Bits and Pieces

Piecing Together Factors of IoT Vulnerability Exploitation

Al Alsadi, Arwa Abdulkarim; Vermeer, Mathew; Sasaki, Takayuki; Yoshioka, Katsunari; Van Eeten, Michel; Gañán, Carlos

DOI
10.1145/3708821.3733875

Publication date
2025

Document Version
Final published version

Published in
ASIA CCS '25: Proceedings of the 20th ACM Asia Conference on Computer and Communications Security

Bits and Pieces: Piecing Together Factors of IoT Vulnerability Exploitation

Arwa Abdulkarim Al Alsadi
Delft University of Technology
Delft, Netherlands
a.alsadi@tudelft.nl

Mathew Vermeer
Delft University of Technology
Delft, Netherlands
m.vermeer@tudelft.nl

Takayuki Sasaki
Yokohama National University
Yokohama, Japan
sasaki-takayuki-yv@ynu.ac.jp

Katsunari Yoshioka
Yokohama National University
Yokohama, Japan
yoshioka@ynu.ac.jp

Michel Van Eeten
Delft University of Technology
Delft, Netherlands
m.j.g.vaneeten@tudelft.nl

Carlos Gañán
Delft University of Technology
Delft, Netherlands
c.hernandezganan@tudelft.nl

Abstract

The proliferation of Internet of Things (IoT) devices has led to a surge in vulnerabilities, with traditional metrics like CVSS and PoC exploits failing to fully explain exploitation patterns. To address this, we leverage features from the-state-of-the-art prediction model EPSS – such as CVSS, CWE, vendors, external references, vulnerability age, and PoCs – and combine it with new features derived from hacking communities. Our study of 23,373 IoT-related CVEs and 25k posts from 25 hacking forums highlights the importance of including insights on attacker behavior from discussions involving vulnerabilities. We identified 38 features with a p-value < 0.05 that impact attackers' selection of IoT vulnerabilities. We use two metrics to evaluate our model with features from hacking forums: McFadden's pseudo R^2, which showed a 21% improvement in explaining variance, and the Brier score for prediction accuracy, with a 17% improvement over EPSS. These results emphasize that current state-of-the-art methods struggle to capture the distinct nuances and complexity of IoT threats, and incorporating available information such as insights into attacker behavior can enhance the factors influencing the targeting of IoT vulnerability better.

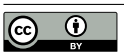
CCS Concepts

• **Computer systems organization** → **Embedded systems**; • **Security and privacy** → **Distributed systems security**; **Vulnerability management**.

Keywords

Vulnerability, Exploits, IoT, Underground forums

ACM Reference Format:
Arwa Abdulkarim Al Alsadi, Mathew Vermeer, Takayuki Sasaki, Katsunari Yoshioka, Michel Van Eeten, and Carlos Gañán. 2025. Bits and Pieces: Piecing Together Factors of IoT Vulnerability Exploitation. In *ACM Asia Conference on Computer and Communications Security (ASIA CCS '25), August 25–29, 2025, Hanoi, Vietnam.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3708821.3733875



ASIA CCS '25, Hanoi, Vietnam

ACM ISBN 979-8-4007-1410-8/25/08
https://doi.org/10.1145/3708821.3733875

1 Introduction

In recent years, the proliferation of Internet of Things (IoT) devices has brought about a corresponding surge in vulnerabilities, amplifying the potential risk of exploitation. For example, accordingly to VARIoT [43], the number of reported IoT vulnerabilities increased from 772 in 2010 to over 3,253 by 2022. This escalation in volume of IoT vulnerabilities necessitates a critical examination of the current state-of-the-art mechanisms for predicting and understanding the factors that lead to exploitation. Since it is unlikely that vendors and users have the time and resources to mitigate all discovered vulnerabilities, it helps to focus on those that actually will get exploited.

Previous research has shown the importance of using IoT-specific features for understanding IoT targeting in the wild [1, 2, 10, 14, 46]. However, the collection of these IoT-specific features such as device types [2, 10, 14, 18, 63] and number of exposed devices in the internet [2, 14, 54] is labor-intensive and currently not scalable and requiring manual work. Therefore, we want to support vendors and users by improving our understanding of what factors make certain IoT vulnerabilities targeted by attackers.

Since 2003, security practitioners have commonly relied on the Common Vulnerability Scoring System (CVSS) public information as a metric for predicting which vulnerabilities are more likely to be targeted by attackers – even though the developers of CVSS actually state it should not be used for this purpose [65]. Higher CVSS scores are interpreted to mean that the vulnerability is more likely to be exploited, hence more urgently needs to be patched. Yet, empirical research has indicated CVSS is unable to predict attacks in the wild [5, 19, 59]. Additionally, on average, 49% of real-world exploits occur before CVSS scores are published [21].

Recognizing the limitations of vulnerability severity scores in exploit prediction, others have explored alternative predictors such as the availability of Proof of Concept (PoC) exploits in public. Unfortunately, this feature has also been questioned as a reliable predictor, as instances most vulnerabilities with PoC exploits never manifest as real-world targets of attacks [4, 5]. Only 4.17% vulnerabilities get associated public exploits within 365 days [36]. Do most attacks occur within that set of vulnerabilities with a PoC?

To fill this gap in understanding which vulnerabilities get attacked in the wild within 30 days, Jacobs et al. introduced the Exploit Prediction Scoring System (EPSS) [41]. It integrated information from various sources such as CVSS scores, vulnerability

features from the National Vulnerability Database (NVD), availability of PoC exploits, vulnerabilities used in offensive security tools, and features derived from social media. At the time of writing, 111 vendors have integrated EPSS in their products, including IoT vendors [29] and platforms such as Vulners [72] and Shodan [61].

While EPSS has shown promise in predicting exploitation for the total set of CVEs, this set consists mostly of CVEs for general-purpose IT systems. EPSS often assigns low scores to IoT vulnerabilities, even to those known to be exploited in the wild, such as the ones included in CISA's Known Exploited Vulnerabilities (KEV) catalog [64]. For instance, CVE-2017-17215 was assigned an EPSS score indicating a near-zero probability (0.05) of being exploited in the wild within 30 days, starting from May 2021, which is the earliest available score in EPSS. It took over three years for the score to increase significantly, reaching 0.96 by April 2023. However, this does not align with real-world findings, as this IoT vulnerability was identified as one of the most frequently and consistently targeted vulnerability over an extended period in multiple studies and security reports conducted as early as January 2015, all the way up to February 2024 [1, 2, 10, 24, 25, 38].

This indicates potential limitations in accurately assessing the severity of IoT vulnerabilities. To raise awareness among vendors and users of EPSS, we aim to reassess the effectiveness of features like CVSS, PoC, and others used in EPSS, specifically for IoT-related vulnerabilities. Any evaluation of EPSS is complicated, however, because the EPSS is unavailable and not fully disclosed, even in its peer-reviewed publications. Without access to the underlying data, model, and source code, independent verification or explanation of its findings is not possible.

We aim to better explain the targeting of IoT vulnerabilities over EPSS by testing whether the inclusion of features from hacking forums can improve the model performance. This extension builds on work that showed black market or underground hacking forums provide information for predicting exploitation [9, 15, 60, 67]. This brings us to the central question for our paper: *What factors determine whether an IoT vulnerability is targeted?*

We create a dataset of IoT vulnerabilities published between January 2016 to June 2023 using VARIoT [43]. This results in a set of 23,373 IoT-related CVEs affecting hardware, software, and applications. Within this set, we need to distinguish between vulnerabilities that were observed to be targeted by attackers and those that were not. We meticulously collected instances of vulnerabilities that were observed to be targeted in the wild from seven data sources: VirusTotal [69], CISA's Known Exploited Vulnerabilities (KEV) Catalog [23], VulnCheck's Known Exploited Vulnerabilities (KEV) [71], AttackerKb [13], Google Project Zero [35], IoTPOT [55], and X-Pot [44] based honeypots.

Next, we set out to explore which features can help us distinguish between targeted and non-targeted vulnerabilities. We identified a lower-bound set of 848 targeted IoT-related CVEs. Subsequently, we extracted features for explaining the targeted set. We extended EPSS with features from the CrimeBB database [57], which contains scrapes of 36 different hacker forums. We found 699 IoT-related CVEs discussed in over 25k posts across 25 different forums, with targeted IoT vulnerabilities being more frequently discussed than non-targeted ones.

Based on these features, we create a model and evaluate its performance using the Brier score and McFadden's pseudo R^2 score. The Brier score measures prediction accuracy, while the R^2 score assesses the overall fit of the model. We incorporate features from hacker forum data, consolidated into a single 'Engagement' feature using PCA. Our model's McFadden's R^2 score increased by 21% with the inclusion of this feature, highlighting the value of hacker forum data in explaining the variance in targeted IoT vulnerabilities. Using the Brier score, our model, with 38 features (fewer than EPSS), improves prediction accuracy by 17% over the EPSS model.

The main contributions of the paper are:

- We develop a theoretical model explaining why certain IoT vulnerabilities are exploited, based on vulnerability, exploitability, and hacking community factors.
- We analyze discussions across 36 hacker forums, identifying 699 IoT-related CVEs mentioned in over 25,000 posts. We find that targeted IoT vulnerabilities are discussed more frequently than non-targeted ones, highlighting the influence of hacking forums on exploitation patterns and risk assessment.
- We develop a logistic regression model that improves EPSS prediction for IoT vulnerability exploitation by 17% by incorporating hacking forum features, which enhance the model's ability to explain variance in targeting by 21% and significantly impact pre-exploitation risk assessment.
- We publicly release the code for our exploit prediction model, ensuring transparency and reproducibility [1].

2 Related Work

Several studies have examined the relationship between CVSS scores and exploit-related factors. Allodi et al. [4] found that using CVSS alone to predict attacks is unreliable, comparable to random selection. This inconsistencies in severity ratings is supported by Wunder et al. [74]. Although the high CVSS scores do not reliably correlate with exploits in the wild, and excluding CVSS had minimal impact on results [19], only 9% of CVSS scores are available at disclosure [59], and nearly half of exploits occurring before CVSS publication [21].

On the other hand, the existence of PoC exploits was found to be a better risk factor [5], but it is not always a reliable indication of exploitation in the wild [4]. A 2020 study [39] found that a low rate, only 5%, of known vulnerabilities are exploited in the wild, consistent with prior research [1, 5, 10, 33, 51, 66, 73]. Yet, features within PoC exploits can be valuable for estimating exploitation frequency [2], predicting functional exploits [66], early exploitability [37], or predicting exploits in the wild [39].

In the realm of exploit prediction in the wild, various methodologies have been explored. For example, Twitter discussions can predict exploitation more accurately than CVSS [21] or PoC-based approaches [59], and help estimate exploitability scores [66]. Other studies have used vulnerability descriptions and online discussions to predict exploitation over time [9, 41, 66] or early exploitability of just disclosed vulnerabilities[37].

Previous studies have explored IoT-specific features to understand IoT targeting in the wild [1, 2, 10, 14, 46]. However, collecting

[1] https://doi.org/10.4121/69d111e7-30e6-4206-b27d-22a5f4e33722

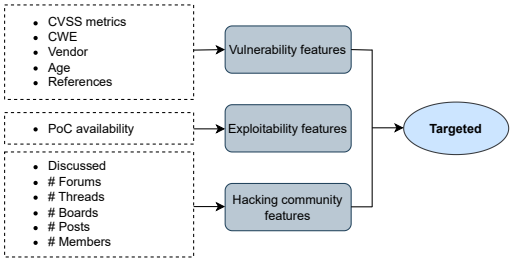


Figure 1: Theoretical Model.

these features, such as device types [2, 10, 11, 14, 18, 63] and the number of exposed devices online [2, 14, 54], is labor-intensive, not scalable, and requires manual effort. Since these features are not readily available online to scale for our dataset of over 23k CVEs and would require similar manual work from vendors and users, we do not include them.

The most comprehensive approach is the Exploit Prediction Scoring System (EPSS), a data-driven framework predicting if a vulnerability will be exploited [40]. EPSS incorporates various data sources including CVSS, vendor information, published exploit code, and references to external websites discussing the vulnerability. While achieving accurate exploitation estimates with a ROC AUC of 0.838, they aimed to improve by adding data like social media and vulnerability scans. In their subsequent version, EPSSv3, they incorporated those additional variables besides more vulnerability features such as CWE and age of vulnerability resulting in 82% performance improvement over previous models, which enriched the predictive capability of the model [41]. However, the lack of full disclosure of the EPSS model hinders its replication and use by others, preventing independent verification of its findings [64].

