Work-in-Progress: A Large-Scale Long-term Analysis of Online Fraud across Multiple Companies and Platforms

Yi Han
F5, Inc.
y.han@f5.com

Shujiang Wu
F5, Inc.
sh.wu@f5.com

Mengmeng Li
F5, Inc.
m.li@f5.com

Zixi Wang
F5, Inc.
zi.wang@f5.com

Pengfei Sun
F5, Inc.
p.sun@f5.com

***Abstract*—Online fraud has emerged as a formidable challenge in the digital age, presenting a serious threat to individuals and organizations worldwide. Characterized by its ever-evolving nature, this type of fraud capitalizes on the rapid development of Internet technologies and the increasing digitization of financial transactions. In this paper, we address the critical need to understand and combat online fraud by conducting an unprecedented analysis based on extensive real-world transaction data. Our study involves a multi-angle, multi-platform examination of fraudsters' approaches, behaviors and intentions. The findings of our study are significant, offering detailed insights into the characteristics, patterns and methods of online fraudulent activities and providing a clear picture of the current landscape of digital deception. To the best of our knowledge, we are the first to conduct such large-scale measurements using industrial-level real-world online transaction data.**

I. Introduction

In the contemporary digital era, the menace of online fraud has become increasingly prevalent, posing a significant risk to the integrity and security of digital transactions and the confidentiality of personal information. The substantial financial damages resulting from such fraudulent activities have garnered sustained interest and concern from sectors spanning industry, academia, and government [12], [21]. Examples of these fraudulent activities include targeting online payment and banking services, leading to considerable financial losses for victims. So, fraud detection has been a crucial area of concern in both academic and industrial research, significantly impacting the prevention of financial loss, the preservation of customer trust, and the protection of privacy.

Amidst this backdrop, it is crucial to distinguish online fraud from another prevalent form of online malicious activity: the use of bots [19], [8], [1], [7]. Bots, automated software programs designed to perform repetitive tasks, can be utilized in both benign and malicious ways. Malicious bots, distinct from fraud, are often employed for purposes such as website scraping, spamming, and launching Distributed Denial of Service (DDoS) attacks. While bots can be a component of fraudulent schemes, they typically do not have the direct financial motivation characteristic of fraud. Instead, their primary aim is often to disrupt services, manipulate data, or overwhelm systems. The distinction between fraud and bot activities is not merely academic but has practical implications in terms of detection and prevention strategies. Fraud typically involves more personalized tactics, targeting specific individuals or organizations with deceptive intent. In contrast, bot-driven activities are usually more generalized and automated, focusing on exploiting systemic vulnerabilities. Understanding these nuances is essential for developing effective countermeasures.

While there is a plethora of research papers [11], [14], [20], [13], [10], [5], [9] discussing methods of fraud detection, there is a notable lack of in-depth fraud behavior analysis. We aim to tackle this widespread issue by conducting a thorough analysis of online fraud, using a large dataset of real-world online transaction data. In this paper, we perform the first large-scale measurement study of fraud and benign traffic collected from 3 commercial websites (including major financial institutes and restaurants). Specifically, we cooperate with a security company to collect web and mobile traffic and classify traffic using its bot and fraud detection and defense system.

Our research contributes to the field by providing a comprehensive examination of the techniques, development, and challenges faced across different platforms in online fraud. By analyzing real-world online transaction data, we offer an empirical insight into the complexity and ever-changing nature of fraudulent activities. Our analysis makes two observations:

Observation-1: Fraudsters, usually being real individuals operating genuine devices, naturally maintain browser environments that resemble those of legitimate users (e.g. up-to-date browser versions and genuine UAs). This contrasts markedly with the behavior of bots which may utilize a range of browser versions and fake/spoofed UAs.

Observation-2: Fraud traffic has unique characteristics compared to benign traffic. Such characteristics includes how fraudsters manipulate the device, their behavioral pattern and their intentions.

In summary, this paper aims to enhance the understanding of online fraud through a detailed examination of real-world data. By analyzing the patterns, methodologies, and evolution

Workshop on Measurements, Attacks, and Defenses for the Web (MADWeb) 2024
1 March 2024, San Diego, CA, USA
ISBN 979-8-9894372-2-1
https://dx.doi.org/10.14722/madweb.2024.23xxx
www.ndss-symposium.org

of online fraud, we contribute valuable insights to the field and highlight the urgent need for ongoing innovation in cybersecurity strategies.

II. Background

In this section, we introduce necessary background knowledge for our measurement study.

