Zero-Day Attack and Cybersecurity on Twitter

Subjects: Computer Science, Cybernetics

Contributor: Ahmet Ercan Topcu, Yehia Ibrahim Alzoubi, Ersin Elbasi, Emre Camalan

In the information era, knowledge can pose risks in the online realm. It is imperative to proactively recognize potential threats, as unforeseen dangers cannot be eliminated entirely. Often, malware exploits and other emerging hazards are only identified after they have occurred. These types of risks are referred to as zero-day attacks since no pre-existing anti-malware measures are available to mitigate them. Consequently, significant damages occur when vulnerabilities in systems are exploited. The effectiveness of security systems, such as IPS and IDS, relies heavily on the prompt and efficient response to emerging threats.

Keywords: zero-day attack ; Twitter ; TensorFlow ; machine learning ; word classification

1. Introduction

In the modern era, the Internet has emerged as the most efficient, accessible, and cost-effective medium for international communication. Social networking platforms, such as Twitter, YouTube, and Facebook, have become essential tools for billions of individuals to connect, exchange information, and enjoy various content. Moreover, governmental organizations and businesses have recognized the potential of these online networks to reach a broader audience and carry out effective advertising campaigns [1]. As of April 2023, approximately 5.18 billion people, accounting for 64.6% of the global population, were using the Internet. Out of these Internet users, around 4.8 billion individuals, making up 59.9% of the world's population, actively engaged with social networking websites [2]. These statistics highlight the widespread adoption and popularity of the Internet and social media platforms globally.

Indeed, the widespread use of the Internet and social media platforms brings significant privacy and security concerns for individuals, data, systems, businesses, and governments. The sheer volume of online activity makes it challenging to monitor every single action. Specific individuals' behaviors can have harmful consequences for themselves or others. It is crucial to be aware of these risks and take proactive measures to protect personal information, maintain secure systems, and promote responsible online behavior [3]. On the other hand, businesses and Internet service providers frequently collect sensitive information for advertising purposes and other uses. This behavior has led to growing concerns among Internet users regarding the security and confidentiality of their personal data. Research conducted on security and privacy in cyberspace indicates that most individuals believe it is primarily the responsibility of the government to ensure security on the Internet [4]. These findings highlight the need for stronger regulations and measures to protect user privacy and enhance overall cybersecurity in online environments [5]. While it is important for governments to establish regulations and frameworks to ensure online security, it is also the responsibility of individuals to take proactive measures to protect themselves online.

One of the most severe cyber threats is the zero-day attack, which remains unrecognized by both the general Internet community and experts [6]. Zero-day attacks, by their very nature, lack pre-existing fixes or protections because they exploit unknown vulnerabilities. These attacks can be leveraged to breach security measures, install malicious software, extract sensitive information, or disrupt system performance [7]. Due to their novelty and unpredictability, zero-day attacks pose significant challenges for defenders, as no known defense mechanisms or countermeasures are available, making them difficult to detect and defend against. This emphasizes the importance of ongoing monitoring, strong cybersecurity measures, and collaborative efforts among experts, researchers, and organizations to identify and address these emerging threats [8] effectively.

2. Zero-Day Attack

Cybersecurity threats encompass a wide range of challenges, including both technical and non-technical attacks [9]. Technical attacks involve exploiting vulnerabilities in software, networks, or systems, whereas non-technical attacks, like Social Engineering, manipulate human psychology to deceive individuals into revealing sensitive information or performing certain actions [10]. To minimize cyber threats, a comprehensive approach is essential. Conducting workshops,

seminars, and simulations can raise awareness among employees, teaching them to recognize and avoid social engineering tactics. Game-based education fosters a proactive cybersecurity mindset [11]. On the technical front, protecting against zero-day attacks requires robust intrusion detection and prevention systems, as well as timely software patching and updates [7]. A multilayered defense strategy, incorporating encryption, access controls, and regular security assessments, can bolster overall cyber resilience and safeguard against evolving threats.

A zero-day attack is a type of cyberattack that exploits previously unknown vulnerabilities in software or systems. It uses security flaws that software makers have not found or fixed, leaving users open to possible attacks [4]. The term “zero-day” indicates that attackers discover and exploit the vulnerability before the software vendor becomes aware of it, leaving zero days to prepare a defense. Zero-day attacks can have severe consequences since they might result in unauthorized access, breaches of data, the execution of malicious software, or the interruption of system operation. Zero-day attacks provide substantial issues requiring proactive security measures, ongoing monitoring, and quick software upgrades to fix new vulnerabilities [12]. Systems remain susceptible until the patch is deployed since it might take a while for software providers to create a patch to cure the vulnerability [6].