Our study advances beyond the state of the art by including data from underground hacking forums, specifically for IoT vulnerabilities. While using vulnerability-related features and PoC exploits as in [41], we expanded our approach by examining discussions within underground hacking forums, specifically targeting IoT vulnerabilities. For example, Tavabi et al. [67] found that dark web discussions outperforms state-of-the-art methods in exploit prediction and improved exploit prediction, with 14% of vulnerabilities mentioned being exploited. Data from dark web also had the highest exploitation rate among four datasets used to predict exploitation in the wild in [9]. Leveraging the CrimeBB dataset, which comprises 36 different underground forums, Moreno-Vera et al. [47] identified 1,498 unique CVEs discussed in these platforms, highlighting the value of incorporating hacker-used platforms for more accurate predictions of vulnerability targeting. This comprehensive approach, integrating insights from online discussions, and underground forums, enhances our understanding and prediction of cyber threats and exploitation activities [8, 15, 37, 59, 60].

3 Theoretical Model

To help answer what determines the targeting by attackers of IoT vulnerabilities, we draw upon previous studies that have investigated various factors and features influencing the selection of a vulnerability as a target. For instance, some studies have explored vulnerability features such as disclosure date [28, 41, 49], severity metrics [5, 28, 41], software and hardware weaknesses, vendors, and label references to advisories, solutions, and tools [28, 41], while others have examined the correlation between exploitation in the wild and the presence of PoC exploits [5, 41, 49, 59]. Additionally, the relationship between discussions of vulnerabilities in social media or underground forums and their exploitation in the wild has also been explored [8, 15, 57, 59, 60, 66]. Our aim is to consolidate these features from prior and recent studies into a comprehensive theoretical model. We develop a model to explain which IoT vulnerabilities will be targeted (Figure 1). Our model shares similarities with EPSS in integrating various factors that influence the targeting of certain vulnerabilities. However, while we focus exclusively on vulnerabilities within the IoT realm, EPSS adopts a broader approach, including vulnerabilities across all domains. Though our model draws inspiration from EPSS's approach to integrate factors derived from online sources to enhance vulnerability analysis and prediction, it extends this by incorporating features extracted from underground forums. Overall, our model consists of three factors: vulnerability features, exploitability features and hacking community features (see Figure 1).

Vulnerability features. Vulnerability features such as disclosure date, CVSS metrics, software and hardware weaknesses, vendors, and label references to advisories and other external websites might predict which vulnerability is likely to be targeted by attackers. Jacobs et al. [41] identified 30 influential features, including count of references to external websites, vulnerability age, specific vendors, and CVSS submetrics, as significant predictors of exploit likelihood. Thus, we integrate vulnerability features into our model.

Exploitability features. The relationship between the existence of PoC exploit code and exploitation in the wild has been studied in previous work [5, 41, 49, 59]. Thus, we include it as feature.

Hacking community features. Discussions of vulnerabilities on social media [21, 59, 66] and underground forums [8, 15, 60] have been found to provide more accurate predictions of vulnerability exploitation than vulnerability features or PoC exploits alone. However, using Twitter feeds was not listed among the top 30 most influential features in EPSS [41]. In fact, Tavabi et al. [67] found that utilizing discussions from the dark web for exploit prediction outperforms state-of-the-art methods, including social media feeds. Thus, we incorporate hacking community features using underground hacking forums to explain whether discussions of IoT vulnerabilities on such platforms can increase the likelihood of these vulnerabilities being targeted. For example, Moreno-Vera et al. [48] identified 1,498 unique CVEs discussed within these underground hacking forums, and Almukaynizi et al. [9] found that 14% of vulnerabilities mentioned on the dark web and deepweb were observed being exploited in the wild. This dataset had the highest rate of exploitation among the four datasets used in the study, highlighting the value of incorporating hacker-used platforms for discussions of vulnerabilities on underground forums.

4 Methodology

The focus, first, lays on gathering data on all IoT vulnerabilities from January 2016 to June 2023, driven by the notable surge in attacks on IoT devices following the 2016 release of the Mirai botnet source code [12]. Next, we collect data on which subset of the vulnerabilities were targeted in the wild. We then distinguish between

Table 1: Description of variables in the theoretical model.

Variable	Description	Count	Type	Data sources
Targeted IoT vuln.	List of IoT exploits in the wild	1	Binary	CISA, Virustotal, IoTPOT, X-Pot AttackerKB, VulnCheck KEV Google Project Zero
CVSSv3 metrics	Measuring the impact and exploitability of a vulnerability	23	Multi-level	NVD
CWE	List of software and hardware weaknesses	111	Binary	NVD
Vendor	List names of affected vendors	212	Binary	NVD
Age of vulnerability	Number of days from the CVE publication date until the day of the analysis	1	Numeric	MITRE CVE list
References with labels	Number of references with each of the reference labels	19	Numeric	MITRE CVE list, NVD
PoC exploits	Denote whether the PoC exploit is available or not	1	Binary	Exploit-DB
Discussed	Denote whether a certain IoT CVEs was discussed or not	1	Binary	CrimeBB
#Forums	Number of forums in which a particular IoT CVE was discussed	1	Numeric	CrimeBB
#Boards	Number of boards in which a particular IoT CVE was discussed	1	Numeric	CrimeBB
#Threads	Number of threads in which a particular IoT CVE was discussed	1	Numeric	CrimeBB
#Posts	Number of posts in threads in which a particular IoT CVE was discussed	1	Numeric	CrimeBB
#Members	Number of members involved in threads in which a particular IoT CVE was discussed	1	Numeric	CrimeBB

targeted and non-targeted vulnerabilities based on observed attack data. Subsequently, we gather features categorized into vulnerability, exploitability, and hacking community, all listed in Figure 1. Finally, we employ a regression model to assess the significance of each feature in influencing IoT vulnerability targeting. This allows for a deeper understanding of the threat landscape and facilitates recommendations for enhancing the security of IoT systems.

4.1 Collecting IoT vulnerabilities

We identify targeted and non-targeted IoT vulnerabilities through a two-step process. First, we collect IoT vulnerabilities from the VARIoT dataset [68], covering vulnerabilities from January 2016 to June 2023. Second, we determine which were targeted using seven sources: (i) VirusTotal [69], (ii) CISA KEV [23], (iii) VulnCheck KEV [71], (iv) AttackerKB [13], (v) Google Project Zero [35], and honeypots such as (vi) IoTPOT [55] and (vii) X-Pot [44].

We then classify vulnerabilities by subtracting the targeted IoT vulnerabilities from the full list, generating a separate set of non-targeted vulnerabilities.

IoT vulnerabilities. We focus exclusively on IoT-specific CVEs using VARIoT [43], which identifies IoT-related vulnerabilities based on its broad definition: "an item (except a phone, PC, tablet, or data center hardware) with network connectivity and data exchange capabilities"[42]. While acknowledging VARIoT's limitations in accurately classifying IoT-related vulnerabilities, we find its comprehensive repository valuable for our research objectives, as demonstrated in various studies [27, 58, 62].

From VARIoT, we collect all CVE-IDs published between January 2016 and June 2023 to align with the spread of the Mirai source code in 2016 [12]. Some IoT vulnerabilities in VARIoT lack CVE-IDs from NVD, often originating from other sources like CNVD. In line with EPSS, we only include vulnerabilities with CVE-IDs. We identified 23,373 CVEs classified as IoT vulnerabilities in VARIoT.

Targeted IoT vulnerabilities. To identify targeted IoT vulnerabilities, we use VARIoT's comprehensive CVE list and cross-reference it with seven different data sources. Of course, not all attacks are observed in our datasets, or in any datasets, meaning our data provides a lower bound set of targeted CVEs. Additionally, the time frame covered by some datasets does not encompass the entire period from January 2016 to December 2024, during which we collected IoT vulnerabilities.

First, we employed a VirusTotal search query for binaries tagged with a CVE-ID in our VARIoT dataset (*tag:cve-**) [70]. We found 526 instances out of 23,373 total CVEs. This count, however, could be impacted by VirusTotal's search constraints, limited to vulnerabilities targeted within 90 days from our search date. Second, we use the KEV Catalog. This list includes 1,026 CVEs. There is an overlap of 218 CVEs with the 23,373 CVEs that were classified as IoT by VARIoT. The third source was VulnCheck's Known Exploited Vulnerabilities (KEV) [71]. This list includes 3,089 with 426 overlapping CVEs that were classified as IoT by VARIoT. This makes VulnCheck KEV the source with the highest number of IoT CVEs exploited in the wild among our datasets. Next, we used the 23K CVEs to fetch data via cURL and then scraped AttackerKB [13] to collect CVEs tagged as "Exploited in the Wild". We identified 312 out of 23,373 total IoT CVEs. For the fifth source, we leverage Google Project Zero [35], which provides a public list of 325 zero-day CVEs exploited in the wild. Of these, 45 IoT CVEs overlapped with the 23K CVEs in VARIoT. The last two datasets are both from honeypots, IoTPOT [55] and X-Pot [44]. We gathered a list of targeted vulnerabilities and then identified which of these vulnerabilities overlapped with out set of VARIoT CVEs. In IoTPOT logs, collected from September 2018 to September 2021, we observed 21 targeted IoT CVEs. In X-Pot logs, collected from July 2019 - October 2023, we identified 114 targeted IoT vulnerabilities.

We identify 848 unique targeted IoT vulnerabilities within these seven datasets. Thus, we treat the remaining 22,525 CVEs out of the total set of 23,373 CVEs as non-targeted IoT vulnerabilities.

4.2 Collecting Explanatory Factors

To assess the influence of the explanatory factors in the theoretical model (section 3), we identify and gather features that can serve as proxies for those factors. In this section, we outline the process of collecting these features, which are summarized in Table 1.

Vulnerability features. We use the NVD [50] dataset to collect five vulnerability features following the methodology used in [41].

CVSSv3 metrics: Given our dataset's focus on IoT vulnerabilities from 2016 onward, we use CVSSv3 [31] to derive base metrics, split into exploitability and impact. Exploitability includes five components: attack vector, attack complexity, privileges required, user interaction, and scope. Impact covers confidentiality, integrity, and availability. Except for 33 CVEs lacking NVD scores, we applied one-hot encoding to all sub-metrics, resulting in 22 variables. We also

include the overall CVSSv3 base score as a single binary variable, representing the combined score of these sub-metrics.

CWE: We collected the Common Weakness Enumeration (CWE), a list of software and hardware weakness types that can become vulnerabilities [53]. We collect 240 different CWE for 23,317 CVEs. The remaining 95 CVEs lacked CWE assignments due to being unavailable or private. We created binary variables for the 111 CWE categories associated with 10 or more vulnerabilities for modeling (see section 7).

Vendor: As some vendors might be more attractive to attackers than others, we use the Common Platform Enumeration (CPE) from NVD to extract only the vendor name(s). We collect vendors for all but 88 CVEs due to the CVEs being marked as "RESERVED" or "REJECT". We did not alter or fix any typos or misspellings within the records. Out of 1,423 unique vendors we found, only 212 vendors with 10 or more CVEs are included in our model, as including vendors with fewer CVEs did not improve performance, according to [40]. Then, we created a binary variable for each vendor and added a feature to measure the total number of vendors per CVE to evaluate whether CVEs affecting more vendors are more likely to be targeted.

Age: The age of a vulnerability may affect its likelihood of exploitation. To test whether older vulnerabilities might be less appealing to attackers due to reduced vulnerable population [41] within the IoT realm, we used the publish date from MITRE [52] to calculate the age — the number of days between the publish date and our feature extraction date. We collected the publish dates for all vulnerabilities but 56 that were marked "RESERVED" or "REJECT".

References: We measure level of activity and analysis related to vulnerabilities in [41] by quantifying the number of references linked to each CVE in MITRE CVE list [52], excluding 423 CVEs marked as "RESERVED," "REJECT," or those without an assigned label to the hyperlink. Furthermore, also collect the unique reference labels assigned to these CVEs listed by NVD, such as Third Party Advisory, Vendor Advisory, and VDB Entry. We identify 18 distinct labels, which we use as binary features.

Exploitability features. We use the exploit database Exploit-DB [30] to collect PoC availability. While other sources such as Metasploit [45] and GitHub [34] also host PoC exploits, Exploit-DB remains one of the best coverage source for PoC exploits [39].

PoC availability: We evaluate the exploitability risk of vulnerabilities based on their availability. We only identify PoC exploits for 890 (3.8%) out of the 23,373 CVEs, indicating a relatively limited presence of publicly available exploits for these vulnerabilities.

Hacking community features. To analyze IoT vulnerability discussions on underground forums, we use the 2023 CrimeBB dataset [57], provided by the Cambridge Cybercrime Centre [20]. It includes 36 forums in English, Russian, and Spanish; however, "Hackers Armies" was unavailable, leaving 35 forums in our analysis. These forums serve as platforms for exchanging ideas and participating in various activities, some of which are illegal. They follow a structured format with boards covering topics from hacking to marketplaces (see Appendix Table 7). Each board contains threads on specific subjects, composed of posts by members. We extract six features from these discussions:

Discussed: To measure whether an IoT vulnerability was discussed within the underground forums, we match post content with CVE-IDs from our dataset. For this matching, we allow for variations in capitalization.