A. Online Fraud

Online fraud involves unauthorized activities initiated by adversaries against online services, impersonating victims to achieve financial gain. This type of fraud is prevalent in various sectors, including financial institutions, retail stores, and restaurants, posing significant challenges. Online fraud manifests in multiple forms throughout different stages of a user's interaction with an online service. For instance, in the banking sector, a user's journey begins with account creation[3], [4], [6], [16], followed by activities such as login[2], [15], [17], [18], account modification, adding payees, and transferring money. Fraud can occur at any point in this process. Examples include Account Origination Fraud, where fraudsters create numerous bogus accounts for purposes like money laundering. Another example is Payment Fraud, involving unauthorized control over stolen accounts to siphon funds to accounts under the fraudster's control. Similarly, in retail and restaurant industries, once fraudsters gain access to customer accounts, they can make unauthorized purchases and resell the acquired items for profit, an act known as Purchase Fraud.

B. Bot

Bots are automated software applications designed to perform tasks on the internet much faster than a human could. They range from benign programs that index web content for search engines to malicious bots involved in spamming, data theft, and automated attacks. Bots can mimic or replace human interaction in online environments, executing repetitive tasks efficiently. This capability is leveraged in various ways, from customer service chatbots that offer instant responses to queries, to bots that scrape websites for data aggregation or competitive analysis. However, the darker side of online bots includes their use in cyber-attacks, such as Distributed Denial of Service (DDoS) attacks, where multiple bots flood a website with traffic to render it inaccessible. Additional nefarious uses involve ad fraud, which drains advertising budgets through fake clicks and impressions; fake account creation for spamming or scamming; credential stuffing, where bots test stolen login credentials across various websites to commit identity theft or fraud; price scraping by competitors to undermine business strategies; and inventory hoarding in the retail sector, where bots quickly reserve or purchase limited-stock items for resale at higher prices.

C. Fraud v.s. Bot

Online fraud and bots share many similarities but they also differ significantly in various aspects. Bots are automated software programs designed to perform tasks at a speed and scale far beyond human capability. Fraud, on the other hand, often involves manual intervention and decision-making by individuals or groups with malicious intent. While automated tools can assist in fraud schemes, the strategic elements, such as targeting specific victims or exploiting particular vulnerabilities, typically require human oversight. Bots operate on a large scale, capable of executing thousands to millions of tasks or attacks in a short period. This scale is particularly useful for activities such as service interruption (DDoS attacks) or widespread data scraping. Fraud tends to be more narrowly focused, targeting specific individuals, companies, or transactions. Fraudsters often aim to remain undetected for as long as possible to maximize gains from a particular scheme, requiring a more selective and cautious approach compared to the broad and indiscriminate nature of bots. Bots can serve a multitude of purposes, not all of which are malicious. They can be deployed for legitimate functions like automated customer support or illegitimate activities including service interruption, information theft, and financial gain through cyber-attacks. Fraud is primarily financially oriented, with the ultimate goal of illicitly obtaining money or valuable information that can be converted to financial gain. Fraud schemes are designed with the intent to deceive and exploit, focusing on monetary outcomes. Both bots and fraud schemes employ retooling strategies to evade detection by security measures. As bots are identified and blocked by cybersecurity defenses, their developers continuously update their methods to circumvent new security protocols. Similarly, fraudsters adapt their tactics in response to enhanced fraud detection mechanisms, altering their approaches to avoid suspicion and continue their illicit activities without being caught.

III. Detection Architecture and Integration

In this section, we delineate the dual facets of our fraud detection system, which hinges on a learning-based analysis mechanism alongside the integration of customer feedback. This combined approach underpins our system's proficiency in discerning online fraud, encapsulating the essence of our methodology as expounded subsequently.

A. Learning-Based Analysis

In the development of our learning-based analysis framework, we meticulously collect and analyze a wide spectrum of signals to enhance the precision in detecting fraudulent activities. For web-based interactions, our system gathers an extensive set of signals including URLs, cookies and other relevant web traffic indicators, each contributing to a comprehensive understanding of potential security threats.

In parallel, our approach to mobile applications involves the analysis of specialized mobile signals. These encompass device-specific attributes, app usage patterns, and other mobile-centric indicators, allowing for a nuanced detection of fraud that is tailored to the mobile ecosystem.

Additionally, our model incorporates an in-depth analysis of user behavior across both platforms. By examining patterns of interaction, transaction histories, and user engagement metrics,

we can identify deviations from established norms that may signify fraudulent intent.

Together, these diverse data sources—from web and mobile signals to user behavior analytics—equip our system with a multi-layered perspective on potential fraud, ensuring a robust and adaptable defense mechanism against online threats.

B. Incorporating Customer Feedback

The integration of customer feedback serves as the secondary core component of our fraud detection schema. This feedback ranges broadly, covering anomalies flagged by users—ranging from irregular login activities to unauthorized credit card transactions reported by banks. Rigorous human verification processes are applied to each piece of feedback, ensuring its pertinence and veracity. Such a protocol not only accelerates our capacity to respond to new threats but also aids in the ongoing enhancement of our algorithmic frameworks.

