I. Our detection system has led to more cautious fraud strategies from NOKEScam, with fraudsters ensuring success by sacrificing the costs of registering more domain names and reducing the number of URLs with NSKeywords configured on each domain.

Phase III: Decline stage (November 23, 2023 - April 2, 2024). At this stage, we observed that adversaries were unable to circumvent detection, with a marked daily decrease in scam activities. This phase, despite lasting 130 days (3 and 3.6 times longer than Phase I and Phase II), detected only 30,263 NSKeywords (0.25 and 0.53 times that of Phase I and Phase II). These were deployed across 24,186 domains, merely 0.77 times the Phase II's count. Besides, the average daily quantity decreased significantly, with an average of only 186 domains and 233 NSKeywords detected per day (12 and 7 times less than Phase I and Phase II). By this stage, despite using more domains, the average number of NOKEScam URLs per domain decreased to only 1.25 on average, and new scam websites could not be indexed by search engines. In this phase, we observed attempts to enhance domain credibility by using domains with Intent Content Provider (ICP) information, yet these did not bypass our detection system. Specifically, in Phase III, we identified 621 domains with ICP, which is double and quadruple the counts from Phase I and II, respectively. Scammers could no longer leverage search engines to spread scam websites to targeted users, let alone gain their trust. Thus, the NOKEScam operation is naturally dismantled in the Baidu search engine. Compared with user complaint data (see Section 6.1), the volume of complaints in the third phase has decreased by a factor of 194 from the onset. Unfortunately, despite the decline in the Baidu search engine, scammers still maintain NOKEScam efforts on other search engines to keep earning steady scam income (see details in Section 6.2).

Overall detection effectiveness. During the detection period, we noted adversaries attempting to bypass our system, like using more domains with ICP information and reducing the number of NOKEScam pages per domain. Despite the higher costs of domain registration, these efforts did not succeed. Our detection relies on identifying uncommon character combinations and their sparse match relationships, which are essential to NOKEScam and difficult to evade. Furthermore, we found that complaints about NOKEScam decreased to one-194th of their original level, demonstrating its real-world resilience against bypass attempts and effectiveness in countering NOKEScam activities (Section 6).

5.2 Infrastructure of NOKEScam

To understand the infrastructure of 68,863 NOKEScam domains, we examined their domain resolution records and registration information. We sent DNS requests for these domains (with `A` records), revealing that they resolved to 9,102 IPs. We also extracted their geographical IP information [53]. For historical resolutions, we accessed an extensive Passive DNS (PDNS) [1] database to gather all resolution records during the detection period, including record types and resolved IPs. Additionally, we obtained WHOIS records [65] for 33,727 NOKEScam-related second-level domains (SLDs) to gather registration information, collecting data for 20,130 SLDs affected by GDPR restrictions [50] and ICP filing information [55] for 1,048 domain names from Baidu. Table 1

Table 1: Top 10 IP Locations and registrars, TLDs of NOKEScam's domains.

Location	# IP	Regsitrar	# SLD	TLD	# SLD
United States	4,718	DNSPod	10,976	.icu	9,904
Seychelles	2,073	Alibaba Cloud	3,936	.com	7,040
Hong Kong, China	503	West263	2,988	.shop	3,233
China	413	Gname	2,352	.cn	3,013
South Africa	358	GoDaddy	626	.asia	2,474
Singapore	343	NameSilo	618	.fun	2,425
Mauritius	323	Juming	611	.top	1,864
Canada	109	West.cn	427	.cfd	564
United Kingdom	77	BANGNING	337	.net	537
Japan	36	Aceville	251	.site	442

presents the top 10 registrars and TLDs for these domains, along with IP location information.

NOKEScam's domain names are registered in bulk and prefer low prices. Scammers tend to register domains for NOKEScam in batches at specific registrars within a short period. Besides, we observed that 74.08% of domains are registered by DNSPod and Alibaba Cloud, both from China. We infer the preferential policies of these two registrars attracted attackers [71]. Furthermore, new TLDs are becoming a major target for abuse by scammers. As shown in Table 1, 6 out of 10 are new gTLDs. Interestingly, `.icu` has overtaken `.com` to become the most popular TLD for scammers. The reason for this phenomenon may be that the new TLD is cheaper [21], which can reduce the cost of NOKEScam.

To circumvent China's domain authentication system, hosting domains on overseas servers is a common strategy. Specifically, 95.46% of NOKEScam IPs are located outside mainland China, predominantly in the United States (51.83%). This strategy likely aims to bypass China's ICP registration requirements, which mandate national licensing for websites providing information services in China, involving content review and operator verification for legal compliance [55]. Notably, ICP information appears on 1,048 (1.52%) NOKEScam websites, which should not pass this review. Manual inspection revealed that this ICP information is outdated, suggesting that scammers may have acquired these domains to create fraudulent websites. By misusing this ICP information, scammers can more easily gain victims' trust for spear scams.

80% of NOKEScam domain names have less than 12 requests per day. In the PDNS data, we observed 20,112 (29.21%) NOKEScam domain resolution records. The resolution volume is generally low, averaging 32 queries per domain, likely due to the covert nature of NOKEScam, with visits primarily from scammers, victims, and bots. Furthermore, 66% of these domains were active for less than 10 days, based on the interval between the first and last observed resolution requests.

Maliciousness of NOKEScam domains is ignored by threat intelligence (TI). To assess the malicious impact of NOKEScam domains, we checked VirusTotal [29] and six open-source domain TI lists, including URLHaus [74], BlackWeb [9], Stopforum Spam [67], Spamlist [68], Dyn Malware Feeds [18], and Zonefiles [95]. Despite the significant financial harm caused by NOKEScam, these domains are rarely flagged as malicious. VirusTotal identified only 612 (0.89%) domains as malicious, with only 115 (0.17%) appearing in open-source TIs. We speculate that NOKEScam's stealthy, targeted nature contributes to its limited exposure in TI, enabling prolonged covert activity.

5.3 Spear Scam Tactics of NOKEScam

We further understand the NOKEScam tactics by analyzing the generating and distributing characteristics of NSKeywords, search engine indexing, scam bait of this spear scam, and the volume of affected users.

How to generate NSKeywords. In total, we detected 153,975 NSKeywords, which is the crucial component of NOKEScam. The average length of NSKeywords is 3 characters (covering 138,448 NSKeywords), which is convenient for user input. We identified a total of 4,158 Chinese characters appearing in all NSKeywords. By analyzing the frequency of these characters in NSKeywords, we discovered a centralization in character selection, with 65 characters occurring in over 500 NSKeywords each. For example, "游" (Pinyin: you, could represent "game") is the most frequent, appearing in 81,030 NSKeywords, followed by "购" (Pinyin: gou, could represent "buy") in 15,354 words. Such a selection strategy arises from scammers' misleading purposes. For example, embedding "游 (representing game)" at the end deceives victims into believing it is a gaming trading platform, like "诅奴游", which is similar to the regular platform.

In addition to the entirely non-existent phrases composed of single characters, 4,745 NSKeywords contain existing Chinese words, with a total of 322 distinct words observed. Most existing words appeared infrequently, averaging 114 NSKeywords each. Among existing words, names of prominent company websites were used more frequently. For instance, we found "腾讯 (Tencent)" appears in 2,463 NSKeywords, typically combined with numbers or letters, to deceive victims into believing these are trustworthy Tencent-affiliated websites, like "腾讯3344网 (Tencent 3344 Website)". Furthermore, we found 20,910 NSKeywords contain digits and letters. Besides confusing victims when combined with famous platform names, 18,406 NSKeywords use the letter "v" and digits to imply version numbers, such as "诅奴游v1.1.25". This strategy not only introduces deception by resembling the real version number but also enhances the randomness of NSKeyword character combinations and enriches scammers' NSKeyword generation.