Number of forums: We assess the popularity of a vulnerability across the forums by counting the number of forums in which the IoT vulnerability was discussed.

Number of boards: We count how many boards discussed specific IoT CVEs to understand the range of topics and interests they span. In total, these vulnerabilities appeared on 122 boards across 25 forums.

Number of threads: Counting distinct threads discussing each IoT vulnerability helps quantify its activity level. We identified 735 unique threads discussing IoT-related vulnerabilities.

Number of posts: Assuming more active discussion indicates higher risk of real-world targeting, we count posts in threads mentioning IoT CVEs. We found 25,782 posts, including both specific and general discussions involving IoT CVEs.

Number of members: To measure engagement level around IoT vulnerabilities, we counted distinct members involved in related discussions, using the same method as for post counts. A total of 5,129 users participated in threads specifically about or generally on IoT CVEs.

Similar to EPSS, we did not perform experiments to add or filter out intentional misinformation or automated posts in underground forum discussions. However, prior research [9, 59] shows that the introduction of misinformation had minimal impact on exploit prediction accuracy. The work in [9] demonstrated that exploit prediction models remain robust against adversarial noise in discussions from underground forums, where structural barriers such as account verification, skill demonstrations, and reputation systems make large-scale data poisoning far more difficult, compared to social media platforms like Twitter/X. Even if substantial noise would be added (e.g., 20% noise in testing data), this leads to minimal degradation in performance (AUC scores above 0.87). In [59], the authors indicate that adversaries can manipulate public platforms like Twitter/X to poison classifiers, but the impact on precision is limited, with a drop to only around 20% under the most sophisticated attack scenarios. In sum the impact is limited. We would also like to note that the main impact of misinformation would be to increase the false positive rate: more CVEs are being discussed on the forums, and hence some might be predicted as higher risk, yet in reality they will not actually be attacked. A somewhat higher false positive rate does not directly undermine the model's utility, as prioritization remains valuable for defenders. Out of thousands of CVEs, the model predicts a small subset as likely to be exploited. Even if the latter set would double in size because of false positives, it still helps defenders to prioritize a small fraction of a total set of thousands.

5 Vulnerability and Exploitability

In this section, we present the findings of vulnerability characteristics and their PoC exploit code among both targeted and non-targeted IoT vulnerabilities.

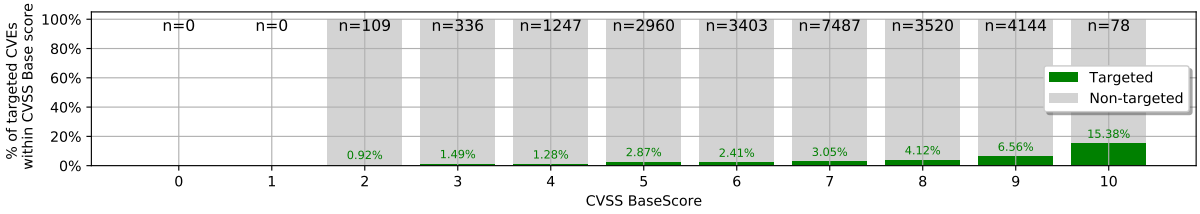


Figure 2: Proportion of IoT vulnerabilities within each CVSS base scores.

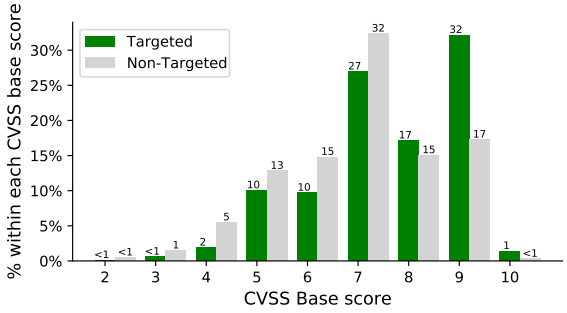


Figure 3: Targeted and non-Targeted IoT vulnerabilities within each CVSS base score.

5.1 Vulnerability features

CVSSv3 metrics. The CVSS base score, ranging from 0 to 10, is derived from Exploitability and Impact metrics, assessing how easily a vulnerability can be exploited and its potential consequences [31].

Analyzing targeting CVEs rates per CVSS score (Figure 2), 15.38% of CVEs with a score of 10 were targeted, but lower scores also saw targeting, indicating a weak correlation between score and targeting. Attackers often choose lower-scored vulnerabilities, despite many high-severity, with scores 9 and 10, remaining unused.

In fact, in absolute numbers, the attackers have chosen many more vulnerabilities with low scores, as we can see in Figure 3. Figure 3 shows the proportion of targeted and non-targeted IoT vulnerabilities for each CVSS base score. While higher CVSS scores (e.g., 9 and 7) account for a significant share of targeted vulnerabilities, lower-scored vulnerabilities (e.g., 5 and 6) are also targeted. For example, 20% of the total targeted vulnerabilities have a CVSS base score of 5 or 6. Also, many high-severity vulnerabilities remain untargeted. Only about 1% of total targeted vulnerabilities have a base score of 10, while the largest subset (32%) has a base score of 9. The average base score for targeted CVEs is 7.5, only slightly higher than 6.9 for nontargeted CVEs. This indicates that adversaries do not solely focus on high-severity vulnerabilities but may also target medium-severity ones, likely considering other factors such as ease of exploitation, availability of PoC exploits, and real-world impact. In short, this suggests that CVSS has weak predictive power, in line with earlier research [6].

Examining the distribution of IoT vulnerabilities within CVSS sub-metrics (Figure 4), we begin with exploitation metrics (left). The majority of targeted CVEs can be remotely exploited, with 67.15% using the network as the attack vector (AV:N). This potentially impacts attack complexity, as nearly 94% exhibit low attack complexity (AC:L), while 66.2% require no privileges (PR:N), and 74.69% do not require user interaction (UI:N). This confirms previous findings that attackers tend to target vulnerabilities with low attack complexity [7].

That said, we do see that non-targeted CVEs have almost the same distribution. In other words, many more CVEs also possess these features and, yet, they are not targeted. We see the same thing for the impact metrics (right side of Figure 4). This all aligns with previous findings that CVSS alone is not an exhaustive predictor on its own for exploitation in the wild [2, 4]. We now see why: these features do not effectively differentiate targeted from non-targeted vulnerabilities. It is unclear whether this is because experts are not reliable in rating these features, or the features themselves are not that relevant to attacker's targeting decision.

Common Weakness Enumeration (CWE). The hardware and software weaknesses in vulnerable systems are categorized using the CWE framework. Among our 23k IoT-related vulnerabilities, we identified 240 CWE categories, with four—CWE-178, CWE-782, CWE-917, and CWE-1336—appearing only in targeted CVEs. All were observed once except CWE-917, which appeared in three vulnerabilities. This may reflect their rarity rather than attacker preference.

We found 88 CWE categories in both targeted and non-targeted CVEs, with the top 20 distributions shown in Figure 5. Among these, CWE-78 (OS Command Injection) was the most common in targeted CVEs (13.6%), aligning with prior research identifying it as a key infection vector [1, 2, 10]. CWE-787 (Out-of-Bounds Write) ranked second (9.16%) and was also the most frequent in non-targeted CVEs for nearly 10%.

In sum, like CVSS scores, CWE categories provide only a weak signal of attacker preference. To put it differently, a random selection of vulnerabilities would yield a similar CWE distribution to that for the observations of targeted vulnerabilities.

Vendors. Examining the associations between vendors and CVEs, we found only 280 out of 1,423 vendors were linked to targeted CVEs, while 1,143 remained untargeted. Figure 6 shows that vendors with more CVEs tend to have more targeted vulnerabilities. For instance, Apple and Cisco lead with 145 and 110 targeted CVEs, respectively, while 45 vendors have only one targeted CVE.

On average, these 280 vendors have 5.9 targeted CVEs (SD = 15.5). Among untargeted vendors, Qualcomm has the most with 507 CVEs, followed by IBM with 254. In short, we see a linear relationship – more CVEs overall will result in more targeted CVEs – with quite some variance.

In addition, targeted CVEs also tend to be associated with more vendors—1.9 per CVE on average, compared to 1.2 for non-targeted ones. This suggests vendor involvement could influence attacker

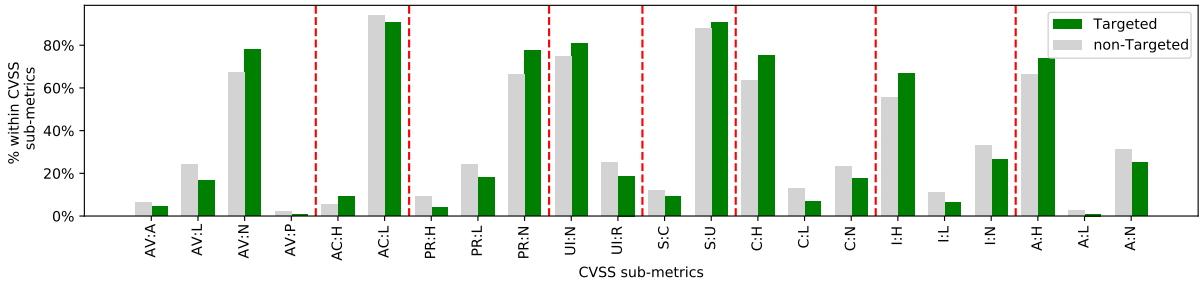


Figure 4: Targeted vs non-Targeted IoT vulnerabilities for each CVSS sub-metrics.

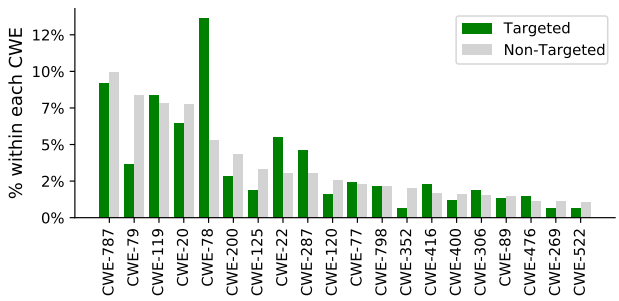


Figure 5: The top 20 CWE categories among targeted and non-targeted IoT vulnerabilities, ordered by CWE the with highest number of vulnerabilities.

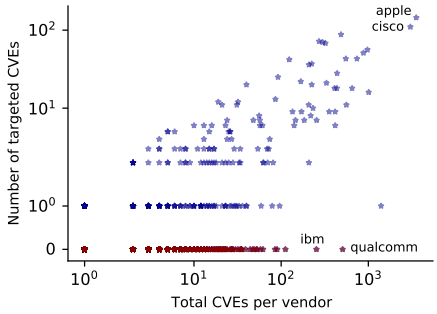


Figure 6: Relationship between targeted and non-targeted CVEs vs. total CVEs per vendor.

selection. Perhaps more vendors per CVE indicates a larger install base of vulnerable systems.

Age. The age of a vulnerability could potentially influence its likelihood of being exploited. We measure the age by the number of days between the vulnerability published date and the time of our feature extraction. The overall average age of all IoT vulnerabilities included in this study is 1,430.5 days.

For targeted vulnerabilities, the average age is slightly lower at about 1,393 days, compared to 1,432 days for non-targeted ones.

While the age of a vulnerability may influence its likelihood of being targeted, there is no significant difference in the average age between targeted and non-targeted CVEs. Despite slight variations, both categories exhibit a wide range of ages, indicating that age alone may is not clear factor for targeting within the IoT vulnerability landscape.

References. The results presented in Table 2 shows the distribution of all 18 reference labels across IoT vulnerabilities. Both targeted and non-targeted CVEs share the top five labels, but these labels appear more frequently in targeted CVEs, except for "Vendor Advisory", which shows only a slight increase. This suggests that targeted vulnerabilities receive more vendor attention.

This observation is further emphasized by the fact that the "Patch" label appears twice as often in targeted CVEs, and "Exploit" label is present in 38.06% of targeted CVEs compared to 18.65% in non-targeted CVEs. Causality might be reversed here: because these CVEs are targeted in the wild, vendors might feel more pressure to develop patches and the attacks might trigger the release of public PoC exploit code available. This can only be resolved by looking at whether a label predates the attacks or the other way around.

Other labels, like "US Gov. Resource", "Mailing List", "Mitigation", "Release Notes", and "Issue Tracking", are more common in targeted CVEs, though many labels show little difference compared to non-targeted ones. These commonalities suggest many similarities across both targeted and non-targeted CVEs within the IoT landscape. A final difference to observe the average count of labels per CVE for targeted versus non-targeted vulnerabilities. For targeted CVEs, this stands at 3, compared to 2.1 for non-targeted CVEs. Both groups share the same range, with a maximum of 10 labels per CVE and a minimum of 1. This supports the idea that the number of references plays a role in exploitation likelihood, aligning with EPSS research [41].