C. Integration and Veracity

Through the amalgamation of insights derived from our learning-based analysis and the incorporation of customer feedback, we have established a comprehensive framework that closely mirrors the ground truth with a minimal margin of error. This fusion of direct performance indicators, aggregated behavioral analyses, and authenticated user reports allows for the high-fidelity detection of fraudulent activities. Our confidence in the system-generated labels is bolstered by this integrative method, underscoring the reliability of our tool in the proactive mitigation of online fraud

IV. Dataset

In this section, we will describe the database we have collected. We have gathered data from three companies, totaling 797,273,912 HTTP(S) requests. Among these, 1.1% traffic identified as fraudulent data.

A. Data Collection

Our data collection is done by implanting a JavaScript code snippet onto the target page for web and an SDK into the native app for mobile. The collect raw data contains network- (e.g. IP, ASN,), browser- (e.g. browser fingerprint) and user- (e.g. user behaviors) level information.

B. Feature Collection

Table I presents the features collected and utilized for our analysis. We have categorized these features into four groups: Meta Information, Browser Fingerprint, Web Signal, and Mobile Signal. Each feature is detailed with a comprehensive description and exemplified. It is important to note that these features are not simultaneously collected on both web and mobile platforms, and are not gathered across multiple platforms at once.

C. Attack Statistics For Web Traffic

Table II exclusively contains data related to web-based activities and illustrates the categorization of data based on attack types. For different companies, we have covered various classifications of attacks, although there are some types of attacks not applicable to certain companies due to the absence of relevant services. For instance, Company Rest.A would not have services related to 'transfers'. In Table II, we have enumerated the total number of requests for each type and the proportion of fraudulent attacks within these categories. This analysis reveals a significant variation in the types of attacks faced by different companies. Despite being in the same industry, companies encounter distinct types of attacks based on their unique service offerings. For instance, Bank B is more susceptible to attacks during the sign-up process, while Bank A is more prone to fraudulent attacks during the login phase. Due to the data for mobile apps originating from various sources, we haven't categorized it based on attack type.

1) Attack Pattern Across a Five-Month Period: Figure 1 to 4 display the benign and fraudulent traffic for Bank A, Bank B, and Rest A (both web and app) over a five-month period, segmented with vertical lines on a weekly basis. Unlike patterns observed with bots, a striking similarity is noticeable between fraudulent attacks and benign traffic, characterized by regular oscillations on a weekly cycle. Intriguingly, fraudulent attacks markedly differ from bot activities. Being predominantly manual in nature, these attacks seem to follow a weekly schedule, humorously suggesting that even fraudsters might respect the concept of weekends, perhaps keeping them free for leisure.

2) Attack Pattern Across a 24-Hour Period: Figure 5 displays the distribution of benign and fraudulent traffic for Bank A, Bank B, and Rest A, showcasing the data across the 24-hour cycle of each day over a period of five months. Similar to the patterns observed on a weekly basis, the distribution of fraudulent attacks, being manually operated, closely mirrors that of benign traffic throughout the day.

V. Observation

A. Web Browser Traffic User Agent (UA) Analysis

Table III showcases the classification of our web traffic data based on User Agent (UA), divided by device, operating system, and browser. We have limited our display to mainstream UAs. The data reveals a discernible gap between the browser usage frequencies of attackers and benign users. Particularly noteworthy is the usage of iOS's Safari browser, where the proportion of use by attackers significantly exceeds that of benign users.

As established in previous measurement studies [19], most UA data in bot attacks is fabricated. Consequently, in the context of fraud attacks, it is crucial to discern whether differences in UA distribution arise from attackers having a limited range of devices or from deliberate UA spoofing.

Based on the browsers predominantly used by both benign users and attackers, as indicated in Table III, we selected

TABLE I
FEATURE USED IN OUR MEASUREMENT STUDY

Name	Description	Example
HTTP(s) Request	Requests to target websites	-
Meta-information	-	-
Traffic type	Benign or malicious (account takeover attempts, content scraping, fake account login, giftcard cracking, and fraud transactions)	Label
Source IP	Source Internet Protocol (IP) address	65.78.121.109
Source ASN	Source Autonomous System Number (ASN)	3,356
Browser Fingerprint	A combination of all features	-
User-Agent	"User-Agent" HTTP header	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/93.0.4577.82 Safari/573.36
Screen resolution	Screen size, color depth, and available Height/Left/Top/Width	1440×900×24, availHeight: 823, availLeft: 0, availTop: 25, availWidth: 1440
Web signal	A combination of Fraud detection features	-
Blur event	Window/tab loses focus	True/False
Invisibility event	Window/tab becomes invisible on screen	True/False
Time on page	Time in seconds a user spent on the page	100 seconds
Proxy Network	If a IP is from a proxy Network	True/False
Data center	If a IP comes from a data center	True/False
Mobile signal	A combination of Mobile features	-
Android OS version	Fraudsters prefer to use old versioned device	14/13/12 .etc
Android device brand	Fraudsters prefer to use popular or cheap Android devices	samsung
Device up time	How long has been since last boot	3.5 days
Emulator	If the current Android device is emulator	True/False
Rooted	If the current Android device is rooted	True/False
Hooked	If the current Android device has installed some rooting frameworks such as Frida	True/False