Moreover, NSKeywords are also reused for *indexing-retry* when scammers discover that they are not indexed. In light of our comprehension of NOKEScam (Section 3.2), it is evident that in order to maintain the similarity of searching results between NOKEScam and certified ads, it will only establish a connection with a single website. However, upon examination of the complete detected NSKeywords dataset, we discovered that some NSKeywords correspond to multiple websites. While they still satisfied a single unique mapping within each day's data. This drew our attention. After further analysis, we found that these websites were posted on different dates and were all within Baidu's governance cycle. Consequently, we infer that the deployment of detection solutions resulted in the direct blocking of NSKeywords from search engines. And scammers reused them after discovering that NSKeywords were not indexed. We conjecture this retry strategy arises on NSKeywords because creating an NSKeyword is harder, as ensuring non-related URLs for an NSKeyword in the huge search engine database is very difficult. In our detection results, we identified 15,931 (10.35%) reused NSKeywords, with an average of 6 times per keyword. The average time interval between retries is 2.39 days. Besides, we found the retry strategy for 14,309 NSKeywords is strict, even frequently changing the domain name of URLs. For example, the porn-related NSKeyword "色鸨阁 (Pinyin: SeBaoGe)" was first detected on September 10, 2023, and indexing attempts continued until April 2024. A total of 106 domains were utilized, with an average retry interval of 1.63 days. These domains have a similar pattern like *[digit1].sebaoge[digit2].top*, where "sebaoge" is the pinyin form of this NSKeyword, *digit1* is a random 3-digit number, and *digit2* is an incrementing 3-digit number, indicating this scammer had bulk registered a batch of domains to keep trying indexing.

In summary, NSKeywords are critical for NOKEScam. Attackers utilize them to enhance the visibility of fraudulent results, thereby increasing scam success rates. However, generating obscure NSKeywords that are not observed on other websites is challenging; thus, attackers reuse unindexed NSKeywords to construct new pages for indexing retry attempts. Additionally, throughout the detection process, despite the various NSKeyword construction techniques employed by attackers, our method for identifying uncommon combinations remains effective.

How to distribute NSKeywords. NOKEScam's targeted nature necessitates stealthy, one-to-one distribution of NSKeywords to specific victims. While analyzing this distribution method is challenging, we strive to provide novel insights from multiple perspectives. Among 2,000 user complaints of ground-truth data, 1,347 cases mentioned how victims obtained the NSKeywords. We found that 64.2% of the cases indicate victims acquired NSKeywords from in-game chats, and 22.3% obtained them from chat platforms like WeChat. The user complaint shows scammers are highly patient in distributing NSKeywords and carefully deploy scams in one-to-one chats. In addition to non-public channels like private conversations, we also sought evidence from public sources, such as Baidu's services. In one day's data from Baidu Knows[6], we only identified 230 posts containing NSKeywords. The scarcity of NSKeywords distribution within the Baidu ecosystem likely aims to preserve the unique mapping between NSKeywords and NOKEScam URLs in Baidu's search engine, as Baidu-generated webpages have a high probability of being indexed. Through the above analysis, we consider that scammers use highly stealthy distribution methods of NOKEScam to maximize the targeted scam success rate, while also trying to contact victims through diverse channels.

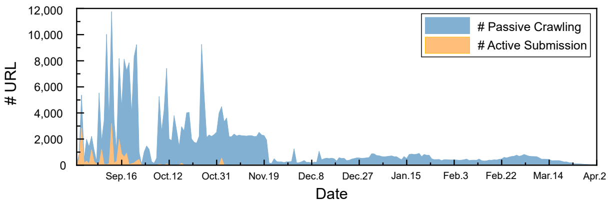


Figure 7: URL distribution for different indexing options.

How to get scam websites indexed by search engines. Baidu records two indexing approaches by logging referrer for passive crawling and user ID for active submission. As shown in Figure 7, we observed that scammers employ diverse and evolving methods to ensure NOKEScam URL indexing. Initially, we found that NOKEScam leverages both indexing avenues. To expedite website indexing, we observed scammers employing active submission techniques, especially in the initial stage. This method covered 8.47% of detected NOKEScam URLs. As noted in Section 2.1, this method requires Baidu accounts and we observed 437 related accounts. Given China's real-name registration policy for online accounts [87], these Baidu accounts used for active submission of NOKEScam pages serve as crucial leads for combating such fraud (see Section 6). Analysis of Baidu account information revealed these accounts were typically registered using false personal data, likely to evade traceability to the fraudsters. Moreover, being passively crawled by crawlers is the main distribution method, concerning 206,984 URLs under 67,128 domains. Figure 7 depicts the shift in scammers' indexing methods from a mixed approach to primarily passive crawling. In Phase I, active submission (33,814 URLs, 985 domains) and passive crawling (96,324 URLs, 12,886 domains) were comparable. Subsequent phases showed almost exclusive reliance on passive crawling. We hypothesize this shift resulted from increased difficulty in active submission due to scammer account suspensions on Baidu.

In addition, to make the indexing process faster and ensure high rankings (ideally the top 1), scammers deploy cloaking techniques to selectively show two different pages on one URL to search engine crawlers versus victims [77, 82, 84]. Specifically, we found scammers display timely content like the latest news to search engine crawlers to expedite indexing. Search engines prioritize the processing of such pages for

[6] Baidu Knows is a free questioning platform.

information freshness. However, when victims visit the URL from search results, scammers display fraudulent webpages. Furthermore, we identified that 3.35% of URLs are still active until April 2024. Examination of accessible pages unveiled scammers' cautious tactics in masking fraudulent activities. We identified 467 URLs utilizing redirection techniques for detection evasion, with some requiring specific user interactions (e.g., clicking a button) to access fraudulent content. Notably, 16.95% of active websites offered download links to fraudulent mobile applications, directing victims to specific fraudulent APPs.

What bait is used to deceive victims. To discern the illicit business categories of NOKEScam, we employed multiple classifiers from the *Semantic Detector* (in Section 4) to categorize NOKEScam pages. NOKEScam primarily involves online game account transactions (94.36%), with a smaller number related to pornography (4.81%) and gambling (0.83%). Notably, within this complex targeted fraud scheme, we identified a novel fraud type: online game transaction fraud (OGTF). OGTF provides an excellent scenario for NOKEScam, since chat tools in online games provide scammers with a convenient way to find targets. The proliferation of online game account trading platforms, exemplified by PlayerAuctions [60], stems from spontaneous player-to-player transactions. The abundance of these platforms complicates users' ability to discern the legitimacy of unfamiliar trading sites [59]. This uncertainty in the game trading ecosystem creates opportune exploitation points for NOKEScam operations. Moreover, to attract victims, scammers also embed popular game names in their NOKEScam webpages. By extracting embedded games in NOKEScam webpages, we found that the top 2 popular games in NOKEScam are Lost Ark from Amazon and Smilegate's famous free-to-play action MMO game, appearing on 6,874 websites. Furthermore, user complaint records help us to reconstruct how the scammers induced the victims to enter the fake trading platform and forced the victims to recharge step by step (see Appendix D).