5.2 Exploitability features

We looked at the presence of PoC exploits to see if it can explain attackers' decision on targeting certain vulnerabilities. We found public PoC exploits for only 890 IoT vulnerabilities out of the total 23,373. This aligns with earlier work which found that around 4% of vulnerabilities get associated public exploits within a year [36].

Only 169 (18.9%) of the PoC exploits were associated with targeted CVEs, while the majority 721 (81%) pertained to non-targeted CVEs. This means that out of the 848 targeted vulnerabilities, only 19.9% have public PoC exploits. This reinforces findings from prior research indicating that the existence of PoC exploits is not the main driver of attacker's selection of vulnerabilities. It seems that most develop their own exploit code or at least get it from another source than PoC exploits being published [4, 5].

Table 2: Summary of total reference labels in targeted and non-targeted vulnerability.

Reference label	Non-targeted	Targeted
Vendor Advisory	15,842(71.63%)	554(66.51%)
Third Party Advisory	11,822(53.45%)	608(72.99%)
VDB Entry	5,566(25.17%)	299(35.89%)
Exploit	4,124(18.65%)	317(38.06%)
Patch	2,907(13.14%)	214(25.69%)
US Gov. Resource	1,693(7.65%)	81(9.72%)
Mailing List	1,265(5.72%)	146(17.53%)
Mitigation	878(3.97%)	65(7.80%)
Release Notes	756(3.42%)	67(8.04%)
Broken Link	670(3.03%)	39(4.68%)
Technical Description	528(2.39%)	21(2.52%)
Issue Tracking	421(1.90%)	74(8.88%)
Product	251(1.13%)	13(1.56%)
Permissions Required	232(1.05%)	8(0.96%)
Not Applicable	154(0.70%)	19(2.28%)
Press/Media Coverage	21(0.09%)	7(0.84%)
URL Repurposed	4(0.02%)	2(0.24%)
Tool Signature	2(0.01%)	1(0.12%)

6 Hacking Community

Analyzing posts from the CrimeBB dataset, we identified discussions about IoT vulnerabilities and their potential exploitation within the hacking forums. Our analysis uncovered over 25k posts discussing 699 unique IoT vulnerabilities, as summarized in Table 3.

We observed that the discussions that mention IoT vulnerabilities within these underground hacking forums primarily focused on a wide range of exploit-related subjects. Notably, there was a substantial presence of inquiries regarding PoC exploits for certain IoT vulnerabilities, as well as requests for help or hiring to exploit such vulnerabilities with marketplace sections dedicated to buying, selling, or exchanging hacking-related products and services. Other discussions involved shared exploit kits, scripts, and tutorials. The forums also featured discussions on security advisories and attack news, including vulnerabilities utilized in zero-day exploits and ransomware attacks. It is evident that these forums serve as significant hubs for discussing IoT vulnerabilities and potential exploits.

6.1 IoT vulnerabilities in hacking forums

We used post content to identify IoT vulnerabilities discussed in CrimeBB using their CVE-ID (see Section 4.2). These discussions covered 699 unique IoT vulnerabilities, of which 222 were targeted. These vulnerabilities were discussed in 25 different forums, as detailed in Table 3. In total, 93 boards, 735 threads, 25,760 posts, and 5,101 members were involved in these discussions.

We found 93 different boards discussing IoT CVEs. Interestingly, while "hack-forums" was not the forum with the highest number of discussed IoT CVEs, it was the forum with the highest number of 25 different boards, suggesting a significant diversity of IoT vulnerabilities discussed across various boards compared to other forums. Following closely behind is the "xss-forum" with nearly a similar number of boards, totaling 24.

The predominant IoT vulnerability identified across 16 boards was CVE-2016-5195, a targeted vulnerability also known as Dirty COW, a Linux kernel vulnerability allowing attackers to gain write access to read-only memory mappings, potentially leading to privilege escalation.

The variation in thread activity highlights the dynamic nature of discussions surrounding IoT vulnerabilities. For instance, the "xss-forum" forum exhibited the highest number with 215 unique threads, indicating a significant focus on IoT vulnerabilities within that community.

Forums with the highest volume of posts do not always feature the most discussions on IoT vulnerabilities, e.g., despite "hack-forums" including over 42 million posts, only 1,021 were related to IoT vulnerabilities, with just 55 focusing on specific CVE(s) (Table 7). In fact, the "antichat" forum led with 19,202 posts on IoT-related CVEs, with a total volume of posts exceeding 2.6 million. "safe-sky-hacks" had only one post discussing an IoT CVE. CVE-2018-10561 (a vulnerability in the DSL-2640B router) garnered the highest total posts at 8,078, in both specific and general discussions, while CVE-2022-40684 had the most specific posts at 115, noted as targeted.

But when looking at forums discussing specific IoT CVEs in Table 3, only 11 out of the 25 forums exhibited such discussions. Among these, the "breached" forum showed the highest number of posts, with a total of 158 posts. Interestingly, while general discussions were prevalent across all forums except for one, namely "cracked" where all 47 posts were exclusively linked to specific IoT CVE, it's noteworthy that "antichat" forum recorded the highest number of posts addressing IoT CVEs within general threads. Nevertheless, despite that, "antichat" forum did not feature any posts within threads focusing on specific IoT vulnerabilities.

We analyzed forum posts mentioning IoT vulnerability using their CVE-IDs to shed light on the extent of discussions. The "xss-forum" had the most posts, with CVE-2020-1472 (critical vulnerability in Microsoft's Netlogon) and CVE-2018-13379 (critical vulnerability found in Fortinet FortiOS SSL VPN) being mentioned 27 times each. Both were observed as targeted in our dataset.

A total of 5,101 individuals participated in threads discussing IoT vulnerabilities using their CVE-IDs. Of these, 345 were involved in threads dedicated to specific IoT CVEs, while the remaining 4,756 contributed to general discussions. This correlates the volume of posts with the number of participants, with forums having more posts also attracting more members. For instance, "xss-forum" led in both post volume and member engagement.

6.2 Targeted vs non-targeted hacking discussions

The distribution across forums where IoT vulnerability discussions took place made it apparent that among the 25 forums, 13 of them predominantly focused on targeted CVEs over non-targeted ones (see Table 3). Conversely, 10 forums showcased a greater inclination toward non-targeted vulnerabilities while the remaining two forums exhibited an equal number of posts discussing both targeted and non-targeted vulnerabilities. Among these forums, "underc0de" emerges as the most active in discussing IoT vulnerabilities, with a total of 289 IoT vulnerabilities, including the highest number of targeted vulnerabilities at 121. That was also reflected in the number of threads. Regarding the most observed IoT vulnerabilities across forums, CVE-2017-5754 (`Meltdown`, hardware vulnerability affecting modern microprocessors [26, 56]) and CVE-2017-5715 (`Spectre Variant 2`, vulnerability that exploits speculative execution in modern microprocessors [26, 56]) were the most prevalent, appearing in 9 forums and were also identified as targeted vulnerabilities.

Table 3: Summary of forums discussed IoT vulnerabilities.

Forum	#Boards	#Threads	#Posts			#Members			#CVEs	
			Specific	General	Total	Specific	General	Total	Targeted	Total
antichat	4	67	0	19202	19202	0	2287	2287	60	135
blackhatworld	2	3	0	11	11	0	9	9	4	5
breached	10	21	158	817	975	151	596	747	13	16
cracked	1	1	47	0	47	47	0	47	1	1
dread	1	1	0	2	2	0	2	2	4	4
elhacker	7	71	10	120	130	6	35	41	45	139
forum-team	3	4	0	9	9	0	4	4	3	9
freehacks	1	1	0	2	2	0	2	2	4	5
garage-for-hackers	1	2	1	1	2	1	1	2	2	2
greysec	2	4	3	41	44	2	20	22	2	4
hack-forums	25	73	55	1021	1076	37	472	509	38	78
ifud	1	1	0	9	9	0	8	8	1	2
kernelmode	1	2	0	13	13	0	6	6	0	2
lolzteam	5	8	0	59	59	0	41	41	7	15
offensive-community	2	2	1	1	2	1	1	2	2	3
probiv	6	59	0	342	342	0	26	26	27	99
raidforums	8	20	11	788	799	10	529	539	18	31
runion	2	2	0	224	224	0	93	93	3	4
safe-sky-hacks	1	1	0	1	1	0	1	1	1	1
torum	4	5	0	16	16	0	11	11	9	12
underc0de	6	163	8	172	180	5	23	28	121	289
unknowncheats	3	6	5	100	105	4	61	65	4	5
v3rmillion	1	2	0	5	5	0	5	5	0	2
xss-forum	24	215	154	2340	2494	81	599	680	103	255
zismo	1	1	0	33	33	0	18	18	1	1
Total (unique)	93	735	453	25308	25760	345	4824	5101	699	222

Table 4: Distribution of hacking community features per IoT CVE, split between targeted CVEs/non-targeted CVEs.

	#Forum	#Boards	#Threads	#Posts			#Members		
				Total	Specific	General	Total	Specific	General
Average	20.57\32.3	5.59\10.75	1.5\2.25	706\660	16\3	690\658	153.24\93.48	12\2.08	143.24\91.8
Min.	1\1	1\1	1\1	1\1	1\1	1\1	1\1	1\1	1\1
Max.	121\168	117\158	12\26	12,828\14,519	140\25	12,828\14,519	1,646\1,444	133\18	1,646\1,444

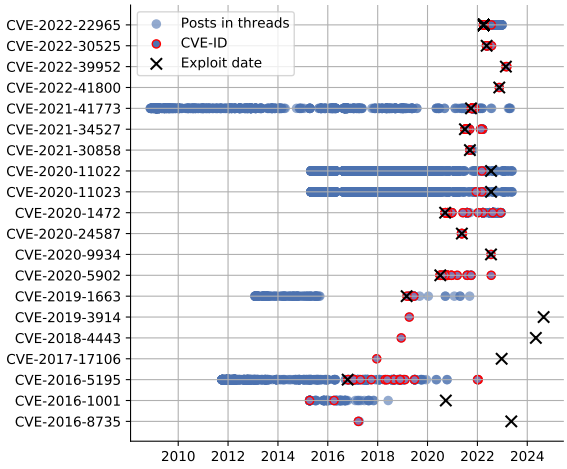


Figure 7: Top 10 IoT CVEs for both the longest and shortest time-to-exploit after post discussing CVE-ID, number of days between the first discussion mentioning the CVE-ID and its first observed exploitation in the wild, ordered by CVE publication year. The ● highlights posts mentioning CVE-IDs.

We measured the number of forums, boards, threads, posts and members per IoT vulnerability in Table 4. Overall, targeted vulnerabilities show higher averages compared to non-targeted ones across all features in the hacking community. For instance, the average number of forums per targeted IoT vulnerability is 20.57, which decreases to 5.59 at the board level and further drops to 1.5 at the thread level. Conversely, the average number of forums, boards and threads per non-targeted CVE are 32.3, 10.75, and 2.25, respectively. Similarly, targeted vulnerabilities received more posts on average, 706 compared to 660 for non-targeted ones, with specific targeted vulnerabilities garnering more discussion than general ones with 16 posts compared to 3 for non-targeted ones. Moreover, discussions on targeted vulnerabilities involved 153.24 members on average, which drops to 93.48 for non-targeted ones, with specific targeted vulnerabilities attracting more participants than general discussions, involving 12 members, which is around three times the number of participants engaged in non-targeted discussions that stands at 2.08 members.

6.3 Time to exploit in the wild

We measure "Time to Exploit" as the number of days between the first discussion of an IoT-related CVE in hacking forums and its first observed exploitation in the wild. To reduce false positives, we exclude posts that did not explicitly mention the CVE-ID in the same thread. However, Figure 7 presents both cases for the Top 10 IoT CVEs with the longest and shortest time-to-exploit, where ● highlights posts mentioning CVE-IDs, and ● highlights posts in the same thread. Due to missing or unclear exploit dates in some sources, we relied on five datasets: AttackerKB, VulnCheck KEV, Google Zero Project, IoTPOT, and X-Pot. Among 222 discussed and targeted IoT CVEs, 159 had available exploit dates. Of these, 78 (49.05%) were discussed before or on the same day as their first exploitation, while 81 were discussed afterward. The time gap between discussion and exploitation varies, with an average time-to-exploit of 482.6 days. The shortest time-to-exploit is zero days, exploitation occurred on the same day as forum discussion, were observed in

Table 5: Brier score and R^2_{MF} scores for all models.