A number of tools, such as OSINT, and other methods for zero-day threat detection are available to examine abnormal activity. OSINT, a crucial component of efficient security intelligence processes, is a collection of signs pointing to heightened risk and highlighting certain threats [8]. These data are essential for assisting security officials in identifying possible risks and deciding what course of action to take next. The data obtained by OSINT can come from various publicly available sources, which is why it is described as open-source intelligence [13].

3. Cybersecurity on Twitter

Social media users often fall victim to security vulnerabilities by unknowingly exposing private data, images, or messages through their accounts and profiles. This is a significant contributing factor to the security flaws observed in these platforms. Both organizations and individuals increasingly rely on social media as a means of communication and find it convenient to target specific audiences [14]. Cybersecurity encompasses a set of principles, rules, techniques, recommendations, reliability, and technologies aimed at protecting the resources of companies and customers and the online environment [15]. These resources include personnel, infrastructure systems, facilities, network connections, and all the data transmitted and stored within the online environment. The primary objective of cybersecurity is to establish and maintain effective safeguards against relevant security threats in the online realm [16].

The importance and global attention towards cybersecurity have significantly increased. Notably, an action document has been published, which features the official positions of over 50 countries on cyberspace, cybercrime, or cybersecurity. This demonstrates the growing recognition and efforts to address cybersecurity challenges internationally [17]. Multiple systems should be employed simultaneously to provide cybersecurity since the chances of an attack are growing as Internet usage grows. In the current climate, where attacks are growing exponentially more complicated, corporate cybersecurity is a notion that cannot be disregarded. There are several approaches to security intelligence, as described above. These can be terrifying, but tracking individuals is vital for the safety of governmental systems or communities when it comes to terrorism, threats, or cyberwarfare. Because of this, scientists and engineers are working to develop fresh approaches to intelligence technology [5].

Twitter is widely recognized as one of the most popular text-based social media platforms on the Internet. Its unique text-oriented format makes it particularly suitable for this research compared to other social platforms. Users can easily share personal information, news, meeting details, or relevant texts on Twitter. Therefore, it becomes a valuable platform to monitor and gather information about new zero-day threats. However, it is important to note that Twitter imposes certain restrictions on data collection. While users can access and retrieve information from Twitter using automated tools, like robots or crawlers, there are limitations in place to prevent the unrestricted collection of all users’ information. These restrictions aim to safeguard user privacy and ensure responsible data usage [5]. By utilizing Twitter as a social media platform, this research can leverage its text-focused nature and the availability of publicly shared information to analyze and identify potential zero-day threats. The research acknowledges and adheres to Twitter’s policies and guidelines concerning data collection and respects users’ privacy while conducting the research.

Cyber-Twitter is a program designed for real-time social media research, specifically Twitter, to identify and assess risks and potentially hazardous communications. It functions as a system to discover and analyze cybersecurity intelligence on Twitter, acting as an OSINT source [8]. When the program is active, it collects data from Twitter and analyzes tweets to identify those that pertain to potential attacks or threats. This program allows for identifying, tagging, and extracting various real-world conceptual entities related to cybersecurity vulnerabilities. These entities may include the methods or

means of an attack, the consequences of an attack, and the software, hardware, or vendors affected by the vulnerabilities. The program employs a security vulnerability concept extractor to perform these tasks [18].

4. Social Media Zero-Day Attack Detection

Various research papers in the literature were identified, all of which centered on investigating zero-day attacks. **Table 1** provides a summary of these studies, including their specific focus, and highlights the distinctions between this current study and the previously conducted ones. Altalhi and Gutub [19] examined Twitter data to identify and predict security threats. They compared various previously published papers that utilized Twitter streaming data to gather information on ongoing and potential cyberattacks. The investigation considered aspects such as detection scope, performance measurements, feature extraction methods, information summarization levels, algorithm complexity, and scalability over time. Several recommendations were proposed to improve the accuracy of forecasts. The results indicated that the SYNAPSE strategy achieved the highest summation value of 490 and an average score of 82, making it the preferred overall approach. On the other hand, the DataFreq scheme performed well, but could not match SYNAPSE in terms of average and total scores, positioning it as the second contender for enhancement [19].

Table 1. Summary of related work on zero-day attack detection.

