Towards AI-powered Cybersecurity Attack Modeling with simulation tools: Review of attack simulators

Aws Jaber$^{1*[0000-0003-2534-9551]}$ and Lothar Fritsch$^{1[0000-0002-0418-4121]}$

Department of Information Technology,
Faculty of Technology, Art and Design,
Oslo Metropolitan University
lotharfr@oslomet.no
aws@ieee.org
https://www.oslomet.no

Abstract. Cybersecurity currently focuses primarily on defenses that detect and prevent cyber-attacks. However, it is more important to regularly verify an organization's security posture to reinforce its cybersecurity defenses as the IT environment becomes more complex and competitive. Confronted with an increasing use of artificial intelligence (AI) in cyber attacks, attack simulation platforms need to allow software vulnerabilities to be found against AI-powered attacks too. Such simulators will enable defenders to maintain a basic safety level and gain control over their security posture. Gradually, we are moving towards smart and autonomous platforms. This paper reviews established cyberattack simulation scientific research techniques with the goal of presenting a selection of tools and platforms that minimize the biases and inaccuracies inherent in traditional, isolated ad hoc research on A-powered cyberattacks.

Keywords: Cybersecurity Modeling · Cybersecurity Simulation· Open Source Attack Toolkit · Modeling and Simulation of AI-powered Cybersecurity · machine learning · cyber range

1 Introduction

The attack surface is expanding due to the more complex and drastically evolving IT environment. [3]. Increasingly, attackers deploy tools using AI algorithms to increase attack performance [8]. As a result, precise target attacks with only defenses that detect and block attacks at specific points have become difficult to prevent. Many security professionals are currently using security equipment and services, according to a survey conducted at the RSA conference [1]. The attack's surface through security assessment and establishing a response strategy has been stressed in recent years . Some studies have been devoted to challenges linked to cybersecurity modeling and their simulation in cyberspace [25]. As shown in Fig. 1, which shows there have been an increase in the quantity of articles from 2020 until the present, and there has been a massive duplication of manuscripts related with cyber attack model and cybersecurity simulations[15]. However, today, there is a shortage of attack simulation platforms that will incorporate AI-powered attack functionality.

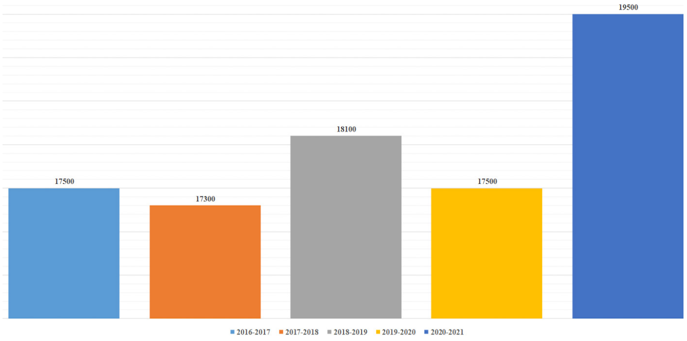

Fig. 1. Google scholar search results between 2016-2021

1.1 Research Questions and objectives

- **RQ1** What is the development status for modeling attack, and attack simulation tool?
- **RQ2**What and How is Breach and Attack Simulation (BAS) have effectiveness in cybersecurity ?

While, the research objectives are:

- **RO1**: To review systematically the most important of development status modeling attack, and attack simulation tool using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)
- **RO2**: To summarize the most trend tools for the breach and attack simulation, which will came from two rounds of snowballs[17]. And figure out the most trend of the simulation tools that came from the systematic review.

1.2 Contribution

The main contribution of this work is threefold: (a) Authors identified the main security issues with attack models and simulation attack tools; (b) Authors conducted a systematic review of attack modeling methods using two snowball rounds to inform the audience about how and what types of simulation tools are used for modeling.(c) Further, the authors provided the necessary insight into how security can be improved by adding AI-powered attack mechanisms to the simulation platforms.

2 Methodology

In this systematic review, based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[22]. The result are showed in Fig.2, the Web of Science search engine was consulted achieving a total of 330 articles. ACM Digital Library indexing engine, achieving a total of 2 articles. While, for the Google Scholar achieved 26 articles Finally, other indexed sources of articles were consulted achieving a total of 15 articles.