Model	Brier score	R^2_{MF}
No Engagement feature	0.031	0.183
Include Engagement feature	0.028	0.222
EPSS scores	0.035	-0.157

seven cases all for CVEs published between 2020 and 2022. The longest time-to-exploit is 2,240 days (CVE-2016-8735, an Apache Tomcat vulnerability). Notably, CVEs with the shortest time-to-exploit had just 1-5 posts mentioning their CVE-ID, while those with the longest delays were mentioned only once, except for three cases appearing up to four times. These findings underscore the importance of monitoring hacking forums for early warnings on IoT vulnerabilities, allowing organizations to identify and mitigate threats before widespread exploitation.

7 Explaining exploitation

7.1 Model and feature selection

In the first version of EPSS, Jacobs et al. use logistic regression to build an exploit prediction system [40], showing that the features used share a somewhat linear relationship. EPSSv2 and EPSSv3 adopt the more complex XGBoost model [22] to capture non-linear relationships [41], sacrificing interpretability for improved predictive accuracy.

Our goal is to analyze and explain the factors influencing whether an IoT vulnerability is targeted by attackers, without relying on time-based patterns. To avoid added complexities, we use cross-sectional data, capturing targeting information at any given point. By choosing a logistic regression model, we can easily make predictions while also clearly identifying which features contribute to those predictions, allowing us to both predict vulnerability targeting and explain the significance of the influencing factors.

Upon further investigation of our potential feature set, we find a collection of linearly dependent features. Including such features in a model can introduce multicollinearity into the eventual model, undermining the statistical significance of the other features [3], leading to an overfitted model. We exclude the eight linearly dependent features from the dataset. That leaves a set of 365 features with which to fit the model.

Jacobs et al. [40] use elastic net regularization (among other methods) to condense their collection of 3,587 features down to 16 though. We aim to reproduce as much as the original methodology to maintain a level of comparability, and to be able to accurately judge the value of the additional hacking community features. Thus, we similarly use elastic net regularization for our feature selection.

We first split the dataset features into two groups: (1) hacking community features and (2) all other features. We use two feature selection methods, both applying elastic net regularization to select the most important features in (2). The first method trains a model using only the regularized features from (2), serving as a control. The second method adds these steps: (i) perform PCA on the hacking features, (ii) select the principal component (PC) with the highest variance, (iii) add this PC to the dataset, and (iv) train the model with the updated dataset.

PCA helps summarize many correlated features [16] while retaining key information, making it useful for our analysis. We select the most significant PC, labeled **Engagement**, to simplify further analysis. This PC accounts for 64% of the variance in hacking community features. As in [40], we tune the hyperparameters by performing a grid search for the α and λ that minimize the Bayesian Information Criteria (BIC). We find the following optimal values: $\alpha = 0.3$ and $\lambda = 0.005$. The regularization selects 93 features (see Figure 8).

7.2 Model evaluation

We split our dataset into a training and testing set using a split of 0.33. After regularization and PCA processing, we fit two logistic regression models using the training set according to the different feature selection approaches described in subsection 7.1. For each approach, we then use each model to estimate the probability that an IoT vulnerability in the test set will be targeted by malicious actors. To evaluate model performance, we compute several metrics: the Brier score and McFadden's pseudo R^2 (R^2_{MF}) score (see Table 5 for the full list of metrics). We also calculate the models' Brier score andR^2_{MF} score performance improvement (PI) as a percentage increase (see [75].)

The Brier score measures the mean square error of probabilistic predictions, i.e., how close the predictions are to the actual value. The R^2_{MF} score—being a normalized version of the MSE—is scale-independent and is more concerned with the general quality of the fit of the model. The closer the Brier score is to 0, the better. Contrarily, the closer the R^2_{MF} score is to 1, the better.

We compare the metrics with the EPSS scores we collected using EPSS API [32] for all vulnerabilities in our dataset except for 88, of which two were observed as targeted. This is because these CVEs are stored in the NVD but labeled as "REJECT" or "RESERVED" in the CVE List and do not appear in search results. EPSS predicts the likelihood of exploitation within a 30-day window [41], so we collect EPSS scores for our set of vulnerabilities for 1 September 2023, 30 days before the last day of data collection (1 October 2023). This allows for a consistent evaluation and comparison with EPSS.

The poor performance of EPSS reiterates its unsuitability for IoT-specific vulnerabilities. Both the control and Engagement models, despite using fewer features, outperformed EPSS on the same set of vulnerabilities. For example, the Brier score for both of our models, performed similarly, with scores of 0.031 and 0.029, which represent an improvement over EPSS's score of 0.035, though the difference is small in absolute terms. Furthermore, when comparing the prediction accuracy of our second model with the Engagement feature to EPSS using their Brier scores, we obtained a 17% performance improvement (PI) over EPSS.

Not only did the model with the Engagement feature perform best in Brier score, but also had a 21% PI in the R^2_{MF} score when comparing it with the control model. This highlights the value of extracting insights from the hacking community (see Figure 9).

While Brier scores offer insight into prediction accuracy, the R^2_{MF} scores more clearly differentiate the models. The control model had an R^2_{MF} score of 0.183, whereas the model with the Engagement feature achieved 0.222, reflecting the 21% PI. This suggests that while both models performed similarly in terms of individual prediction errors, the Engagement model was better at explaining the variance in targeted IoT vulnerabilities. Incorporating hacking community data greatly improves the model's predictive power, emphasizing the value of monitoring such activity. In contrast, the EPSS model

yielded an R^2_{MF} score of -0.157, indicating it performed worse than a mean-based prediction. This further demonstrates EPSS's limitations for IoT vulnerabilities across both metrics, as it struggles to capture the unique nuances of IoT-related threats.

7.3 Risk assessment using hacking forum

To evaluate the impact of hacking forum discussions on the risk assessment of vulnerabilities before their first exploitation, we conducted an experiment comparing model predictions in two scenarios. First, with the first-exploitation date in hand for the 78 CVEs that were both discussed in hacking forums and targeted before or on the same day as their first observed exploitation (see subsection 6.3), we applied the trained model from subsection 7.1 to this subset. To specifically assess the impact of discussions prior to first exploitation, we recalculated and aggregated forum activity data only up to their first exploitation date. Since all CVEs in this subset were targeted, the predicted probability of targeting should ideally be close to 1. The mean prediction for this subset was 0.57. To assess whether hacking forum features significantly influence targeting predictions, we compared these results to a case where the same 78 CVEs were assumed to be targeted but had no prior forum discussions. So, we excluded all hacking forum-related features while keeping all other characteristics unchanged, assuming no prior forum discussions. After applying the same data processing steps, we predicted the targeted label for this second scenario. The mean prediction dropped to 0.26. The observed reduction in predicted targeting probability of the second scenario confirms that removing forum discussions significantly reduced the model's confidence in identifying these vulnerabilities as targeted and reinforces the role of hacking forum discussions in enhancing pre-exploitation risk assessment. In addition, we computed "Thiel's U" [17] for these features to help identify their predictive power. Since "Thiel's U" closer to 1 indicates a feature almost perfectly explains the target, we found that BaseScore with a Thiel's U of 0.995, followed by count of references at 0.984, Engagement at 0.978, and number of vendors per CVE at 0.977, have the highest predictive power.

7.4 Model interpretation

7.4.1 Features significance. The fitted model includes 38 features with < 0.05, as shown in Figure 9. Each coefficient indicates the feature's influence on the likelihood of an IoT vulnerability being targeted. Among the 38 features linked to vulnerability, exploitability, and hacking community factors, all but five increase the likelihood of an IoT vulnerability being targeted.

Significance of vulnerability features. Of the 38 features, 36 were derived from vulnerability data, including 1 CVSS metric, 22 vendors, 8 CWE categories, and 5 from references features.

CVSSv3 metrics: The only CVSS metric with a significant yet negative impact on targeting is "Privileges Required: Low" (PR:L); a one-unit increase reduces the likelihood of an IoT vulnerability being targeted by 0.56 times, perhaps because attackers prioritize vulnerabilities requiring no authentication for broader impact or favor those with higher privileges for deeper system access.

Vendor: This represent the most influential feature, with 22 out of 38 features linked to targeting likelihood. All vendors, except "Huawei" and "FreeBSD," increase the likelihood of an IoT vulnerability being targeted by 1.8 to 34.6 times, with "Microchip" having the highest impact. In contrast, "Huawei" and "FreeBSD" decrease targeting likelihood by 0.02 and 0.03 times, respectively. The sharp decline associated with "Huawei" may be due to only one targeted vulnerability, CVE-2017-17215, compared to 1,400 non-targeted ones in our dataset. While this vulnerability, which allows authenticated attackers to remotely exploit Huawei HG532 routers, was the most frequently and consistently targeted IoT vulnerability [1, 2, 10, 24, 25, 38], the high standard error suggests some uncertainty in the estimate (see Table 8 in the Appendix).

CWE: All eight CWE categories increase the likelihood of an IoT vulnerability being targeted, except for CWE-352 (Cross-Site Request Forgery), which decreases it by 0.23 times. The remaining CWEs raise targeting likelihood between 1.75 and 12.29 times, with CWE-116 (Improper Encoding or Escaping of Output) having the highest impact, increasing it by 12.29 times per unit increase.

References: We identified five reference-related features assigned to CVEs by NVD, all positively correlated with a 1.43 to 2.11-fold increase in the likelihood of an IoT vulnerability being targeted, except for "Third Party Advisory", which decreases it by 0.68 times. "Issue Tracking" has the highest impact, increasing the likelihood of targeting by 2.11 times.

Significance of exploitability features. *PoC availability:* Although the availability of PoC exploits alone do not fully explain IoT targeting, the presence of PoC exploits alongside other features make IoT vulnerabilities 3.57 times more likely to be targeted compared to the ones without PoC exploits.

Significance of Engagement. This feature captures whether an IoT vulnerability was discussed in underground hacking forums, including details like the number of forums, boards, threads, posts, and members involved. For example, when community members actively engage in discussions about specific vulnerabilities, the likelihood of those vulnerabilities being targeted in the wild increases by 1.27 times. This indicates that increased forum discussions lead to a higher probability of malicious targeting. These findings align with previous results in subsection 7.2, where including forums data improved the model's variance explanation by 21% and its targeting prediction of IoT vulnerabilities by 17% over EPSS.

7.4.2 Features importance. The SHAP value distribution highlights feature influence, with the top 10 in our model ranging from 0.017 to 0.002. The highest SHAP value is for the CVSS base score, followed by CVE Age (0.013) and Engagement (0.008), while the lowest is for the vendor "Intel" (see Figure 10). Notably, the Engagement feature underscores the role of attacker-driven insights. In contrast, EPSS top features have SHAP values between 0.42–0.09, led by the count of references, Remote tag (0.34), and Code Execution (0.29), with CVSS metric (C:H) ranking lowest. We compare the feature importance derived from our model including the Engagement feature, with that from EPSS using SHAP values, as it's the only public information provided by EPSS [41]. This comparison allows us to assess the relative significance of various features in predicting exploitation, contrasting general vulnerabilities in EPSS with our model's focus on IoT vulnerabilities. Our model and EPSS exhibit considerable overlap in their top 10 influential features, with 8 of our top 10 features appearing in EPSS and 7 of EPSS's features present in our model (see Table 6). This alignment underscores shared key predictors, such as Age of CVE, exploit availability (Exploit-DB),

Table 6: Top 10 SHAP features: Our model with Engagement vs. EPSS

Model with Eng.	EPSS rank	EPSS model	Model with Eng. rank
CVSS:3.1/Scored	#28	CVE: Count of Ref.	#8
CVE: Age of CVE	#5	Tag: Remote	-
Engagement	-	Exploit: Exploit DB	#5
Ref: VDB Entry	#14	Tag: Code Exec.	-
CVSS:3.1/I:H	#24	Vendor: Microsoft	#19
Exploit: Exploit DB	#4	CVSS:3.1/PR:N	-
CVSS:3.1/A:H	#7	CVE: Age of CVE	#2
CVE: Count of Ref.	#1	CVSS:3.1/AV:N	#9
CVSS:3.1/AV:N	#7	CVE: Age of CVE	#2
Vendor: Intel	-	CVSS:3.1/C:H	#14

CVSS metrics, reference label "VDB Entry", and vendor "Intel". For example, the top predictor in our model is the CVSS base score but it ranks lower in EPSS (#28), while the count of references is EPSS top feature yet ranks lower for IoT in our model (#8). Notably, hacker forum engagement ranks as the 3rd most influential feature in our model but is absent in EPSS, emphasizing the model with Engagement's stronger focus on attacker community discussions in risk assessment.

8 Discussion

EPSS was trained on a dataset dominated by general-purpose computing vulnerabilities, as they make up the bulk of the total vulnerability population. However, IoT vulnerabilities differ in important ways that make them harder to predict using these same general features.