Study	Focus	Method	Description
[19]	Predictions of cyberattacks	Survey	Comparing different proposed work against detection scope, performance measurements, feature extraction methods, information summarization levels, algorithm complexity, and scalability over time
[1]	Zero-day prediction	Autoencoder and deep anomaly detection	Test three datasets from the real world totaling 222,541 URLs
[20]	Zero-day detection	Deep learning technique	Developing an Intrusion IDS model with a high recall rate and minimal false negatives
[21]	Zero-day detection	tDCGAN	Generating synthetic malware and distinguishing it from real malware
[22]	Zero-day detection	Semi-supervised machine learning	Deploying Benford's law that locates abnormal behavior based on the distribution of leading digits in numerical data
[23]	Zero-day detection	Neural Network classifier	Generating synthetic zero-day data and applying NN classifier to predict the zero-day attack
[24]	Zero-day detection	Zero-shot learning approach	Evaluating the effectiveness of machine learning-based IDSs in recognizing zero-day attack
[25]	Zero-day detection	Deep learning-based IDS	Using deep novelty-based classifiers and conventional clustering based on specialized layers of deep structures
[26]	Analogous zero-day detection	PlausMal-GAN	A malware training framework based on the generated analogous malware data using generative adversarial networks

Study	Focus	Method	Description
[27]	Classify various ransomware variants	Multi-tier streaming analytics model	Numerically grouping ransomware variants into ancestor groups and statistically combining those from multiple-descendant families
This research	Zero-day detection	Tensorflow technique	Collecting and analyzing real data from the Twitter platform to detect potential zero-day attacks

Bu and Cho [1] presented a method for identifying zero-day attacks utilizing a convolutional autoencoder and deep character-level anomaly detection. They conducted rigorous tests using three datasets from the real world totaling 222,541 URLs, and their strategy outperformed other recent deep learning approaches. Evaluation of the receiver operating characteristic curve, tenfold cross-validation, and contrasts with deep learning techniques based on categorization was used to demonstrate the superiority of the suggested method. The findings showed that the suggested strategy improved sensitivity by 3.98% compared to classification-based deep learning approaches. The enhancement was credited to using an operation designed for the unique properties of URLs and installing a neural network structure optimized for URL modeling. The study highlighted the potential for boosting cybersecurity measures by demonstrating the efficacy of the deep character-level anomaly detection technique in detecting zero-day threats [1].

Hindy et al. [20] proposed using an autoencoder technique for detecting zero-day vulnerabilities. The objective was to develop an Intrusion Detection System (IDS) model with a high recall rate and minimal false negatives. To demonstrate the effectiveness of the model, its results were compared with those of a one-class Support Vector Machine (SVM). The study focused on assessing the one-class SVM's performance when zero-day attacks deviated from expected behavior. The autoencoder's encoding–decoding features proved to be highly beneficial for the proposed IDS model. The results of the study revealed that autoencoders are capable of effectively identifying sophisticated zero-day attacks. The findings showed an accuracy range of 89–99% in zero-day detection, underscoring the efficacy of the approach [20].

Kim et al. [21] proposed a technique called transferred deep-convolutional generative adversarial network (tDCGAN), aiming to generate synthetic malware and effectively distinguish it from real malware. The method leverages actual data and data modified by tDCGAN through a deep autoencoder, which extracts essential features and enhances GAN training stability. The deep autoencoder learns malware characteristics, generates generic data, and passes this knowledge to facilitate reliable GAN training. Through transfer learning, the trained discriminator imparts its ability to recognize malware traits to the detection system. As a result, tDCGAN achieved an average classification accuracy of 95.74%, contributing to increased learning stability in malware detection [21].

Mbona and Eloff [22] presented a method for employing semi-supervised machine learning to detect zero-day attacks. The use of Benford's law, a statistical theory that locates abnormal behavior based on the distribution of leading digits in numerical data, is part of their approach. The study reveals that this method may successfully recognize key network characteristics that point to aberrant behavior, helping to identify zero-day attacks. Their findings emphasize using semi-supervised machine learning techniques for making the best choice of pertinent attributes. According to experimental findings, one-class SVMs performed best in identifying zero-day attacks, with a 74% Matthews correlation coefficient and an 85% F1 score. These results highlight the potential of semi-supervised machine learning approaches for identifying zero-day attacks, mainly when used in conjunction with suitable feature selection techniques [22].

Peppes and Alexakis [23] presented an approach involving the generation of authentic zero-day-type data in tabular format and evaluating a neural network trained with and without synthetic data to detect zero-day attacks. By employing Generative Adversarial Networks (GANs), they successfully created a larger dataset of synthetic information on zero-day exploits. This dataset was used to train a Neural Network classifier specifically designed for identifying zero-day attacks, and its performance was evaluated. The results showed that after approximately 5000 iterations, the synthetic zero-day attack data in tabular format reached a state of equilibrium, producing data that closely resembled the original data samples.