2.1 search criteria

We used a Boolean operators [21] , which are shown in Appendix for essential systematic keyword search. However, After refining the search terms for all databases, which shown in Table 1, the results have reported 358 research papers.

Table 1. The quantity of articles taken from three main indexed databases

No.	Keywords	Database	articles No.
1	S1	Web of science	330
2	S2	ACM Digital Library	2
3	S3	Google Scholar	26

Inclusion Brief inclusion: Look for technological content, attacks simualtions, Exclusion: focus on attacking AI.

Exclusion Result: references (scientific articles, preprints, book chapters).

3 Results

A cyber-attack simulation is a tool that can be used to check an organization's security policy and status, and it allows a security reinforcement plan to be developed by defining seven errors and four warnings, the surface of the attack that needs to be handled beforehand. [12]. The cyber-attack simulation comprises of BAS, attack graph, and penetration testing.

Needless to say, during the attack point, penetration testing , typically also used for simulated hacking or analysis of the vulnerability, is primarily used manually, and its importance varies according to the user's expertise [18]. The latest penetration testing aims to provide a vulnerability scanning feature to locate vulnerabilities and a feature to exploit an actual vulnerability. The attack graph can define all possible paths for an attack by considering the connections between points, moving one step further from the 'points'-based vulnerability search, which was the aim of the current penetration testing [19] . The graph

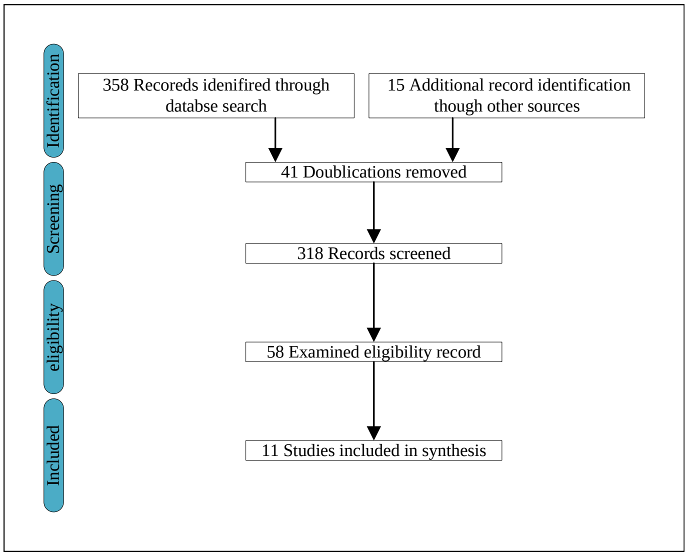
Identification
Screening
eligibility
Included
358 Recoreds idenifired through databse search
15 Additional record identification though other sources
41 Doublications removed
318 Records screened
58 Examined eligibility record
11 Studies included in synthesis

Fig. 2. Research Methodology based on PRISMA

path's priority, which is likely to be used for the attack, is being determined and presented to the user at this time. However, there is no assurance that all attack paths discovered using the attack graph will be valid or that the attacker will use those paths[13]. The BAS can obtain the evaluation result by simulating a valid simulated hacking based on the attack scenario, unlike the attack graph , which identifies the attack path through static analysis[5]. Most BAS products include scenarios based on existing attacks and allow users to create their attack scenarios. However, scenarios identified by internal staff who are well aware of the organization's IT infrastructure environment are standardized, and BAS is also vulnerable to zero-day.

To enable penetration testing, a number of open-source and paid resources are available[11] . However, most professionals use more than one tool to use different functions according to specific attack techniques[20], or it is useful for simulated hacking testing for more advanced hackers[26]. In most of the cases, the tools are developed and updated by themselves. Recent penetration testing technologies, as described above, aim to automate vulnerability analysis and exploitation steps, and this paper describes representative tools that provide an automated penetration testing function.

Metasploit is an open-source community-supported penetration testing platform that offers scanning and penetration testing of vulnerabilities[10]. Metasploit uses more than 4,000 exploit modules and more than 150,000 vulnerabilities to provide penetration testing and analysis and response methods, and the database is continually updated with the help of the open-source community.