For instance, while our model and EPSS share similarities in their top 10 most influential features (see Table 6), only 11 of the top 30 most influential features values overlapped when considering the full list in Figure 10. In terms of the differences, we find more vendor-related features dominating our model, with 10 appearing in the top 30, compared to only two in EPSS. This underscores the role vendors play in IoT exploitability, aligning with prior work identifying Zyxel and Sonicwall among the most frequently exploited vendors [1, 2, 10]. Similarly, we found that two categories of CWE weaknesses correlate with attacker targeting for IoT vulnerabilities, yet EPSS does not rank any CWE-related features in its top 30.

These discrepancies highlight the unique and complex nature of IoT vulnerabilities. Spring [64] argued that EPSS often assigns low scores to IoT vulnerabilities that are actively exploited, such as those listed in CISA's (KEV) catalog. For example, CVE-2017- 17215 (Huawei HG532 routers) remained a frequent target yet received a near-zero EPSS score (0.05) for exploitation within 30 days as of May 2021. It took three years for the score to rise to 0.96 by April 2023—delayed recognition that does not align with real-world attacker behavior [1, 2, 10, 24, 25, 38]. Prior work emphasizes the importance of IoT-specific features, such as device types [2, 10, 14, 18, 63] and internet exposure levels [2, 14, 54], which significantly influence attacker targeting [1, 2, 10, 14, 46]. While we acknowledge that collecting such data is labor-intensive and difficult to scale for integration into automated models like EPSS, these findings suggest IoT vulnerabilities may be exploited differently from general ones.

Additionally, EPSS relies on social media discussions, such as Twitter/X, to detect exploitation, whereas research suggests that dark web discussions are more predictive [9]. To address this gap of improving the prediction for IoT vulnerabilities, we incorporated hacking community discussions into our model, enhancing the model's ability to explain variance by 21%. Not only did this inclusion improve the model's explanatory power, but it also increased its predictive accuracy by 17% over EPSS [41]. To further assess the impact of hacking forum discussions on pre-exploitation risk assessment, we found that including forum data increased the mean predicted targeting probability for these CVEs to 0.57, whereas excluding forum-related features reduced it to 0.26.

These cases accounted for 49% of discussed CVEs with known exploit dates, averaging a 482.6-day 'time-to-exploit". Notably, 9% were exploited on the same day they were first discussed. In contrast, EPSS did not rank social media feeds among its top 30 predictors, highlighting the importance of hacking forum discussions in identifying targeted IoT vulnerabilities.

This suggests that IoT vulnerabilities should likely be treated separately in prediction models, as the factors influencing their exploitation differ significantly from those of general vulnerabilities. This provides an area for future research and model development, particularly in the realm of IoT security, that might account for IoT-specific features.

9 Limitations

Several limitations exist within our methodology. Our approach solely relied on CVE identifiers for identifying IoT vulnerabilities, both those targeted in the wild and those discussed in underground forums. This may exclude vulnerabilities not assigned CVE-IDs, potentially overlooking significant threats. Additionally, the automated search methods employed using VARIoT dataset to identify IoT-related vulnerabilities might have resulted in missed or irrelevant entries. In underground forums, there is a risk of input data manipulation by malicious actors, who may discuss random IoT vulnerabilities to disrupt classification. Furthermore, the lack or ambiguity of exploit capture dates for targeted IoT vulnerabilities in some datasets, such as CISA and VirusTotal, poses challenges in understanding the timing of targeting activities. Finally, including IoT-specific features like device type [1, 2, 10, 11, 18] or install base [2, 14, 54] requires manual data collection, which is impractical given the scale of our dataset of over 23k CVEs. Instead, this study focuses on supporting vendors and users in evaluating IoT vulnerabilities by leveraging readily available online information.

10 Conclusion

We analyzed factors influencing IoT vulnerability targeting using features like CVSS severity, CWE categories, vendors, external references, vulnerability age, and PoC exploits derived from EPSS, along with new insights from hacking forums. For over 23k IoT-related CVEs, we leveraged vulnerability features, PoC availability, and discussions from more than 25k posts across 25 hacking forums. We identified 38 features with p-value < 0.05 that affected attackers' targeting choices of IoT vulnerabilities. Our model, utilizing hacking forum data, showed a 21% performance improvement in McFadden's pseudo R^2 score. Using fewer features than EPSS, our Brier score prediction accuracy increased by 17% compared to EPSS, which showed that existing state-of-the-art methods often failed to capture the complexities of IoT threats, and integrating insights into attacker behavior from online sources enhanced the prediction of IoT vulnerability targeting.

Acknowledgments

This work is partly supported by the Dutch Research Council (NWO) under the RAPID project (Grant No. CS.007), commissioned research (No.JPJ012368C08101) by National Institute of Information and Communications Technology (NICT) in Japan, and King Abdulaziz City for Science and Technology (KACST).

References

[1] Arwa Abdulkarim Al Alsadi, Kaichi Sameshima, Jakob Bleier, Katsunari Yoshioka, Martina Lindorfer, Michel van Eeten, and Carlos H Gañán. 2022. No Spring Chicken: Quantifying the Lifespan of Exploits in IoT Malware Using Static and Dynamic Analysis. (2022), 309–321.

[2] Arwa Abdulkarim Al Alsadi, Kaichi Sameshima, Katsunari Yoshioka, Michel Van Eeten, and Carlos Hernandez Gañán. 2023. Bin There, Target That: Analyzing the Target Selection of IoT Vulnerabilities in Malware Binaries. In *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses (RAID '23)*. Association for Computing Machinery, New York, NY, USA, 513–526. https://doi.org/10.1145/3607199.3607241

[3] MP Allen. 1997. *The problem of multicollinearity.* Springer US, Boston, MA, 176–180. https://doi.org/10.1007/978-0-585-25657-3_37

[4] Luca Allodi and Fabio Massacci. 2012. A Preliminary Analysis of Vulnerability Scores for Attacks in Wild: The Ekits and Sym Datasets. In *Proceedings of the 2012 ACM Workshop on Building Analysis Datasets and Gathering Experience Returns for Security* (Raleigh, North Carolina, USA) *(BADGERS '12)*. Association for Computing Machinery, New York, NY, USA, 17–24. https://doi.org/10.1145/2382416.2382427

[5] Luca Allodi and Fabio Massacci. 2014. Comparing Vulnerability Severity and Exploits Using Case-Control Studies. *ACM Trans. Inf. Syst. Secur.* 17, 1, Article 1 (aug 2014), 20 pages. https://doi.org/10.1145/2630069

[6] Luca Allodi and Fabio Massacci. 2017. Security Events and Vulnerability Data for Cybersecurity Risk Estimation. *Risk Analysis* 37 (2017). https://api.semanticscholar.org/CorpusID:22919745

[7] Luca Allodi, Fabio Massacci, and Julian Williams. 2022. The work-averse cyber-attacker model: theory and evidence from two million attack signatures. *Risk Analysis* 42, 8 (2022), 1623–1642.

[8] Luca Allodi, Woohyun Shim, and Fabio Massacci. 2013. Quantitative Assessment of Risk Reduction with Cybercrime Black Market Monitoring. *Proceedings - IEEE CS Security and Privacy Workshops, SPW 2013*, 165–172. https://doi.org/10.1109/SPW.2013.16

[9] Mohammed Almukaynizi, Eric Nunes, Krishna Dharaiya, Manoj Senguttuvan, Jana Shakarian, and Paulo Shakarian. 2019. *Patch before exploited: An approach to identify targeted software vulnerabilities.* Springer Science and Business Media Deutschland GmbH, 81–113. https://doi.org/10.1007/978-3-319-98842-9_4

[10] Omar Alrawi, Charles Lever, Kevin Valakuzhy, Ryan Court, Kevin Snow, Fabian Monrose, and Manos Antonakakis. 2021. The Circle Of Life: A Large-Scale Study of The IoT Malware Lifecycle. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, 3505–3522.

[11] Carlos A. Rivera Alvarez, Arash Shaghaghi, David D. Nguyen, and Salil S. Kanhere. 2021. Is this IoT Device Likely to be Secure? Risk Score Prediction for IoT Devices Using Gradient Boosting Machines. arXiv:2111.11874 [cs.CR]

[12] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, and Yi Zhou. 2017. Understanding the Mirai Botnet. In *26th USENIX Security Symposium (USENIX Security 17)*. USENIX Association, Vancouver, BC, 1093–1110. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

[13] AttackerKB. 2024. AttackerKB. https://attackerkb.com.

[14] Maria Bada and Ildiko Pete. 2020. An exploration of the cybercrime ecosystem around Shodan. In *2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. 1–8. https://doi.org/10.1109/IOTSMS52051.2020.9340224

[15] Randa Basheer, Bassel Alkhatib, and Zhiyong Xu. 2021. Threats from the Dark: A Review over Dark Web Investigation Research for Cyber Threat Intelligence. *J. Comput. Netw. Commun.* 2021 (jan 2021), 21 pages. https://doi.org/10.1155/2021/1302999

[16] Mats Björklund. 2019. Be careful with your principal components. *Evolution* 73, 10 (2019), 2151–2158. https://doi.org/10.1111/evo.13835 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/evo.13835

[17] Friedhelm Bliemel. 1973. Theil's Forecast Accuracy Coefficient: A Clarification. *Journal of Marketing Research* 10, 4 (1973), 444–446. https://doi.org/10.2307/3149394 Accessed: 2025-02-27.

[18] Grzegorz Blinowski, Paweł Piotrowski, and Michał Wiśniewski. 2021. Comparing Support Vector Machine and Neural Network Classifiers of CVE Vulnerabilities. 734–740. https://doi.org/10.5220/0010574807340740

[19] Mehran Bozorgi, Lawrence K. Saul, Stefan Savage, and Geoffrey M. Voelker. 2010. Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Washington, DC, USA) *(KDD '10)*. Association for Computing Machinery, New York, NY, USA, 105–114.

[20] Cambridge. 2024. Department of Computer Science and Technology: Cambridge Cybercrime Centre. https://www.cambridgecybercrime.uk/.

[21] Haipeng Chen, Rui Liu, Noseong Park, and V.S. Subrahmanian. 2019. Using Twitter to Predict When Vulnerabilities will be Exploited. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Anchorage, AK, USA) *(KDD '19)*. Association for Computing Machinery, New York, NY, USA, 3143–3152. https://doi.org/10.1145/3292500.3330742

[22] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 785–794.

[23] CISA. 2023. Known Exploited Vulnerabilities Catalog | CISA. https://www.cisa.gov/known-exploited-vulnerabilities-catalog.

[24] CUJO AI. 2023. The Zerobot Botnet: Vulnerabilities Targeted and Exploits Used in Detail. https://cujo.com/blog/the-zerobot-botnet-vulnerabilities-targeted-and-exploits-used-in-detail/.

[25] CUJO AI. 2024. IoT Botnet Report 2021: Malware and Vulnerabilities Targeted. https://cujo.com/blog/iot-botnet-report-2021-malware-and-vulnerabilities-targeted/.

[26] Dell. 2018. Meltdown/Spectre (CVE-2017-5715, CVE-2017-5753, CVE-2017-5754): Impact on Dell Products. https://www.dell.com/support/kbdoc/nl-nl/000177783/meltdown-spectre-cve-2017-5715-cve-2017-5753-cve-2017-5754-effect-op-dell-producten.

[27] Antoine d'Estalenx and Carlos Gañán. 2022. NURSE: eNd-UseR IoT malware detection tool for Smart homEs. In *Proceedings of the 11th International Conference on the Internet of Things* (St.Gallen, Switzerland) *(IoT '21)*. Association for Computing Machinery, New York, NY, USA, 134–142. https://doi.org/10.1145/3494322.3494340

[28] Michel Edkrantz. 2015. Predicting Exploit Likelihood for Cyber Vulnerabilities with Machine Learning. https://api.semanticscholar.org/CorpusID:60325485

[29] EPSS. 2024. Who is using EPSS? https://www.first.org/epss/who_is_using/.

[30] ExploitDB. 2024. Exploit Database - Exploits for Penetration Testers, Researchers, and Ethical Hackers. https://www.exploit-db.com/.

[31] FIRST Inc. 2023. Common Vulnerability Scoring System v3.0: Specification Document. https://www.first.org/cvss/specification-document

[32] Forum of Incident Response and Security Teams, Inc. 2024. Exploit Prediction Scoring System (EPSS). https://www.first.org/epss/api.

[33] GCA. 2022. *GCA Internet Integrity Papers: IoT Policy and Attack Report II.* Technical Report. Global Cyber Alliance. https://www.globalcyberalliance.org/reports_publications/iot-policy-and-attack-report-ii/.

[34] Github. 2022. . https://github.com/.

[35] Google. 2024. Google Project Zero. https://googleprojectzero.blogspot.com/p/0day.html.