Despite gambling and pornography being common underground industries known for broad promotion, they exhibit more severe potential harm to users in NOKEScam. Analysis of related NSKeywords and page titles reveals use of obscure terms mimicking gambling games or unique pornographic content names, likely to evade chat platform detection. This indirect method guides victims to underground sites while avoiding direct blocking. Examination of page analysis indicates covert profit mechanisms in gambling and pornography sectors, by enticing victims with free trials while failing to deliver promised services. In gambling scenarios, attackers utilize demo games to entice victims to deposit funds without providing complete games and may also promote fraudulent gambling apps [32]. In adult scenarios, they use incomplete videos to tempt victims into making payments without delivering full content.

How many users were affected. NOKEScam resulted in serious consequences in the wild that caused direct financial harm to victims. We assess NOKEScam's user impact using Baidu's search logs. Specifically, we collect page view (PV) data for detected NOKEScam-related domain names, measuring user visits originating from search engine queries and subsequent clicks. Statistics revealed an average daily PV of about 30k for these fraudulent domains, a significant figure as each PV potentially represents a victim accessing a scam website. Moreover, the substantial PV indicates that these domains were not accidentally accessed. Typically, unintended search results receive fewer clicks, as they don't match users' search intent, resulting in lower PV counts. Furthermore, we estimated the severity of victims' financial losses using victim complaint records. By examining the last 100 complaint reports, we find 20 complaints mentioned the amount of money defrauded, ranging from 100 to 33,443 dollars. The average fraud amount per NOKEScam victim was $2,896, which could be used to estimate potential losses through the Baidu search engine. Furthermore, Baidu recorded around 218 million daily active users [70]. Assuming a rate of one-millionth of users falling victim to NOKEScam, approximately 218 people are defrauded daily, resulting in losses of up to $631,328.

5.4 Campaign Analysis of NOKEScam

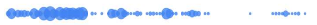
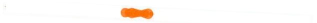
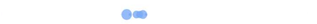
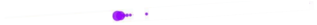
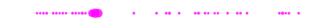
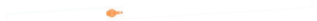
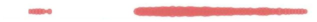
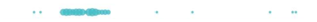
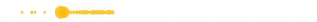
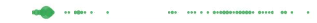
Beyond the macro-level insights, we developed a clustering algorithm to analyze the scam strategies of NOKEScam campaigns from a micro-perspective.

Cluster NOKEScam campaigns. We designed a clustering methodology for NOKEScam based on two assumptions: 1) websites with similar title patterns, and 2) websites using the same domain registration information, belong to the same campaign. Specifically, first, we clustered titles with a Levenshtein distance [93] of less than three into groups. Then, we merged clusters with the same domains or registrant information.

Overview of Campaigns. In total, we identified 143 NOKEScam campaigns. NOKEScam exhibits a pronounced head effect, with the top 10 campaigns accounting for 80.02% of NSkeywords. Table 2 shows the top 10 campaigns. The largest campaign used 104,204 URLs with unique NSKeywords and was active in 68 days. Online gaming fraud accounts for 21.0% of campaigns and has larger groups than other business types, with 8 of the top 10 in this business. While 95% of campaigns focus on one underground industry, 5% engage in multiple businesses. For example, campaign 7 involves both online game fraud and gambling, likely due to similarities between gambling and gaming platforms.

Activity of NOKEScam Campaigns. We calculated campaign active days by summing the days when the NOKEScam URL was observed. Most campaigns have short active days, with 67.83% active in less than 7 days. However, some short-lived campaigns have many NSKeywords, indicating high activity. For example, campaign 3 was active for just 6 days

Table 2: Information of Top 10 NOKEScam campaigns.

Cam.	Business	# Domain	# NSKeyword	Active Time (Aug.17, 2023 - Dec.31, 2023*)	# Active Day
1	Online game transaction fraud	8,898	104,204		68
2	Online game transaction fraud	532	8,316		9
3	Online game transaction fraud	5	2,557		6
4	Online game transaction fraud	153	2,378		6
5	Online game transaction fraud	5	2,158		37
6	Online game transaction fraud	4	1,422		5
7	Online game transaction fraud & Gambling	136	746		60
8	Pornography & Gambling	387	513		24
9	Online game transaction fraud	87	489		21
10	Pornography	391	433		45

* : Since the crackdown from detection, these groups had no active information in 2024. Therefore, we set the cutoff date for the plots as December 31, 2023.

but generated up to 426 websites daily. Longer campaigns often use many domains with frequent changes and short periods per domain. Campaign 10 used 391 domains, and each domain was active for just 1.16 days on average, with the shortest domain active for only 1 day. We speculate this domain rotation is an evasion strategy to avoid detection.

Evasion Strategy. During governance, we observed adversaries probing and attempting to evade our detection. Initially, campaign 1 used 253 domains with an average of 16 pages each. However, they later abandoned these domains, introduced 34 new ones, and reduced the average pages per domain to 8. Besides, modified page titles were observed, as seen in campaign 1. They added longer normal text, such as news or encyclopedia content, to obscure malicious semantics. Additionally, they included meaningless punctuation, numbers, or letters around the NSKeyword. Evasion attempts can still be effectively detected. Although attackers lengthen titles and add distracting characters, NSKeywords enable the *NSKeyword Identifier* to detect anomalies. Additionally, semantic conflicts between fraudulent and normal text can be identified by the *Semantic Checker*. Thus, the detection system remains effective due to its integration of three key observations. This resilience is evidenced by user complaints analysis in Section 6, demonstrating minimal impact on the system's performance.

6 Real-World Impact

6.1 NOKEScam in Baidu

Our detection system is now stably deployed at Baidu, daily scanning NOIW data. Detected NOKEScam URLs are reported to Baidu's governance department for blocking, thus preventing user exposure to malicious content. Specifically,

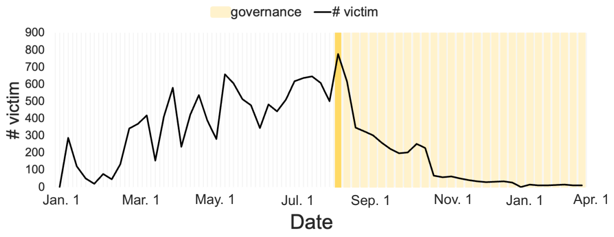


Figure 8: Number of user complaints received by Baidu from January 2023 to April 2024. The yellow area on the right is the post-government feedback.

Baidu extracts the domain names from detected NOKEScam URLs, ensuring that all URLs under that domain are not indexed in the NOIW data and are removed from the indexed database. This enables swift blocking of identified NOKEScam domains. Notably, Baidu manually confirms detection results before removal to prevent false positives affecting normal users. The absence of user complaints concerning erroneous blocking of legitimate websites indicates minimal impact from false positives.

Blocking NOKEScam websites from reaching victims. From real user feedback, our lightweight detection system on NOIW significantly contained NOKEScam spread on Baidu. Figure 8 shows the number of Baidu user complaints about encountering NOKEScam over time. Yellow areas indicate changes after deploying our approach. Before governance, Baidu received 384 victim complaints per week on average. After blocking the indexing of scam sites, adversaries could not add new ones. Thus, complaints dropped precipitously, decreasing 194-fold from the most 777 to 4 per week, reaching an average of 6 per week. Despite minimal weekly user complaints, an in-depth analysis of these cases reveals attackers' evasion attempts primarily focus on obfuscating titles and summaries. As shown in the second case in Appendix B, it embeds benign text far exceeding malicious content within titles. While embedding extensive benign text in titles may bypass our detection system, user feedback indicates high anomaly recognition in such cases. This reduces fraud success rates, making it an unfavorable evasion tactic for attackers. In summary, user complaints significantly reflect real-world conditions. The sharp decline in such complaints indicates a substantial reduction in NOKEScam activities.