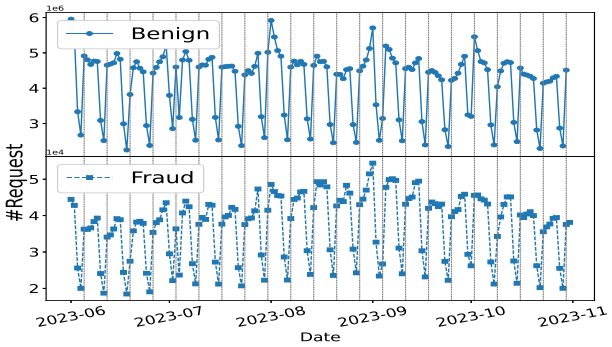

Fig. 1. Bank A 5 Months Traffic

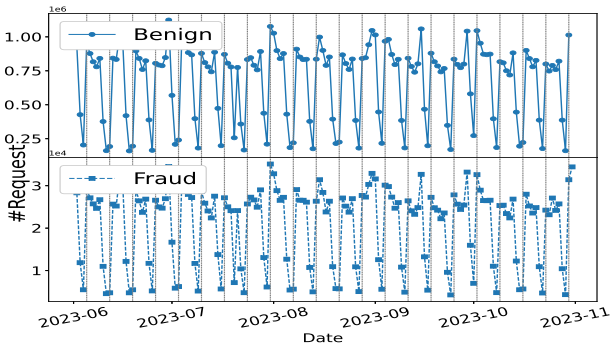

Fig. 2. Bank B 5 Months Traffic

TABLE II
BREAKDOWN OF WEB BENIGN VS. ADVERSARIAL REQUESTS BASED ON ATTACK TYPES

Company	sign up		login in		Transfer		Order	
	# request	Attack	# request	Attack	# request	Attack	# request	Attack
Bank A	-	-	98,439k	3.07%	-	-	9,768k	1.51%
Bank B	2,423k	9.24%	625,093k	0.86%	10,464k	0.29%	-	-
Rest. A Web	2,993k	0.01%	12,916k	0.03%	-	-	13,565k	0.03%

Windows Chrome, macOS Safari, and Android Chrome for a detailed comparative analysis of minor browser version numbers. As depicted in Figure 6 to 8, we observed that the percentages of browser versions used by attackers and benign users are strikingly similar, suggesting that attackers, like regular users, keep their browsers updated and provide authentic UAs. This observation elucidates the pattern seen in Table 3, where the differences in UA proportions are attributed to attackers having limited device options rather

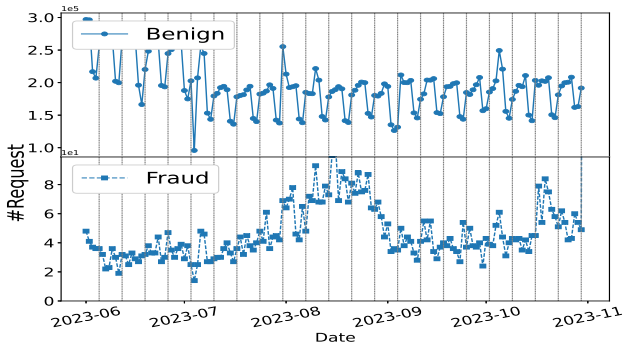

Fig. 3. Rest. A Web 5 Months Traffic

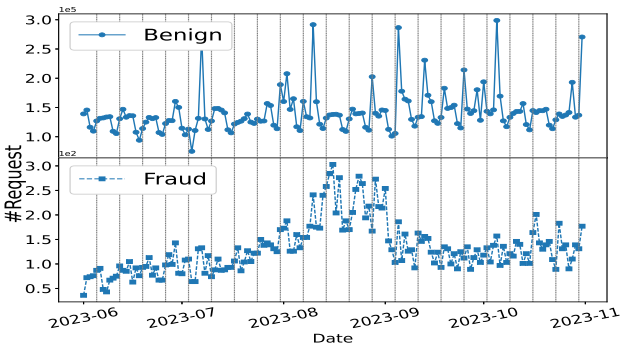

Fig. 4. Rest. A App 5 Months Traffic

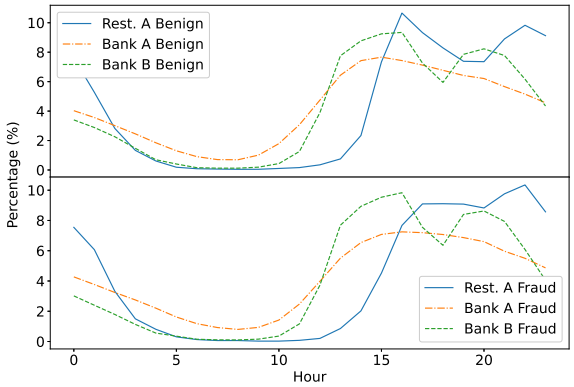

Fig. 5. Number of Benign and Adversarial Requests over One Day

than supplying false UAs, as is common with bots. The reason fraud attackers might use genuine UAs could be that fraud attacks require more complex and varied interactions. Given the heightened scrutiny in data analysis for attacker features, such as inconsistency checks, using real UAs could paradoxically increase the success rate of the attackers.