[36] Allen D Householder, Jeff Chrabaszcz, Trent Novelly, David Warren, and Jonathan M Spring. 2020. Historical analysis of exploit availability timelines. In *13th USENIX Workshop on Cyber Security Experimentation and Test (CSET 20)*.

[37] Emanuele Iannone, Giulia Sellitto, Emanuele Iaccarino, Filomena Ferrucci, Andrea De Lucia, and Fabio Palomba. 2024. Early and Realistic Exploitability Prediction of Just-Disclosed Software Vulnerabilities: How Reliable Can It Be? *ACM Trans. Softw. Eng. Methodol.* (mar 2024). https://doi.org/10.1145/3654443 Just Accepted.

[38] J. Salvio, R. Tay. 2022. Fresh TOTOLINK Vulnerabilities Picked Up by Beastmode Mirai Campaign. https://www.fortinet.com/blog/threat-research/totolink-vulnerabilities-beastmode-mirai-campaign.

[39] Jay Jacobs, Sasha Romanosky, Idris Adjerid, and Wade Baker. 2020. Improving vulnerability remediation through better exploit prediction. *Journal of Cybersecurity* 6 (01 2020). https://doi.org/10.1093/cybsec/tyaa015

[40] Jay Jacobs, Sasha Romanosky, Benjamin Edwards, Idris Adjerid, and Michael Roytman. 2021. Exploit Prediction Scoring System (EPSS). *Digital Threats* 2, 3, Article 20 (jul 2021), 17 pages.

[41] J. Jacobs, S. Romanosky, O. Suciu, B. Edwards, and A. Sarabi. 2023. Enhancing Vulnerability Prioritization: Data-Driven Exploit Predictions with Community-Driven Insights. In *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE Computer Society, Los Alamitos, CA, USA, 194–206. https://doi.org/10.1109/EuroSPW59978.2023.00027

[42] Marek Janiszewski, Anna Felkner, Piotr Lewandowski, Marcin Rytel, and Hubert Romanowski. 2021. Automatic Actionable Information Processing and Trust Management towards Safer Internet of Things. *Sensors* 21, 13 (2021). https://doi.org/10.3390/s21134359

[43] Marek Janiszewski, Marcin Rytel, Piotr Lewandowski, and Hubert Romanowski. 2022. VARIoT - Vulnerability and Attack Repository for the Internet of Things. In *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. IEEE, 752–755. https://doi.org/10.1109/CCGrid54584.2022.00085

[44] Seiya Kato, Rui Tanabe, Katsunari Yoshioka, and Tsutomu Matsumoto. 2021. Adaptive Observation of Emerging Cyber Attacks targeting Various IoT Devices.

In *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. 143–151.

[45] Metasploit. 2024. Metasploit | Penetration Testing Software. https://www.metasploit.com/.

[46] Trend Micro. 2019. Uncovering IoT Threats in the Cybercrime Underground. https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/the-internet-of-things-in-the-cybercrime-underground.

[47] Felipe Moreno-Vera. 2023. Inferring Discussion Topics about Exploitation of Vulnerabilities from Underground Hacking Forums. In *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*. 816–821. https://doi.org/10.1109/ICTC58733.2023.10393244

[48] Felipe Moreno-Vera, Mateus Nogueira, Cainã Figueiredo, Daniel Sadoc Menasché, Miguel Bicudo, Ashton Woiwood, Enrico Lovat, Anton Kocheturov, and Leandro Pfleger de Aguiar. 2023. Cream Skimming the Underground: Identifying Relevant Information Points from Online Forums. arXiv:2308.02581 [cs.CR]

[49] Antonio Nappa, Richard Johnson, Leyla Bilge, Juan Caballero, and Tudor Dumitras. 2015. The Attack of the Clones: A Study of the Impact of Shared Code on Vulnerability Patching. In *2015 IEEE Symposium on Security and Privacy*. 692–708. https://doi.org/10.1109/SP.2015.48

[50] National Institute of Standards and Technology (NIST). 2021. National Vulnerability Database. https://nvd.nist.gov/

[51] Kartik Nayak, Daniel Marino, Petros Efstathopoulos, and Tudor Dumitraş. 2014. Some vulnerabilities are different than others: Studying vulnerabilities and attack surfaces in the wild. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 8688 LNCS. Springer Verlag, 426–446. https://doi.org/10.1007/978-3-319-11379-1{_}21

[52] NVD. 2022. CVE - Common Vulnerabilities and Exposures (CVE). https://cve.mitre.org/index.html.

[53] NVD. 2024. Common Weakness Enumeration (CWE). https://cwe.mitre.org/.

[54] Ofri Ouzan. 2023. Advanced Shodan Use for Tracking Down Vulnerable Components. https://medium.com/@ofriouzan/advanced-shodan-use-for-tracking-down-vulnerable-components-7b6927a87c45.

[55] Yin Minn Pa Pa, Shogo Suzuki, Katsunari Yoshioka, Tsutomu Matsumoto, Takahiro Kasama, and Christian Rossow. 2015. IoTPOT: Analysing the Rise of IoT Compromises. In *9th USENIX Workshop on Offensive Technologies (WOOT 15)*. USENIX Association, Washington, D.C. https://www.usenix.org/conference/woot15/workshop-program/presentation/pa

[56] Packtsecurity. 2022. Spectre or Meltdown vulnerabilities on IoT. https://security.packt.com/spectre-or-meltdown-vulnerabilities-on-iot/.

[57] Sergio Pastrana, Daniel R. Thomas, Alice Hutchings, and Richard Clayton. 2018. CrimeBB: Enabling Cybercrime Research on Underground Forums at Scale. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1845–1854. https://doi.org/10.1145/3178876.3186178

[58] S. Rivera Pérez, M. van Eeten, and C. H. Gañán. 2024. Patchy Performance? Uncovering the Vulnerability Management Practices of IoT-Centric Vendors. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society, Los Alamitos, CA, USA, 153–153. https://doi.org/10.1109/SP54263.2024.00154

[59] Carl Sabottke, Octavian Suciu, and Tudor Dumitraş. 2015. Vulnerability disclosure in the age of social media: exploiting twitter for predicting real-world exploits. In *Proceedings of the 24th USENIX Conference on Security Symposium* (Washington, D.C.) *(SEC'15)*. USENIX Association, USA, 1041–1056.

[60] Paulo Shakarian. 2018. Dark-Web Cyber Threat Intelligence: From Data to Intelligence to Prediction. *Information* 9, 12 (2018). https://doi.org/10.3390/info9120305

[61] Shodan. 2022. Shodan Search Engine. https://www.shodan.io/dashboard.

[62] Yuba R. Siwakoti and Danda B. Rawat. 2023. Detect-IoT: A Comparative Analysis of Machine Learning Algorithms for Detecting Compromised IoT Devices. In *Proceedings of the Twenty-Fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing* (Washington, DC, USA) *(MobiHoc '23)*. Association for Computing Machinery, New York, NY, USA, 370–375. https://doi.org/10.1145/3565287.3616529

[63] Jinke Song, Shangfeng Wan, Min Huang, Jiqiang Liu, Limin Sun, and Qiang Li. 2023. Toward Automatically Connecting IoT Devices with Vulnerabilities in the Wild. *ACM Trans. Sen. Netw.* 20, 1, Article 6 (oct 2023), 26 pages. https://doi.org/10.1145/3608951

[64] Jonathan Spring. 2022. Probably Don't Rely on EPSS Yet. Carnegie Mellon University, Software Engineering Institute's Insights (blog). https://insights.sei.cmu.edu/blog/probably-dont-rely-on-epss-yet/

[65] Jonathan Spring, Eric Hatleback, Allen Householder, Art Manion, and Deana Shick. 2021. Time to Change the CVSS? *IEEE Security & Privacy* 19, 2 (2021), 74–78. https://doi.org/10.1109/MSEC.2020.3044475

[66] Octavian Suciu, Connor Nelson, Zhuoer Lyu, Tiffany Bao, and Tudor Dumitras. 2022. Expected Exploitability: Predicting the Development of Functional Vulnerability Exploits. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 377–394. https://www.usenix.org/conference/usenixsecurity22/presentation/suciu

[67] Nazgol Tavabi, Palash Goyal, Mohammed Almukaynizi, Paulo Shakarian, and Kristina Lerman. 2018. DarkEmbed: Exploit Prediction With Neural Language Models. *Proceedings of the AAAI Conference on Artificial Intelligence* 32 (04 2018). https://doi.org/10.1609/aaai.v32i1.11428

[68] VARIoT. 2024. VARIoT – Vulnerability and Attack Repository for IoT. https://www.variot.eu/.

[69] VirusTotal. 2024. VirusTotal. https://www.virustotal.com.

[70] VirusTotal. 2024. VirusTotal – Learning resources - Vulnerability management. https://www.virustotal.com/getstarted/vulnerability-management.

[71] VulnCheck. 2024. VulnCheck KEV. https://vulncheck.com/kev.

[72] vulners. 2024. All-in-one vulnerability intelligence. https://vulners.com/.

[73] Suzanne Widup, Marc Spitler, David Hylender, and Gabriel Bassett. 2018. 2018 Verizon Data Breach Investigations Report.

[74] J. Wunder, A. Kurtz, C. Eichenmüller, F. Gassmann, and Z. Benenson. 2024. Shedding Light on CVSS Scoring Inconsistencies: A User-Centric Study on Evaluating Widespread Security Vulnerabilities. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society, Los Alamitos, CA, USA, 57–57. https://doi.org/10.1109/SP54263.2024.00058

[75] Wei Yang, Jiakun Jiang, Erin Schnellinger, Stephen Kimmel, and Wensheng Guo. 2022. Modified Brier score for evaluating prediction accuracy for binary outcomes. *Statistical Methods in Medical Research* 31 (08 2022), 096228022211223. https://doi.org/10.1177/09622802221122391

A CrimeBB dataset

A.1 Summary of CrimeBB dataset

Table 7: Summary of CrimeBB forums ordered by the number of posts.

Forums	# Posts	Oldest post
hack-forums	42,474,326	2007-01-27
zismo	11,127,167	2010-05-26
blackhatworld	10,320,120	2005-10-31
multiplayer-game-hacking	10,217,579	2005-12-26
nulled	6,675,498	2013-04-02
lolzteam	6,196,005	2013-03-10
ogusers	3,608,306	1900-01-01
cracked	2,977,801	2018-04-03
mmo4me	2,899,930	2010-04-01
antichat	2,630,906	2002-05-29
v3rmillion	2,459,519	2016-02-02
unknowncheats	2,403,995	2002-11-02
raidforums	1,231,126	2015-03-20
elhacker	980,523	2002-08-21
probiv	822,671	2014-11-05
breached	737,922	2022-03-16
forum-team	431,695	2017-10-31
indetectables	328,024	2006-02-20
xss-forum	310,796	2004-11-13
dread	294,596	2018-02-15
runion	240,632	2012-01-11
offensive-community	161,492	2012-06-30
underc0de	92,247	2010-02-10
the-hub	88,753	2014-01-09
ifud	72,851	2012-05-10
piratebay-forum	60,678	2013-10-23
torum	28,485	2017-05-25
safe-sky-hacks	27,018	2013-03-28
kernelmode	26,815	2010-03-11
freehacks	26,471	2013-07-27
deutschland-im-deep-web	20,185	2018-11-22
greysec	11,925	2015-06-10
garage-for-hackers	8,710	2010-07-06
stresser-forums	7,069	2017-04-09
envoy-forum	2,163	2019-07-06
Hackers Armies	-	-
Total	11,0003,999	

B Model

B.1 Features with factors influencing the targeting of certain IoT vulnerability using GLM.

Table 8: Regression results of the model including the Engagement feature. Only the 38 features with p-value < 0.05 are shown.