Combating with scammers. As outlined in Section 5.3, Baidu logs indexing sources for each URL. Active submission requires Baidu accounts, enabling the tracking of NOKEScam URL submitters. Our detection identified 437 highly suspicious accounts, reported to Baidu. Through account interaction analysis, Baidu expanded this to 2,335 scam accounts. These accounts were then managed by removing their URLs and revoking submission privileges. This account tracking and governance process forms a crucial component in effectively combating NOKEScam.

6.2 NOKEScam in other Search Engines

While our study focuses on Baidu, NOKEScam represents a general scam approach. To demonstrate its pervasiveness, we assessed its presence on other major search engines: Google, Bing, 360 Search, and Sougou. Our investigation reveals NOKEScam's prevalence across all major search engines, highlighting a significant industry-wide challenge.

First, we actively created a non-existent English word, "Navigrake", to avoid current results focusing solely on Chinese. We configured a webpage for it and, by submitting this page to four search engines without any SEO efforts, achieved the top (and only) ranking for searching "Navigrake". This experiment demonstrates the potential for abuse of NOKEScam by these search engines, as adversaries could easily exploit such "high-search-ranking effects" for fraud.

Subsequently, we attempted to leverage Baidu's detection results to validate the presence of NOKEScam in other search engines. Using 500 randomly selected NSKeywords, we analyzed the top 5 search results from each engine, employing our NSKeyword identification method, followed by manual verification. Despite the absence of any previously identified NSKeywords from other search engines, we discovered 653 new NSKeywords through the search engine's search term recommendations. Considering that adversaries can actively submit indexing requests to search engines, we hypothesize that scammers use distinct NSKeywords for each search engine to more precisely attract victims. In addition, we found 445 posts or reports about NOKEScam in Bing and 360 Search. Victims stated how they were deceived or applied for legal aid, which also proves the widespread impact of NOKEScam in the wild. Although crawling restrictions impede external assessment of NOKEScam's scope, our validation using Baidu's partial detection results still confirms NOKEScam's impact across other search engines. We've informed affected search engines and will offer assistance in mitigating this threat.

7 Discussion

Limitations. Despite effectively combating NOKEScam on Baidu, analyzing a single search engine has limitations. However, as the largest Chinese search engine, Baidu-specific detection still provides an overview of this novel scam and confirms NOKEScam severity. Although a comprehensive evaluation of other search engines is challenging, we still assessed the presence of NOKEScam on them based on our detection results in Baidu (Section 6.2). We also disclosed results to affected search engines and will actively assist them in addressing this threat with our open-sourced detection methodology. Since Baidu mainly serves Chinese users, our detection method implemented is primarily for Chinese text (especially for identifying NSKeywords). However, the overall approach utilizes general perplexity to calculate character combination probabilities, making it language-agnostic. Although we lack real-world observation on other languages (e.g., English), we have implemented a detection method for uncommon English words and phrases, with the code open-sourced and detailed in Appendix C. Additionally, our detection system's significant combating impact on NOKEScam, with only minimal errors persisting. Regarding false negatives, user complaints consistently declined over seven months post-deployment, reducing to merely four per week. This demonstrates our detection system's robust performance, effectively countering attackers' evasion attempts. The evaluation shows a low false positive rate (0.9%), with Baidu implementing manual verification before blocking detected URLs. The absence of user complaints regarding erroneous blocking of legitimate websites to date indicates that false positive impacts are manageable.

8 Related Works

Spear phishing. Previous research on spear phishing has focused on several common scenarios, including email, social medias, telephone, and SMS [8, 39]. Email is a typical context for spear phishing [12, 33, 39], with two main detection aspects. On one hand, some works detect spear phishing emails based on their impersonation behavior [31] and other works detect them based on compromised email accounts [15, 30, 57]. In addition, some works detect spear phishing emails from multiple dimensions, including detection based on URL component features [10, 14, 23, 75], detection based on webpage content [2, 45, 49, 81, 85, 94], and detection based on textual features of the emails (such as the subject and body content) [3, 17, 20, 24, 90]. Social media-related spear phishing studies propose tracking methods for compromised accounts [66, 72]. Considering the vast user base of telecommunications for phone calls and SMS, telecom services are also a major context for spear phishing. Besides, existing works conduct in-depth analyses of spear phishing scams via telephone [62–64, 73], and methods for detecting and measuring SMS-based spear phishing attacks [48, 58].

BlackHat SEO. Numerous studies demonstrate that search engine results may not be entirely reliable, as they may contain illicit promotion information [19, 42, 44, 52, 76]. The primary goal of illegal promotion is to attract more users of illicit products, including but not limited to gambling sites, pornographic services, and illegal drugs. Additionally, existing research has confirmed that high-ranking effects in search engines have been exploited in phishing fraud, such as tech support scams [47, 69] and fake search ad scam [46]. Existing studies identified that perpetrators of BlackHat SEO frequently target popular keywords [40, 41, 56], since these keywords have substantial user traffic for effective promotion. Multiple functions of search engines are exploited in BlackHat SEO practices, such as automatic search recommendation [79] and nearby business search services [78]. Since BlackHat SEO injects a substantial volume of illicit promotional content, search engines face significant harm. Consequently, existing studies proposed various detection methods, including detecting spider pools based on the identification of wildcard DNS [19], detection based on link relationships [13, 36, 51, 83, 92], and content-based detection [88]. *In this work, we first reveal a novel spear scam activity, NOKEScam.* Unlike other fraud methods, NOKEScam spreads obscure NSKeywords rather than fraudulent links. These unique keywords ensure top search engine results without using any complex Black Hat SEO techniques, exploiting user trust by misleading victims to perceive results as legitimate (e.g., certified ads). Despite challenges in detecting this novel fraud in vast search engine data, we collaborated with prominent search engine providers to develop effective detection methods. Real-world results demonstrate successful mitigation of this emerging fraudulent behavior.

9 Conclusion

This paper reveals a novel spear scam, the uncommon and usually non-sense keyword spear scam, NOKEScam. It lures victims into searching for obscure NSKeywords, leading them to a unique matching website at the top of search results, which exploits the endorsement of search engines to convince victims of the site's legitimacy. In partnership with Baidu, China's leading search engine, we conducted the first detection and measurement study of NOKEScam. Our deployed detection system identified 153,975 NSKeywords over a seven months. Based on our findings, we provide the first comprehensive analysis of the NOKEScam ecosystem, encompassing tendencies in domain infrastructure usage, evolution of indexing strategies and title generation mechanisms. Notably, we discovered a novel fraud category, online game transaction fraud (94.36%). Besides, our detection system's real-world impact is substantial, reducing NOKEScam-related user complaints 194-fold.

Ethics

Our analysis strictly followed ethical guidelines, namely the Belmont Report [22] and the Menlo Report [38]. Potential privacy risks relate to three data types: user complaints, submission accounts, and domain registration. First, the user complaint dataset may contain personal user information, despite Baidu's explicit privacy protection statement [5, 6]. Therefore, before analysis, Baidu manually reviewed the data and anonymized any personal information by computing MD5 hashes. Thus our complaint analysis does not involve personal data. For submission accounts, we provided only scam URLs to the search engine for handling, without accessing account info. For other data, we only used publicly available information without ethical issues, like URLs and page content. For domain registration, we relied on WHOIS data but could not obtain all registrant info due to GDPR, limiting governance. Overall, we took measures to protect privacy and limit analysis to public data. Finally, we will anonymize our open-source detection results to mitigate ethical risks.