To evaluate the effectiveness of machine learning-based IDSs in recognizing zero-day attack scenarios, Sarhan et al. [24] developed a unique zero-shot learning approach. The learning models translate data characteristics to semantic attributes that distinguish between known attacks and benign activity during the attribute learning step. To identify zero-day attacks as malicious, the models create connections between known and zero-day attacks during the deductive phase. The

efficiency of the learning model in identifying unknown attacks is measured by a new assessment metric called Zero-day Detection Rate. Two machine learning models, as well as two current IDS datasets, were used to assess the proposed system. The findings show that several of the study's identified zero-day attack groups provide problems for ML-based IDSs since they are difficult to identify as malicious. Further investigation revealed that these assaults with low-zero-day detection percentages have distinctive characteristic distributions and a wider Wasserstein distance than assaults in other assault classes. These results show the need to consider certain feature distributions to overcome such obstacles and illustrate the shortcomings of ML-based IDSs in successfully identifying specific zero-day attack situations [24].

Soltani et al. [25] proposed a thorough IDS system that uses deep learning methods to counter new assaults efficiently. This system stands out as the first of its type to use both deep novelty-based classifiers and conventional clustering based on specialized layers of deep structures. The study presented DOC++, an improved version of DOC that acts as a deep novelty-based classifier. In the preprocessing stage, the Deep Intrusion Detection framework was used to improve the capacity of deep learning algorithms to recognize content-based assaults. Four distinct algorithms—DOC, DOC++, OpenMax, and AutoSVM—were contrasted as the framework's novel detectors. According to the results, the open set identification module was most successfully implemented by DOC++. Furthermore, the clustering and post-training stages validated the model's applicability for supervised labeling and updating procedures, which showed good levels of thoroughness and uniformity. The results highlight the proposed framework's robustness and effectiveness in handling novel attack situations, confirming its value as a deep learning-based IDS strategy [25].

Won et al. [26] developed a malware training methodology called PlausMal-GAN, which utilizes generative adversarial networks to create similar malware data. They used a combination of authentic and artificially generated malware images to train the discriminator, which acts as a detector, to recognize various virus characteristics. The proposed approach outperformed other methods, especially for equivalent zero-day malware images, which are considered analogous zero-day malware data. Moreover, the architecture offers significant advantages for antivirus systems, as it does not require time-consuming malware signature evaluation. This highlights the potential of PlausMal-GAN as an effective and efficient tool for enhancing malware detection capabilities [26].

Zuhair et al. [27] suggested a multi-tiered streaming analytics technique known as a hybrid machine learner, which utilizes 24 static and dynamic features to classify various ransomware variants belonging to 14 different families. The suggested methodology involves numerically grouping ransomware variants into ancestor groups and statistically combining those from multiple-descendant families. To evaluate the effectiveness of the approach, the methodology was applied to categorize ransomware variants among a dataset consisting of 40,000 samples, including malicious software, goodware, and different variations of ransomware, in both semi-realistic and realistic scenarios. In a realistic comparison test, the ransomware streaming analytics model demonstrated an average classification accuracy of 97%, an error rate of 2.4%, and a miss rate of 0.34%. These results indicate that the proposed model outperforms competing anti-ransomware technologies in terms of performance and accuracy [27].

References

1. Bu, S.-J.; Cho, S.-B. Deep character-level anomaly detection based on a convolutional autoencoder for zero-day phishing URL detection. Electronics 2021, 10, 1492.
2. Statista. Number of Internet and Social Media Users Worldwide as of April 2023. Available online: https://www.statista.com/statistics/617136/digital-population-worldwide/ (accessed on 26 June 2023).
3. Marinho, R.; Holanda, R. Automated emerging cyber threat identification and profiling based on natural language processing. IEEE Access 2023, 11, 58915–58936.
4. Cheng, X.; Zhang, J.; Tu, Y.; Chen, B. Cyber situation perception for Internet of things systems based on zero-day attack activities recognition within advanced persistent threat. Concurr. Comput. Pract. Exp. 2022, 34, e6001.
5. Pattnaik, N.; Li, S.; Nurse, J.R. Perspectives of non-expert users on cyber security and privacy: An analysis of online discussions on twitter. Comput. Secur. 2023, 125, 103008.
6. Zahoora, U.; Rajarajan, M.; Pan, Z.; Khan, A. Zero-day ransomware attack detection using deep contractive autoencoder and voting based ensemble classifier. Appl. Intell. 2022, 52, 13941–13960.
7. Ahmad, R.; Alsmadi, I.; Alhamdani, W.; Tawalbeh, L.A. Zero-day attack detection: A systematic literature review. Artif. Intell. Rev. 2023, 5, 1–79.
8. Yadav, A.; Kumar, A.; Singh, V. Open-source intelligence: A comprehensive review of the current state, applications and future perspectives in cyber security. Artif. Intell. Rev. 2023, 15, 1–32.