TABLE III
BREAKDOWN OF WEB BENIGN VS. ADVERSARIAL REQUESTS BASED ON USERAGENT

Device	OS	Browser	Benign	Adv
PC	Windows	Chrome	32.8%	23.3%
		Edge	15.7%	8.8%
		Firefox	<0.5%	1.1%
	Mac OS	Chrome	7.1%	3.8%
		Firefox	0.8%	<0.5%
		Safari	10.6%	3.9%
	Linux	Chrome	<0.5%	<0.5%
	Chrome OS	Chrome	1.5%	0.6%
Mobile	iOS	Safari	11.8%	39.3%
		Chrome	<0.5%	2.6%
		Safari WebView	1.2%	3.8%
	Android	Samsung Internet	1.2%	0.7%
		Chrome	9.4%	8.9%
		Chrome WebView	<0.5%	<0.5%

B. Device Shifting

One common approach fraudsters consider to stay untracked is to shift the devices they use (frequently). Device shifting does not really means one has to physically switch to a new device but simply clear the cookie of the browser. This is because web applications use cookie to identify and track devices. Once the cookie is cleared, the application will regenerate a new one and it will be treated as a new device. Benign users do not usually clear the the cookie often.

We study the device shifting pattern of fraudsters by looking at the age of a device. To compute the device age, we subtract the timestamp of the cookie from the timestamp when the request occurred. We define multiple time intervals (e.g., from 0 to 5 minutes) and count the number of fraud requests that fall into each interval. Figure 9 shows the results. The results are averaged over all the customers. It can be seen from the figure, most devices associated with fraudsters have an age below 5 minutes while benign devices tend to "live" longer.

C. Anonymization

Another weapon of fraudsters that helps them avoid detection is anonymization. The utilization of proxy and data center virtual machines are two popular approaches of anonymization. The aforementioned two approaches can be identified based on IP and ASN. To identify proxy, we compare the request IP with an IP intelligence dataset we maintained. As for ASN, we compare the request ASN also with an internal dataset we maintained. We show the percentage of proxy and data center in Table IV. Based on the results, fraudsters do utilize proxy and data center more frequently than benign users.

TABLE IV
PERCENTAGE OF PROXY IP AND DATA CENTER ASN. FRAUD TRAFFIC HAS SIGNIFICANTLY HIGHER PERCENTAGE ON BOTH FEATURES.

	% of Proxy IP		% of Data Center ASN	
	fraud	benign	fraud	benign
Bank A	87.2%	59.3%	15.3%	2.2%
Bank B	85.1%	61.5%	18.4%	3.7%
Rest. A	98.1%	60.7%	20.5%	4.8%
Rest. A App	84.7%	67.0%	7.5%	0.0%

D. Fraudster Behavior

Fraudsters usually exhibits distinct behavioral patterns compared to benign users as their purposes are different. The user journey of a benign user usually only involves signing up for their own account, signing in their own account