Open Science Considerations

We released our detection script and detected data on http://nokescam.com. However, as our detection data source (Baidu's actual indexing data) cannot be disclosed, thus a direct evaluation of the detection script's functionality and reproducibility is not feasible. To address this limitation and provide comprehensive awareness to the security community, we published all detection results, including NSKeywords and NOKEScam URLs. Additionally, our method is universally applicable to search engine providers. We responsibly disclosed findings from other search engines to related vendors and assisted in implementing our detection system to comprehensively mitigate this security risk.

Acknowledgments

We sincerely thank all anonymous reviewers and our shepherd for their valuable comments on improving the paper. This work is in part supported by the National Natural Science Foundation of China (62102218), the National Key Research and Development Program of China (No. 2023YFC3321303). Haixin Duan is supported by the Taishan Scholars Program.

References

[1] 114 DNS. 114 DNS. https://www.114dns.com/, 2024.

[2] Sahar Abdelnabi, Katharina Krombholz, and Mario Fritz. Visualphishnet: Zero-day phishing website detection by visual similarity. In Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna, editors, *CCS '20: 2020 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, USA, November 9-13, 2020*, pages 1681–1698. ACM, 2020.

[3] Shivam Aggarwal, Vishal Kumar, and S. D. Sudarsan. Identification and detection of phishing emails using natural language processing techniques. In Ron Poet and Muttukrishnan Rajarajan, editors, *Proceedings of the 7th International Conference on Security of Information and Networks, Glasgow, Scotland, UK, September 9-11, 2014*, page 217. ACM, 2014.

[4] Baidu. Anti-fraud Lesson - Counterfeit game trading platforms. https://110.baidu.com/fqz/page/classroom/79, 2023.

[5] Baidu. Baidu Complaint Handling Statement. https://tousu.baidu.com/statement, 2024.

[6] Baidu. Baidu Privacy Protection Statement. https://www.baidu.com/duty/yinsiquan.html, 2024.

[7] Baidu Search Engine. Baidu Link submission. https://ziyuan.baidu.com/linksubmit/url, 2023.

[8] Zinaida Benenson, Freya Gassmann, and Robert Landwirth. Unpacking spear phishing susceptibility. In Michael Brenner, Kurt Rohloff, Joseph Bonneau, Andrew Miller, Peter Y. A. Ryan, Vanessa Teague, Andrea Bracciali, Massimiliano Sala, Federico Pintore, and Markus Jakobsson, editors, *Financial Cryptography and Data Security - FC 2017 International Workshops, WAHC, BITCOIN, VOTING, WTSC, and TA, Sliema, Malta, April 7, 2017, Revised Selected Papers*, volume 10323 of *Lecture Notes in Computer Science*, pages 610–627. Springer, 2017.

[9] BlackWeb. Blackweb. https://github.com/maravento/blackweb/, 2023.

[10] Aaron Blum, Brad Wardman, Thamar Solorio, and Gary Warner. Lexical feature based phishing URL detection using online learning. In Rachel Greenstadt, editor, *Proceedings of the 3rd ACM Workshop on Security and Artificial Intelligence, AISec 2010, Chicago, Illinois, USA, October 8, 2010*, pages 54–60. ACM, 2010.

[11] Sergey Brin and Lawrence Page. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Comput. Networks*, 56(18):3825–3833, 2012.

[12] Deanna D. Caputo, Shari Lawrence Pfleeger, Jesse D. Freeman, and M. Eric Johnson. Going spear phishing: Exploring embedded training and awareness. *IEEE Secur. Priv.*, 12(1):28–38, 2014.

[13] Carlos Castillo, Debora Donato, Aristides Gionis, Vanessa Murdock, and Fabrizio Silvestri. Know your neighbors: web spam detection using the web topology. In Wessel Kraaij, Arjen P. de Vries, Charles L. A. Clarke, Norbert Fuhr, and Noriko Kando, editors, *SIGIR 2007: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, July 23-27, 2007*, pages 423–430. ACM, 2007.

[14] Juan Chen and Chuanxiong Guo. Online detection and prevention of phishing attacks. In *2006 First International Conference on Communications and Networking in China*, pages 1–7. IEEE, 2006.

[15] Asaf Cidon, Lior Gavish, Itay Bleier, Nadia Korshun, Marco Schweighauser, and Alexey Tsitkin. High precision detection of business email compromise. In Nadia Heninger and Patrick Traynor, editors, *28th USENIX Security Symposium, USENIX Security 2019, Santa Clara, CA, USA, August 14-16, 2019*, pages 1291–1307. USENIX Association, 2019.

[16] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019.

[17] Prateek Dewan, Anand Kashyap, and Ponnurangam Kumaraguru. Analyzing social and stylometric features to identify spear phishing emails. In *2014 APWG Symposium on Electronic Crime Research, eCrime 2014, Birmingham, AL, USA, September 23-25, 2014*, pages 1–13. IEEE, 2014.

[18] Dyn DNS. Dyn dns. http://security-research.dyndns.org/pub/malware-feeds/, 2023.

[19] Kun Du, Hao Yang, Zhou Li, Hai-Xin Duan, and Kehuan Zhang. The ever-changing labyrinth: A large-scale analysis of wildcard DNS powered blackhat SEO. In Thorsten Holz and Stefan Savage, editors, *25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016*. USENIX Association, 2016.

[20] Sevtap Duman, Kubra Kalkan-Cakmakci, Manuel Egele, William K. Robertson, and Engin Kirda. Emailprofiler: Spearphishing filtering with header and stylometric features of emails. In *40th IEEE Annual Computer Software and Applications Conference, COMPSAC 2016, Atlanta, GA, USA, June 10-14, 2016*, pages 408–416. IEEE Computer Society, 2016.

[21] Dynadot. Domain Prices. https://www.dynadot.com/domain/tlds-prices?price_level=0, 2024.

[22] United States. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont report: ethical principles and guidelines for the protection of human subjects of research*. Department of Health, Education and Welfare, 1979.

[23] Sujata Garera, Niels Provos, Monica Chew, and Aviel D Rubin. A framework for detection and measurement of phishing attacks. In *Proceedings of the 2007 ACM workshop on Recurring malcode*, pages 1–8, 2007.

[24] Hugo Gascon, Steffen Ullrich, Benjamin Stritter, and Konrad Rieck. Reading between the lines: Content-agnostic detection of spear-phishing emails. In Michael D. Bailey, Thorsten Holz, Manolis Stamatogiannakis, and Sotiris Ioannidis, editors, *Research in Attacks, Intrusions, and Defenses - 21st International Symposium, RAID 2018, Heraklion, Crete, Greece, September 10-12, 2018, Proceedings*. Springer, 2018.

[25] Google. Google Link Submission. https://www.google.com/webmasters/tools/submit-url, 2023.

[26] Google. Limits and Quotas on API Requests. https://developers.google.com/analytics/devguides/config/mgmt/v3/limits-quotas?, 2023.

[27] Google. How Google Search Work. https://www.google.com/search/howsearchworks/how-search-works, 2024.

[28] Google Search Engine. What is fuzzy search? https://cloud.google.com/discover/what-is-fuzzy-search?hl=en, 2024.