Feature	Coefficient	Standard Error
PoC	1.272^{***}	(0.165)
PR:L	-0.595^{*}	(0.240)
Vendor:INTEL	1.159^{***}	(0.261)
Vendor:FEDORAPROJECT	0.891^{**}	(0.338)
Vendor:CITRIX	2.013^{***}	(0.480)
Vendor:MITEL	2.051^{**}	(0.733)
Vendor:DLINK	0.785^{***}	(0.223)
Vendor:DRAYTEK	3.399^{***}	(0.629)
Vendor:WESTERNDIGITAL	2.111^{*}	(0.920)
Vendor:WAVLINK	1.849^{*}	(0.787)
Vendor:MINIO	2.591^{**}	(0.853)
Vendor:APACHE	2.304^{***}	(0.356)
Vendor:HUAWEI	-3.882^{**}	(1.217)
Vendor:ZYXEL	1.440^{**}	(0.463)
Vendor:MICROCHIP	3.544^{*}	(1.495)
Vendor:FORTINET	1.183^{***}	(0.273)
Vendor:FREEBSD	-3.447^{***}	(0.909)
Vendor:EMBEDTHIS	1.943^{*}	(0.803)
Vendor:OPENSSL	2.992^{***}	(0.555)
Vendor:GRANDSTREAM	1.566^{*}	(0.798)
Vendor:LINKSYS	1.546^{*}	(0.736)
Vendor:SONICWALL	1.446^{**}	(0.492)
Vendor:DAHUASECURITY	2.230^{**}	(0.722)
Vendor:NUUO	1.540^{*}	(0.653)
CWE-22	0.590^{*}	(0.266)
CWE-352	-1.488^{*}	(0.739)
CWE-843	1.658^{**}	(0.535)
CWE-287	0.562^{*}	(0.250)
CWE-78	1.152^{***}	(0.182)
CWE-94	1.107^{*}	(0.493)
CWE-116	2.509^{*}	(1.086)
CWE-294	1.733^{*}	(0.747)
Ref:Third Party Advisory	-0.389^{*}	(0.175)
Ref:Exploit	0.358^{*}	(0.171)
Ref:Release Notes	0.562^{*}	(0.261)
Ref:VDB Entry	0.416^{*}	(0.196)
Ref:Issue Tracking	0.746^{**}	(0.266)
Engagement	0.237^{***}	(0.017)
Observations	15,600	
Akaike Inf. Crit.	3,810.212	
Note:	*p<0.05; **p<0.01; ***p<0.001	

B.2 Model feature coefficients

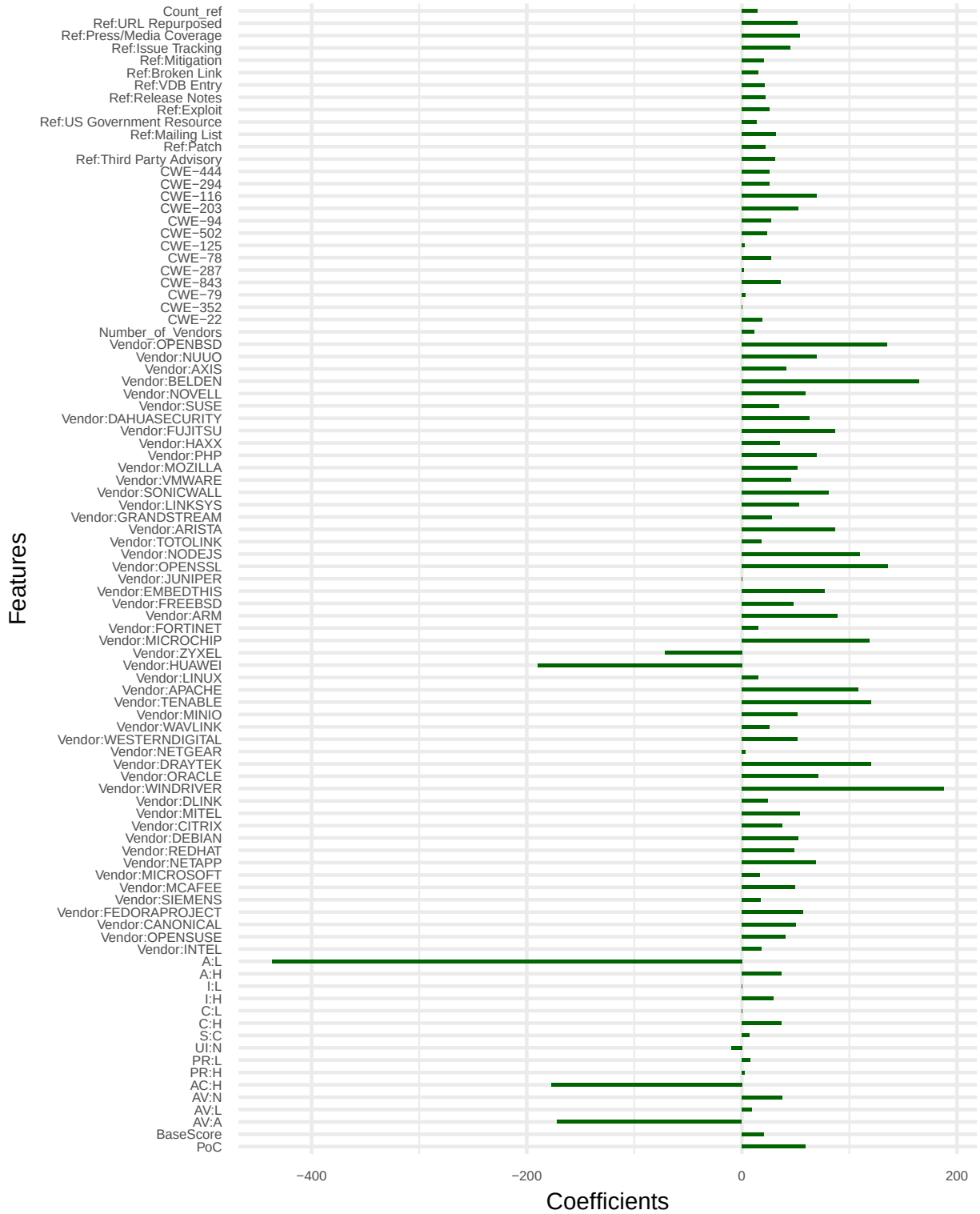


Figure 8: The 93 features used selected by elastic net regularization.

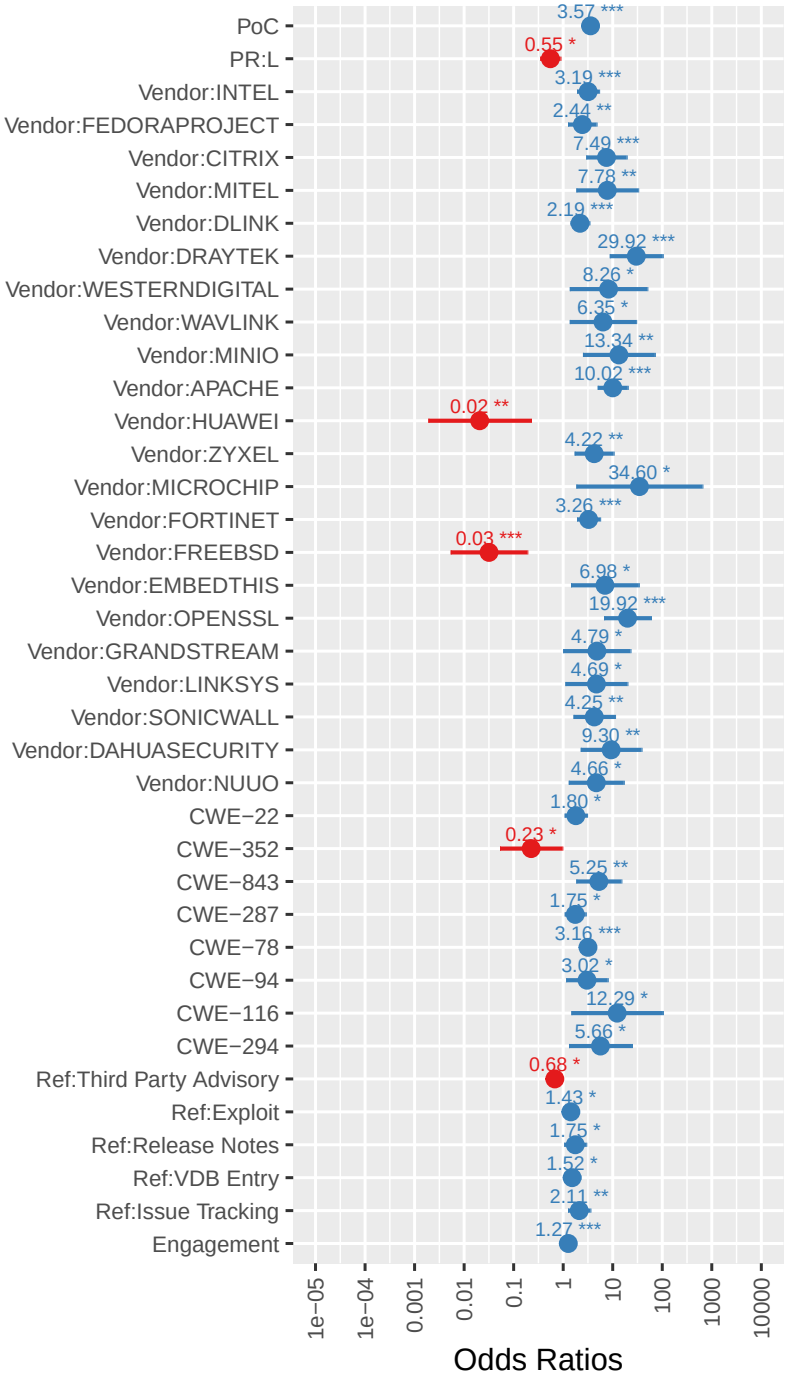


Figure 9: Odds ratio of the regression model with hacking community PC in regularization. Only the features with p-value < 0.05 are shown.

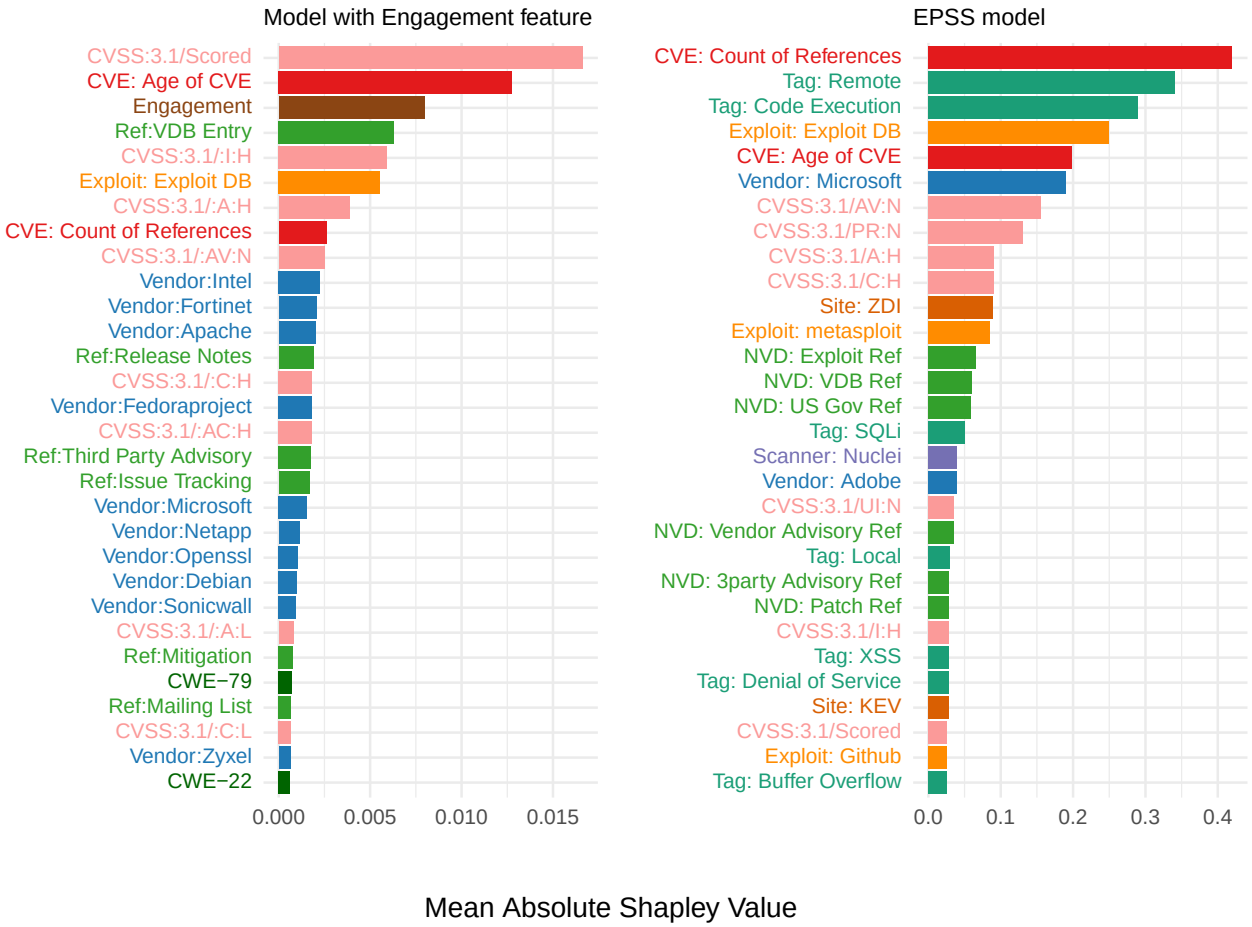


Figure 10: Mean absolute Shapley value for the top 30 most influential feature of our model with Engagement vs. EPSS.