9. Malatji, M.; Marnewick, A.; von Solms, S. Validation of a socio-technical management process for optimising cybersecurity practices. Comput. Secur. 2020, 95, 101846.
10. Fatima, R.; Yasin, A.; Liu, L.; Wang, J. How persuasive is a phishing email? A phishing game for phishing awareness. J. Comput. Secur. 2019, 27, 581–612.
11. Fatima, R.; Yasin, A.; Liu, L.; Jianmin, W. Strategies for counteracting social engineering attacks. Comput. Fraud Secur. 2022, 2022, S1361–S3723.
12. Ali, S.; Rehman, S.U.; Imran, A.; Adeem, G.; Iqbal, Z.; Kim, K.-I. Comparative evaluation of AI-based techniques for zero-day attacks detection. Electronics 2022, 11, 3934.
13. Fjelland, R. Why general artificial intelligence will not be realized. Humanit. Soc. Sci. Commun. 2020, 7, 10.
14. Mishra, A.; Alzoubi, Y.I.; Anwar, M.J.; Gill, A.Q. Attributes impacting cybersecurity policy development: An evidence from seven nations. Comput. Secur. 2022, 120, 102820.
15. Fourati, M.; Jedidi, A.; Gargouri, F. A deep learning-based classification for topic detection of audiovisual documents. Appl. Intell. 2022, 53, 8776–8798.
16. Mishra, A.; Alzoubi, Y.I.; Gill, A.Q.; Anwar, M.J. Cybersecurity enterprises policies: A comparative study. Sensors 2022, 22, 538.
17. Mishra, A.; Jabar, T.S.; Alzoubi, Y.I.; NathMishra, K. Enhancing privacy-preserving mechanisms in cloud storage: A novel conceptual framework. Concurr. Comput. Pract. Exp. 2023, e7831.
18. Mittal, S.; Das, P.K.; Mulwad, V.; Joshi, A.; Finin, T. Cybertwitter: Using twitter to generate alerts for cybersecurity threats and vulnerabilities. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM ‘16), San Francisco, CA, USA, 18–21 August 2016; IEEE: New York, NY, USA; pp. 860–867.
19. Altalhi, S.; Gutub, A. A survey on predictions of cyber-attacks utilizing real-time twitter tracing recognition. J. Ambient Intell. Humaniz. Comput. 2021, 12, 10209–10221.
20. Hindy, H.; Atkinson, R.; Tachtatzis, C.; Colin, J.-N.; Bayne, E.; Bellekens, X. Utilising deep learning techniques for effective zero-day attack detection. Electronics 2020, 9, 1684.
21. Kim, J.-Y.; Bu, S.-J.; Cho, S.-B. Zero-day malware detection using transferred generative adversarial networks based on deep autoencoders. Inf. Sci. 2018, 460, 83–102.
22. Mbona, I.; Eloff, J.H. Detecting zero-day intrusion attacks using semi-supervised machine learning approaches. IEEE Access 2022, 10, 69822–69838.
23. Peppes, N.; Alexakis, T.; Adamopoulou, E.; Demestichas, K. The effectiveness of zero-day attacks data samples generated via GANs on deep learning classifiers. Sensors 2023, 23, 900.
24. Sarhan, M.; Layeghy, S.; Gallagher, M.; Portmann, M. From zero-shot machine learning to zero-day attack detection. Int. J. Inf. Secur. 2023, 22, 947–959.
25. Soltani, M.; Ousat, B.; Siavoshani, M.J.; Jahangir, A.H. An adaptable deep learning-based Intrusion Detection System to zero-day attacks. J. Inf. Secur. Appl. 2023, 76, 103516.
26. Won, D.-O.; Jang, Y.-N.; Lee, S.-W. PlausMal-GAN: Plausible malware training based on generative adversarial networks for analogous zero-day malware detection. IEEE Trans. Emerg. Top. Comput. 2022, 11, 82–94.
27. Zuhair, H.; Selamat, A.; Krejcar, O. A multi-tier streaming analytics model of 0-day ransomware detection using machine learning. Appl. Sci. 2020, 10, 3210.