[29] Hispasec Sistemas Company. Virus Total. https://www.virustotal.com/gui/home/search. (Access in October, 2021).

[30] Grant Ho, Asaf Cidon, Lior Gavish, Marco Schweighauser, Vern Paxson, Stefan Savage, Geoffrey M. Voelker, and David A. Wagner. Detecting and characterizing lateral phishing at scale. In Nadia Heninger and Patrick Traynor, editors, *28th USENIX Security Symposium, USENIX Security 2019, Santa Clara, CA, USA, August 14-16, 2019*. USENIX Association, 2019.

[31] Grant Ho, Aashish Sharma, Mobin Javed, Vern Paxson, and David A. Wagner. Detecting credential spearphishing in enterprise settings. In Engin Kirda and Thomas Ristenpart, editors, *26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017*, pages 469–485. USENIX Association, 2017.

[32] Geng Hong, Zhemin Yang, Sen Yang, Xiaojing Liao, Xiaolin Du, Min Yang, and Haixin Duan. Analyzing ground-truth data of mobile gambling scams. In *43rd IEEE Symposium on Security and Privacy, SP 2022, San Francisco, CA, USA, May 22-26, 2022*, pages 2176–2193. IEEE, 2022.

[33] Hang Hu and Gang Wang. End-to-end measurements of email spoofing attacks. In William Enck and Adrienne Porter Felt, editors, *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, pages 1095–1112. USENIX Association, 2018.

[34] Hugging Face. bert-base-chinese. https://huggingface.co/bert-base-chinese, 2023.

[35] Zhenyu Jiao, Shuqi Sun, and Ke Sun. Chinese lexical analysis with deep bi-gru-crf network. *CoRR*, abs/1807.01882, 2018.

[36] John P. John, Fang Yu, Yinglian Xie, Arvind Krishnamurthy, and Martín Abadi. deseo: Combating search-result poisoning. In *20th USENIX Security Symposium, San Francisco, CA, USA, August 8-12, 2011, Proceedings*. USENIX Association, 2011.

[37] Matthew Joslin, Neng Li, Shuang Hao, Minhue Xue, and Haojin Zhu. Measuring and analyzing search engine poisoning of linguistic collisions. In *2019 IEEE Symposium on Security and Privacy, SP 2019, San Francisco, CA, USA, May 19-23, 2019*. IEEE, 2019.

[38] Erin Kenneally and David Dittrich. The menlo report: Ethical principles guiding information and communication technology research. *Available at SSRN 2445102*, 2012.

[39] Daniele Lain, Kari Kostiainen, and Srdjan Capkun. Phishing in organizations: Findings from a large-scale and long-term study. In *43rd IEEE Symposium on Security and Privacy, SP 2022, San Francisco, CA, USA, May 22-26, 2022*, pages 842–859. IEEE, 2022.

[40] Nektarios Leontiadis, Tyler Moore, and Nicolas Christin. Measuring and analyzing search-redirection attacks in the illicit online prescription drug trade. In *20th USENIX Security Symposium, San Francisco, CA, USA, August 8-12, 2011, Proceedings*. USENIX Association, 2011.

[41] Nektarios Leontiadis, Tyler Moore, and Nicolas Christin. A nearly four-year longitudinal study of search-engine poisoning. In Gail-Joon Ahn, Moti Yung, and Ninghui Li, editors, *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, AZ, USA, November 3-7, 2014*, pages 930–941. ACM, 2014.

[42] Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, Damon McCoy, Nicholas Weaver, Vern Paxson, Geoffrey M. Voelker, and Stefan Savage. Click trajectories: End-to-end analysis of the spam value chain. In *32nd IEEE Symposium on Security and Privacy, SP 2011, 22-25 May 2011, Berkeley, California, USA*, pages 431–446. IEEE Computer Society, 2011.

[43] Dianqi Li, Yizhe Zhang, Hao Peng, Liqun Chen, Chris Brockett, Ming-Ting Sun, and Bill Dolan. Contextualized perturbation for textual adversarial attack. In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tür, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 5053–5069. Association for Computational Linguistics, 2021.

[44] Xiaojing Liao, Chang Liu, Damon McCoy, Elaine Shi, Shuang Hao, and Raheem A. Beyah. Characterizing long-tail SEO spam on cloud web hosting services. In Jacqueline Bourdeau, Jim Hendler, Roger Nkambou, Ian Horrocks, and Ben Y. Zhao, editors, *Proceedings of the 25th International Conference on World Wide Web, WWW 2016, Montreal, Canada, April 11 - 15, 2016*, pages 321–332. ACM, 2016.

[45] Yun Lin, Ruofan Liu, Dinil Mon Divakaran, Jun Yang Ng, Qing Zhou Chan, Yiwen Lu, Yuxuan Si, Fan Zhang, and Jin Song Dong. Phishpedia: A hybrid deep learning based approach to visually identify phishing webpages. In Michael D. Bailey and Rachel Greenstadt, editors, *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 3793–3810. USENIX Association, 2021.

[46] Baojun Liu, Zhou Li, Peiyuan Zong, Chaoyi Lu, Hai-Xin Duan, Ying Liu, Sumayah A. Alrwais, XiaoFeng Wang, Shuang Hao, Yaoqi Jia, Yiming Zhang, Kai Chen, and Zaifeng Zhang. Trafficking stop: Detecting and measuring illicit traffic monetization through large-scale DNS analysis. In *IEEE European Symposium on Security and Privacy, EuroS&P 2019, Stockholm, Sweden, June 17-19, 2019*, pages 560–575. IEEE, 2019.

[47] Jienan Liu, Pooja Pun, Phani Vadrevu, and Roberto Perdisci. Understanding, measuring, and detecting modern technical support scams. In *8th IEEE European Symposium on Security and Privacy, EuroS&P 2023, Delft, Netherlands, July 3-7, 2023*. IEEE, 2023.

[48] Mingxuan Liu, Yiming Zhang, Baojun Liu, Zhou Li, Haixin Duan, and Donghong Sun. Detecting and characterizing SMS spearphishing attacks. In *ACSAC '21: Annual Computer Security Applications Conference, Virtual Event, USA, December 6 - 10, 2021*, pages 930–943. ACM, 2021.

[49] Ruofan Liu, Yun Lin, Xianglin Yang, Siang Hwee Ng, Dinil Mon Divakaran, and Jin Song Dong. Inferring phishing intention via webpage appearance and dynamics: A deep vision based approach. In Kevin R. B. Butler and Kurt Thomas, editors, *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*. USENIX Association, 2022.

[50] Chaoyi Lu, Baojun Liu, Yiming Zhang, Zhou Li, Fenglu Zhang, Haixin Duan, Ying Liu, Joann Qiongna Chen, Jinjin Liang, Zaifeng Zhang, Shuang Hao, and Min Yang. From WHOIS to WHOWAS: A large-scale measurement study of domain registration privacy under the GDPR. In *28th Annual Network and Distributed System Security Symposium, NDSS 2021, virtually, February 21-25, 2021*. The Internet Society, 2021.

[51] Long Lu, Roberto Perdisci, and Wenke Lee. SURF: detecting and measuring search poisoning. In Yan Chen, George Danezis, and Vitaly Shmatikov, editors, *Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS 2011, Chicago, Illinois, USA, October 17-21, 2011*, pages 467–476. ACM, 2011.

[52] Xiulin Ma. Research on black hat seo behaviour measurement. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pages 1041–1045, 2018.

[53] MaxMind. Industry leading IP Geolocation and Online Fraud Prevention | MaxMind. https://www.maxmind.com/, 2024.