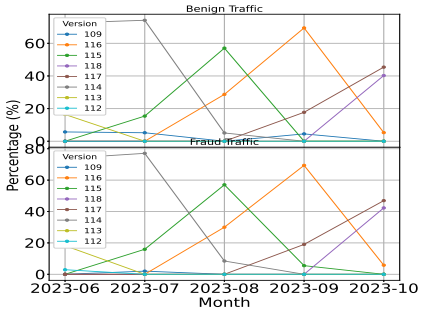

Fig. 6. Windows Chrome

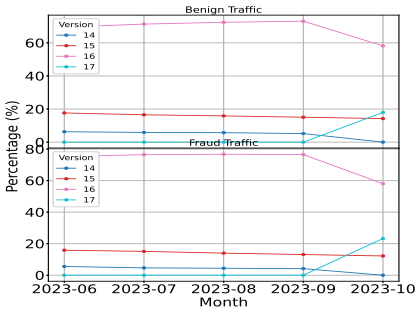

Fig. 7. MacOS Safari

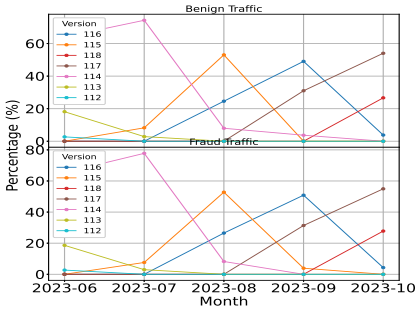

Fig. 8. Android Chrome

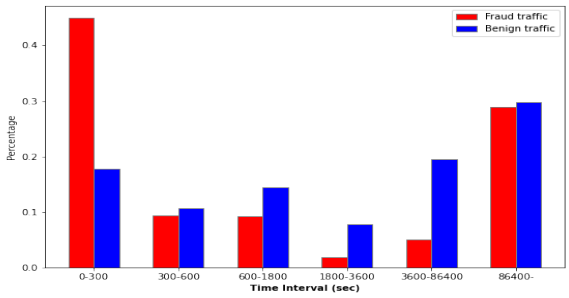

Fig. 9. Distribution of requests based on device age. A large portion of fraud devices have a device age of below 5 minutes.

or placing reasonable number of orders for themselves and family. Fraudsters's job on the other hand, involves signing up multiple accounts, signing in multiple stolen accounts and placing orders for their "customers".

One important set of characteristics related to dealing with large number of stolen accounts is multitasking. Again to efficiently process large number of accounts it is common for fraudsters to open multiple small browser windows, frequently switch between windows and minimize/maximize windows. For these behaviors, we use the percentage of blur events and invisibility events detected as the signals. A blur event is defined as when the window or tab loses focus. An invisibility event is defined as the window or tab becomes invisible on screen.

Fraudsters also tend to spend less time on a page. This is easy to understand as 1) they are very familiar with the page/App as they need to perform malicious tasks on the page/app repeatedly as opposed to benign users who has much less frequent usage of the page/App. 2) They have fixed goals (e.g. placing orders for their "customers") as opposed to benign users who sometimes browse around a bit. We measure the time in seconds on page to represent this behavior.

Table V shows the signals related to all fraudster behaviors mentioned above. All values are averaged over all the requests.

E. Fraudster Intention

Fraudster intention, i.e., how fraudsters exploits stolen accounts, is also quite different compared to how benign users uses their accounts. For example, fraudsters might sign in

TABLE V
SIGNALS RELATED TO FRAUDSTER BEHAVIORS.

	% blur events		% invisibility events		Time on page (sec)	
	fraud	benign	fraud	benign	fraud	benign
Bank A	17.9%	13.0%	16.4%	11.3%	56.2	80.2
Bank B	23.1%	17.2%	19.8%	12.4%	40.1	90.2
Rest. A	19.1%	14.3%	17.2%	12.5%	103.9	161.7

TABLE VI
FRAUDSTER INTENTION RELATED SIGNALS.

	# of login attempts		# of orders	
	fraud	benign	fraud	benign
Bank A	5.5	1.1	-	-
Bank B	4.7	1.0	-	-
Rest. A	3.3	1.3	3.5	1.5
Rest. A App	3.9	1.2	3.2	1.7

different stolen accounts or placing abnormal number of orders in a short period of time. Such scenarios can be represented by measuring the number of distinct account sign in attempts or the number of orders placed. To further hide themselves, fraudsters might choose to place orders at geologically distinct locations. To measure this, we compare the timezone of where the order is placed with the geolocation if the request ip and look for mismatch.

The results are presented in Table VI. All the results are averaged over all the fraud requests. It can be seen from the table, the difference between fraudulent and benign users are obvious.

F. Mobile App

Fraud from the mobile Apps has some unique characteristics. We discuss them in this section. Fraudsters tend to use old versions of the Android system as they are easier to root and are compatible with most common hooking frameworks or other security tools that fraudsters like to use. Figure 10 showed the distribution results of the Android OS version in both benign and fraud traffic.

Fraudsters tend to use Samsung and Xiaomi Android phones either because they are popular or they are cheaper and there are numerous online tutorials on how to root/hook/custom them. Figure 11 showed percentage of each brand. Fraudsters tend to restart the device frequently than the normal user in

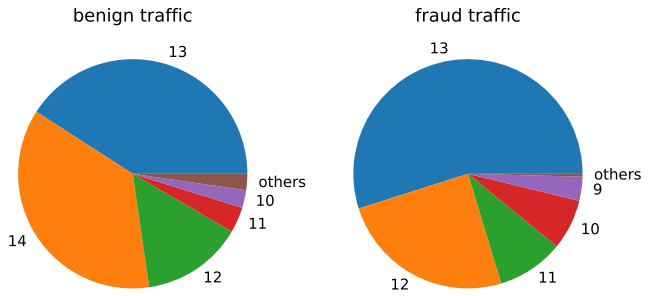

Fig. 10. The distribution of Android OS version in both benign and fraud traffic

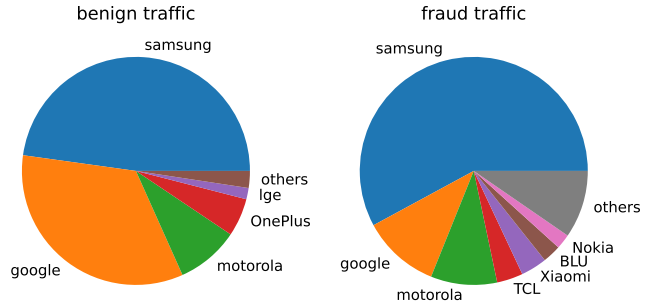

Fig. 11. The distribution of Android device brand in both benign and fraud traffic