[54] Michal Wlosik. What Is Search Advertising and How Does It Work? https://clearcode.cc/blog/what-is-search-advertising/, 2024.

[55] Ministry of Industry and Information Technology of the People's Republic of China. ICP/IP address/domain name registration information query. https://bzxx.miit.gov.cn/bzxx/problem/detail?id=C20C288CE25B449EAC08869F5EDE2F95, 2024.

[56] Tyler Moore, Nektarios Leontiadis, and Nicolas Christin. Fashion crimes: trending-term exploitation on the web. In Yan Chen, George Danezis, and Vitaly Shmatikov, editors, *Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS 2011, Chicago, Illinois, USA, October 17-21, 2011*, pages 455–466. ACM, 2011.

[57] Mohamed Nabeel, Enes Altinisik, Haipei Sun, Issa Khalil, Wendy Hui Wang, and Ting Yu. CADUE: content-agnostic detection of unwanted emails for enterprise security. In Leyla Bilge and Tudor Dumitras, editors, *RAID '21: 24th International Symposium on Research in Attacks, Intrusions and Defenses, San Sebastian, Spain, October 6-8, 2021*. ACM, 2021.

[58] Aleksandr Nahapetyan, Sathvik Prasad, Kevin Childs, Adam Oest, Yeganeh Ladwig, Alexandros Kapravelos, and Bradley Reaves. On sms phishing tactics and infrastructure.

[59] PengPai News. Online game transactions were defrauded by the "recharge and unfreeze"' routine, and a middle school student was defrauded of nearly 70,000 yuan. https://m.thepaper.cn/kuaibao_detail.jsp?contid=8506254&from=kuaibao, 2024.

[60] PlayerAuctions. Sell Game Accounts. https://www.playerauctions.com/about/sell-game-accounts/, 2024.

[61] Victor Le Pochat, Tom van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczynski, and Wouter Joosen. Tranco: A research-oriented top sites ranking hardened against manipulation. In *26th Annual Network and Distributed System Security Symposium*, 2019.

[62] Sathvik Prasad, Elijah Bouma-Sims, Athishay Kiran Mylappan, and Bradley Reaves. Who's calling? characterizing robocalls through audio and metadata analysis. In Srdjan Capkun and Franziska Roesner, editors, *29th USENIX Security Symposium, USENIX Security 2020, August 12-14, 2020*. USENIX Association, 2020.

[63] Sathvik Prasad, Trevor Dunlap, Alexander J. Ross, and Bradley Reaves. Diving into robocall content with snorcall. In Joseph A. Calandrino and Carmela Troncoso, editors, *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, pages 427–444. USENIX Association, 2023.

[64] Bradley Reaves, Logan Blue, Hadi Abdullah, Luis Vargas, Patrick Traynor, and Thomas Shrimpton. Authenticall: Efficient identity and content authentication for phone calls. In Engin Kirda and Thomas Ristenpart, editors, *26th USENIX Security Symposium, USENIX Security 2017, Vancouver, BC, Canada, August 16-18, 2017*, pages 575–592. USENIX Association, 2017.

[65] Richard Penman. whois. https://github.com/richardpenman/whois, 2024.

[66] John Seymour and Philip Tully. Generative models for spear phishing posts on social media. *CoRR*, abs/1802.05196, 2018.

[67] Stop Forum Spam. Stop forum spam. https://www.stopforumspam.com/, 2023.

[68] Spam List. Spam List. https://joewein.net/spam/spam-bl-b.htm, 2023.

[69] Bharat Srinivasan, Athanasios Kountouras, Najmeh Miramirkhani, Monjur Alam, Nick Nikiforakis, Manos Antonakakis, and Mustaque Ahamad. Exposing search and advertisement abuse tactics and infrastructure of technical support scammers. In Pierre-Antoine Champin, Fabien Gandon, Mounia Lalmas, and Panagiotis G. Ipeirotis, editors, *Proceedings of the 2018 World Wide Web Conference on World Wide Web, WWW 2018, Lyon, France, April 23-27, 2018*, pages 319–328. ACM, 2018.

[70] statista. Average number of daily active users of Baidu app from 3rd quarter 2018 to 4th quarter 2021. https://www.statista.com/statistics/1079983/baidu-app-quarterly-daily-active-users/, 2024.

[71] Tencent Cloud. The price of domain name registration. https://buy.cloud.tencent.com/domain/price?type=overview, 2024.

[72] Kurt Thomas, Damon McCoy, Chris Grier, Alek Kolcz, and Vern Paxson. Trafficking fraudulent accounts: The role of the underground market in twitter spam and abuse. In Samuel T. King, editor, *Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013*. USENIX Association, 2013.

[73] Huahong Tu, Adam Doupé, Ziming Zhao, and Gail-Joon Ahn. Users really do answer telephone scams. In Nadia Heninger and Patrick Traynor, editors, *28th USENIX Security Symposium, USENIX Security 2019, Santa Clara, CA, USA, August 14-16, 2019*, pages 1327–1340. USENIX Association, 2019.

[74] URLHaus. Urlhaus. https://urlhaus.abuse.ch/, 2023.

[75] Rakesh M. Verma and Keith Dyer. On the character of phishing urls: Accurate and robust statistical learning classifiers. In Jaehong Park and Anna Cinzia Squicciarini, editors, *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy, CODASPY 2015, San Antonio, TX, USA, March 2-4, 2015*, pages 111–122. ACM, 2015.

[76] David Y. Wang, Matthew F. Der, Mohammad Karami, Lawrence K. Saul, Damon McCoy, Stefan Savage, and Geoffrey M. Voelker. Search + seizure: The effectiveness of interventions on SEO campaigns. In Carey Williamson, Aditya Akella, and Nina Taft, editors, *Proceedings of the 2014 Internet Measurement Conference, IMC 2014, Vancouver, BC, Canada, November 5-7, 2014*, pages 359–372. ACM, 2014.

[77] David Y. Wang, Stefan Savage, and Geoffrey M. Voelker. Cloak and dagger: dynamics of web search cloaking. In Yan Chen, George Danezis, and Vitaly Shmatikov, editors, *Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS 2011, Chicago, Illinois, USA, October 17-21, 2011*. ACM.

[78] Peng Wang, Zilong Lin, Xiaojing Liao, and XiaoFeng Wang. Demystifying local business search poisoning for illicit drug promotion. In *29th Annual Network and Distributed System Security Symposium, NDSS 2022, San Diego, California, USA, April 24-28, 2022*. The Internet Society, 2022.

[79] Peng Wang, Xianghang Mi, Xiaojing Liao, XiaoFeng Wang, Kan Yuan, Feng Qian, and Raheem A. Beyah. Game of missuggestions: Semantic analysis of search-autocomplete manipulations. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society, 2018.

[80] Whatsapp. About suspicious links. https://faq.whatsapp.com/393169153028916/?cms_platform=web&cms_id=393169153028916&draft=false, 2024.

[81] Colin Whittaker, Brian Ryner, and Marria Nazif. Large-scale automatic classification of phishing pages. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2010, San Diego, California, USA, 28th February - 3rd March 2010*. The Internet Society, 2010.

[82] Baoning Wu and Brian D. Davison. Cloaking and redirection: A preliminary study. In *AIRWeb 2005, First International Workshop on Adversarial Information Retrieval on the Web, co-located with the WWW conference, Chiba, Japan, May 2005*, pages 7–16, 2005.