order to apply some system change or app change, such as custom the system to spoof some system settings or properties to bypass detection, so the average device up time in the fraud traffic is much shorter than the benign traffic. This is confirmed based on our measurements: the average device up time of the benign traffic 8.23 days v.s. fraud traffic 3.87 days.

Fraudsters tend to root the Android device to get the full access to it. In this way, they can install any tools they want as well as customize and modify the Android device. In our data, we find that among all the benign traffic, 1.5% are either rooted or hooked, while for fraud traffic, the number is 3.5%. The percentage of emulator in benign traffic is 0.01% and in fraud traffic it is 5.46%.

Fraudsters tend to install hooking frameworks on the Android device to customize and modify them, such as they can spoof Android device identifiers by hooking some system APIs with hooking frameworks like Frida.

Fraudsters use Android emulators to reduce cost, as buying Android phones will cost some money especially if they want to do the fraud in scale, which needs lots of Android devices, while emulators are totally free. So it's no wonder that the percentage of Android emulators in the fraud traffic is much higher than that in the benign traffic, as is shown in the table VII.

VI. Discussion

A. Contributing to Fraud Detection Solutions

The distinct characteristics between fraudsters and benign users we observed in VI can be used in designing fraud detection solutions. The detailed observations from the study can be transformed into quantifiable features that serve as critical inputs for both rule-based systems and machine learning (ML) models. Specifically, for rule-based systems, these features can help in identifying complex patterns indicative of fraudulent behaviors. Detection rules can then be implemented accordingly. These features can also be used to train ML models such as classifiers or anomaly detectors to distinguish fraud requests from benign ones. This study's insights into fraudster characteristics also enable the identification of coordinated fraud campaigns. By analyzing the connectivity of requests based on various identification signals (e.g. IP, device ID, email etc.) together with commonalities in fraudster behavior and intention signals, it is possible to cluster related incidents into campaigns. Understanding the scope and modus operandi of these campaigns is crucial for developing targeted countermeasures and for understanding the broader threat landscape.

B. Limitation

The study relies on data from specific customers, which may not fully represent the entire spectrum of online fraud. Furthermore, online fraud activities are highly dynamic, with tactics and technologies constantly evolving. The study's findings may be time-sensitive, reflecting the state of online fraud behavior only during the period of data collection. This temporal limitation means that the findings may need updates or modifications in the future.

C. Addressing Data Ethics

We adhere to strict data ethics and privacy guidelines when conducting this measurement study. All customer data are anonymized to remove any personally identifiable information (PII). Additionally, we employ data minimization principles, i.e., only collecting and retaining data necessary for the research objectives of this study. We also establish an ethical review board to oversee the research process, ensuring that the study adheres to ethical standards and respects user privacy. We also maintain transparency with users about how their data is being used for research purposes. This includes clear communication through privacy policies, consent forms, and notifications, informing users about the nature of the research, the types of data collected, how it will be used, and measures taken to protect their privacy.

VII. Conclusion

In this paper, we perform the first large-scale measurement study of fraud and benign traffic collected from 3 commercial websites. We introduce features associated with fraud requests. We discuss observations based on these features to distinguish fraud traffic from benign traffic. Our discussion covers multiple angles including fraud traffic characteristics, environment spoofing, fraudster behavior and fraudster intention.

REFERENCES

[1] Babak Amin Azad, Oleksii Starov, Pierre Laperdrix, and Nick Nikiforakis. Web runner 2049: Evaluating third-party anti-bot services. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 17th International Conference, DIMVA 2020, Lisbon, Portugal, June 24–26, 2020, Proceedings 17*, pages 135–159. Springer, 2020.

[2] Minh Hieu Nguyen Ba, Jacob Bennett, Michael Gallagher, and Suman Bhunia. A case study of credential stuffing attack: Canva data breach. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 735–740. IEEE, 2021.

[3] Yazan Boshmaf, Dionysios Logothetis, Georgos Siganos, Jorge Lería, Jose Lorenzo, Matei Ripeanu, and Konstantin Beznosov. Integro: Leveraging victim prediction for robust fake account detection in osns. In *NDSS*, volume 15, pages 8–11. Citeseer, 2015.

[4] Yazan Boshmaf, Dionysios Logothetis, Georgos Siganos, Jorge Lería, Jose Lorenzo, Matei Ripeanu, Konstantin Beznosov, and Hassan Halawa. Íntegro: Leveraging victim prediction for robust fake account detection in large scale osns. *Computers & Security*, 61:142–168, 2016.

[5] Elie Bursztein, Borbala Benko, Daniel Margolis, Tadek Pietraszek, Andy Archer, Allan Aquino, Andreas Pitsillidis, and Stefan Savage. Handcrafted fraud and extortion: Manual account hijacking in the wild. In *Proceedings of the 2014 conference on internet measurement conference*, pages 347–358, 2014.

[6] Buket Erşahin, Özlem Aktaş, Deniz Kılınç, and Ceyhun Akyol. Twitter fake account detection. In *2017 International Conference on Computer Science and Engineering (UBMK)*, pages 388–392. IEEE, 2017.

[7] Hugo Jonker, Benjamin Krumnow, and Gabry Vlot. Fingerprint surface-based detection of web bot detectors. In *Computer Security–ESORICS 2019: 24th European Symposium on Research in Computer Security, Luxembourg, September 23–27, 2019, Proceedings, Part II 24*, pages 586–605. Springer, 2019.

[8] Xigao Li, Babak Amin Azad, Amir Rahmati, and Nick Nikiforakis. Good bot, bad bot: Characterizing automated browsing activity. In *2021 IEEE symposium on security and privacy (sp)*, pages 1589–1605. IEEE, 2021.

[9] Yang Liu, Xiang Ao, Zidi Qin, Jianfeng Chi, Jinghua Feng, Hao Yang, and Qing He. Pick and choose: a gnn-based imbalanced learning approach for fraud detection. In *Proceedings of the web conference 2021*, pages 3168–3177, 2021.

[10] Mingxuan Lu, Zhichao Han, Zitao Zhang, Yang Zhao, and Yinan Shan. Graph neural networks in real-time fraud detection with lambda architecture. *arXiv preprint arXiv:2110.04559*, 2021.

[11] Lin Meng, Hesham Mostafa, Marcel Nassar, Xiaonan Zhang, and Jiawei Zhang. Generative graph augmentation for minority class in fraud detection. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 4200–4204, 2023.

[12] Tahereh Pourhabibi, Kok-Leong Ong, Booi H Kam, and Yee Ling Boo. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*, 133:113303, 2020.

[13] Susie Xi Rao, Shuai Zhang, Zhichao Han, Zitao Zhang, Wei Min, Zhiyao Chen, Yinan Shan, Yang Zhao, and Ce Zhang. xfraud: explainable fraud transaction detection. *arXiv preprint arXiv:2011.12193*, 2020.

[14] Yizhuo Rao, Xianya Mi, Chengyuan Duan, Xiaoguang Ren, Jiajun Cheng, Yu Chen, Hongliang You, Qiang Gao, Zhixian Zeng, and Xiao Wei. Know-gnn: An explainable knowledge-guided graph neural network for fraud detection. In *International Conference on Neural Information Processing*, pages 159–167. Springer, 2021.

[15] Steven Rees-Pullman. Is credential stuffing the new phishing? *Computer Fraud & Security*, 2020(7):16–19, 2020.

[16] Somya Ranjan Sahoo and BB Gupta. Real-time detection of fake account in twitter using machine-learning approach. In *Advances in Computational Intelligence and Communication Technology: Proceedings of CICT 2019*, pages 149–159. Springer, 2021.

[17] Kurt Thomas, Jennifer Pullman, Kevin Yeo, Ananth Raghunathan, Patrick Gage Kelley, Luca Invernizzi, Borbala Benko, Tadek Pietraszek, Sarvar Patel, Dan Boneh, et al. Protecting accounts from credential stuffing with password breach alerting. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 1556–1571, 2019.

[18] Ke Coby Wang and Michael K Reiter. Detecting stuffing of a {User's} credentials at her own accounts. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 2201–2218, 2020.

[19] Shujiang Wu, Pengfei Sun, Yao Zhao, and Yinzhi Cao. Him of many faces: Characterizing billion-scale adversarial and benign browser fingerprints on commercial websites. In *30th Annual Network and Distributed System Security Symposium, NDSS*, 2023.

[20] Jianke Yu, Hanchen Wang, Xiaoyang Wang, Zhao Li, Lu Qin, Wenjie Zhang, Jian Liao, and Ying Zhang. Group-based fraud detection network on e-commerce platforms. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 5463–5475, 2023.

[21] Ge Zhang, Jia Wu, Jian Yang, Amin Beheshti, Shan Xue, Chuan Zhou, and Quan Z Sheng. Fraudre: Fraud detection dual-resistant to graph inconsistency and imbalance. In *2021 IEEE International Conference on Data Mining (ICDM)*, pages 867–876. IEEE, 2021.