[83] Baoning Wu and Brian D. Davison. Identifying link farm spam pages. In Allan Ellis and Tatsuya Hagino, editors, *Proceedings of the 14th international conference on World Wide Web, WWW 2005, Chiba, Japan, May 10-14, 2005 - Special interest tracks and posters*, pages 820–829. ACM, 2005.

[84] Baoning Wu and Brian D. Davison. Detecting semantic cloaking on the web. In Les Carr, David De Roure, Arun Iyengar, Carole A. Goble, and Michael Dahlin, editors, *Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh, Scotland, UK, May 23-26, 2006*, pages 819–828. ACM, 2006.

[85] Guang Xiang, Jason I. Hong, Carolyn P. Rosé, and Lorrie Faith Cranor. CANTINA+: A feature-rich machine learning framework for detecting phishing web sites. *ACM Trans. Inf. Syst. Secur.*, 14(2):21:1–21:28, 2011.

[86] Qinge Xie, Shujun Tang, Xiaofeng Zheng, Qingran Lin, Baojun Liu, Haixin Duan, and Frank Li. Building an open, robust, and stable voting-based domain top list. In Kevin R. B. Butler and Kurt Thomas, editors, *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, pages 625–642. USENIX Association, 2022.

[87] Xinhua Net. Baidu will implement real-name registration, search will not be affected. http://news.xinhuanet.com/politics/2017-05/12/c_1120959353.htm, 2017.

[88] Ronghai Yang, Xianbo Wang, Cheng Chi, Dawei Wang, Jiawei He, Siming Pang, and Wing Cheong Lau. Scalable detection of promotional website defacements in black hat SEO campaigns. In Michael Bailey and Rachel Greenstadt, editors, *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 3703–3720. USENIX Association, 2021.

[89] Yoast - SEO for everyone. How to create the right meta description. https://yoast.com/meta-descriptions/, 2023.

[90] Weider D. Yu, Shruti Nargundkar, and Nagapriya Tiruthani. Phishcatch - A phishing detection tool. In Sheikh Iqbal Ahamed, Elisa Bertino, Carl K. Chang, Vladimir Getov, Lin Liu, Ming Hua, and Rajesh Subramanyan, editors, *Proceedings of the 33rd Annual IEEE International Computer Software and Applications Conference, COMPSAC 2009, Seattle, Washington, USA, July 20-24, 2009*. IEEE Computer Society, 2009.

[91] Kan Yuan, Di Tang, Xiaojing Liao, XiaoFeng Wang, Xuan Feng, Yi Chen, Menghan Sun, Haoran Lu, and Kehuan Zhang. Stealthy porn: Understanding real-world adversarial images for illicit online promotion. In *2019 IEEE Symposium on Security and Privacy, SP 2019, San Francisco, CA, USA, May 19-23, 2019*, 2019.

[92] Jialong Zhang, Chao Yang, Zhaoyan Xu, and Guofei Gu. Poisonamplifier: A guided approach of discovering compromised websites through reversing search poisoning attacks. In Davide Balzarotti, Salvatore J. Stolfo, and Marco Cova, editors, *Research in Attacks, Intrusions, and Defenses - 15th International Symposium, RAID 2012, Amsterdam, The Netherlands, September 12-14, 2012. Proceedings*, volume 7462 of *Lecture Notes in Computer Science*, pages 230–253. Springer, 2012.

[93] Shengnan Zhang, Yan Hu, and Guangrong Bian. Research on string similarity algorithm based on levenshtein distance. In *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pages 2247–2251. IEEE, 2017.

[94] Yue Zhang, Jason I. Hong, and Lorrie Faith Cranor. Cantina: a content-based approach to detecting phishing web sites. In Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider, and Prashant J. Shenoy, editors, *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada, May 8-12, 2007*, pages 639–648. ACM, 2007.

[95] ZoneFiles. Zonefiles. https://zonefiles.io/detailed-domain-lists/, 2023.

A User Complaint

Table 3 shows the key information extracted from an example complaint.

Table 3: Scam information extracted from user compliant.

Search Keyword	WuZhuang Account
Domain in Search Result	hz7re.huagong1.top/ebook/q13in
Redirection Domain	gg1.cfcfcg.top/vvvvv
Scamed or Not	yes
Scamed Money	800 RMB
Scamed Date	August 13, 2023
Payment Channels	Bank Card
Receiving Account	623******647

B NOKEScam Title Example

Table 4 shows several examples of NOKEScam webpage title.

Table 4: Examples of NOKEScam website title patterns.

Business	NOKEScam Website Title	English Title
Online Game Transaction Fraud	绒斌购-快速网游在线交易平台	RongBinGou-Online game trading fast portal
Online Game Transaction Fraud	蚁裔游 /热门游戏-“一带一路研讨会”重点演讲内容…引发热烈讨论	YiYiYou / Hot Games-Key presentations at “Belt and Road Seminar” including…Sparks Intense Discussion
Gambling	僿卜站··彩票预测网站	SaiBoZhan··Lottery prediction website
Porn	色脐阁：便捷视频网站	QinQiZhan： Convenient video website

* The underlined part is the NSKeyword in the title.

C NSKeyword Identifier for English

We perform word segmentation on English text using frequency data from the Wikipedia corpus via wordninja[7]. This process identifies non-existent words, e.g., “Navigrake”[8]. To thwart adversaries using uncommon word combinations to create NSKeywords, we trained an N-gram model on diverse text sources to predict word probabilities based on preceding words. This approach facilitates the detection of English NSKeywords, and we have open-sourced code for both English and Chinese for community reference and validation[9].

D Real-World Case Study

This real-world case provides valuable insights for the security community and offers concrete educational material for user awareness.

Is your account for sale? I'll buy your account for $1,000.

Script 1: Money Temptation. Scammers discreetly transmit NSKeywords to victims via platforms like in-game chat, offering to purchase accounts at premium rates.

> *Let's trade on the “Zunu game”. I have traded on it many times before. You can search it with a search engine, and he is the one who ranks first, and it has been certified by the search engine.*

Script 2: Search engine endorsement. They suggest a specific trading platform (NSKeyword), using high-search-ranking effects to imply certification and credibility. To overcome victim hesitation, scammers may claim prior successful transactions on the platform, offer higher prices, and create urgency for quick decision-making.

> ***Warning****: Because the information you entered is incorrect, your account has been frozen and you need to recharge $200 to unblock the account.*

Script 3: Deposit of funds. Upon attempting fund withdrawal, the victim's account is frozen by the fraudulent platform, citing incorrect information. To unfreeze, the platform demands a deposit. After compliance, the account remains blocked, with escalating deposit requirements. This tactic exploits the sunk cost fallacy, preying on victims' aversion to losing their initial investment.

> *I have already paid. It was* ***your mistake*** *that caused the account to be frozen.*
> *In this way, not only did I lose money, but I also lost access to your account. You solve it quickly, or I will call the police.*

Script 4: Defeat the defense. When victims express suspicion about deposit demands, adversaries employ psychological manipulation. They shift blame to the victim, claiming the account freeze resulted from the victim's errors. This tactic aims to induce guilt and pressure compliance, suggesting the adversary has also suffered losses due to the victim's alleged mistakes. Once the victim feels guilty, they post a deposit. If the complaint is unsuccessful, the adversary may escalate the situation by threatening to involve law enforcement unless the victims post a bond and resolve the matter promptly.

[7] https://github.com/keredson/wordninja.

[8] An example of an English NOKEScam website: http://epptest.shop/.

[9] We released the script on http://nokescam.com.