Burp is a security testing program for web applications developed by Portswigger Web Protection. It is used as a web vulnerability scanner because its scanning feature has more power than the penetration function.You can use Burp Proxy in the Community version to intercept web traffic and analyze and manipulate content. Moreover, it offers a function to use the Clickbandit tool to conduct a clickjacking attack against a vulnerable service.

Currently, the **Canvas of Immunity** offers an automated exploit system and runs a development framework based on 800 exploits[24]. It is a commercial tool and consists of a framework for the attack and a suit to test penetration.It is possible to conduct a simulated attack on web applications through one system and the network range setting.It could be used by changing the Canvas engine according to your requirements. Currently, Canvas generates shell code automatically. This set of procedures enables the user to conduct penetration testing by choosing them according to the menu. Although Canvas contains modules and information that can be performed for the host, users have to use hacking professionals to decide what to do with the target host and interpret and use the results.

Core Impact Pro is the Core Security's penetration testing tool, which offers the most commercial-grade exploits[14]. Each month, the exploit and SCADA exploit packages provided by Metasploit can be combined and used with 40 exploit codes generated by themselves. The SCADA exploit kit explicitly includes more than 140 exploits that target ICS and SCADA. It can also be combined with other tools such as PowerShell Empire or Metasploit for penetration testing. Core Impact has an advantage in allowing automatic exploitation and pivoting of adjacent devices as compared to Metasploit, and intuitively presenting pivoting flow also in its stealth mode.

BAS enables an automated simulation of multi-level cyber-attack scenarios [4]. For this, we are modeling a chain of attacks that would possibly be used by actual attackers to target the IT environment. The key difference between the BAS and the previous simulation is that it can automatically execute the user-selected attack scenario.

The **FireDrill** platform of **AttackIQ** constantly verifies an organization's security program's efficacy and ensures that security products and services respond to simulated attack scenarios appropriately [4]. A library of more than 1,500 individual attacks is being developed as a platform for evaluating compliance. It provides the function of automatically performing security evaluations using the attack scenario library generated or provided by the security individual in charge. The AttackIQ platform contains a management console and an agent that can be installed on-premises, in virtual environments, or in the cloud, and the agent is responsible for executing the exploit by obtaining the selected scenario from the AttackIQ platform's sensor[7]. To provide such a function of scenario-based attack simulation, community support must continuously provide a library that allows an attack scenario to be implemented, which is considered a difficult point when adopting the scenario-based simulation.

SafeBreach offers a simulation platform to test the infrastructure's security status using simulated assets and actual attack techniques[2]. It positions fake assets, simulating real assets in selected segments, executes scenarios of attacks between these simulators, and allows for endpoint simulators, network, and cloud[6]. It also enables us to choose the type of data to attack, such as source code, personal identifying information, and credit cards. The SafeBreach platform enables us to test the attack safely without impacting real assets and resources by executing an attack against a simulated environment.

Cymulate offers an automated method for determining company security infrastructure and activity using real asset attacks without actual security breaches [16]. The BAS platform of Cymulate runs between software agents named ”Cymulate's Hopper modules” deployed on real assets or between software agents and the cloud of Cymulate. The software agent is installed on the target device, and the actual malware is downloaded from the SaaS solution. The evaluation is performed by additionally extending the attack vector, beginning with one attack vector as the target. Threat vectors are lateral movement, corporate web application attacks, and web browsing or e-mail[9]. A cyber-attack simulation needs to be developed to represent different social engineering factors to find vulnerabilities in the system's operations and business processes beyond finding the system's vulnerabilities[23].

4 Discussion and Conclusion

4.1 Discussion

The trend of cyber-attack simulation for security assessment has been identified as a preemptive countermeasure against cyber-attacks. Cyber-attack simulation supports a feature that can execute a simulated attack to prepare for an actual attack, and BAS and attack graph are gaining much attention from the already widely used penetration testing technologies. Cyber-attack simulation is in the stage of progressive development from the current manual analysis, which relies on the user's ability to use an automated analysis. This change is due to the reason that a one-time assessment by costly manual inspection has restrictions on the rapid detection and diagnosis of bugs and vulnerabilities bound to occur in rapidly and drastically evolving IT environments. Thus, safety can be measured by repeated and continuous execution, not just at a specific point in time, and it is required that can help non-professional users understand the security status. The existing tools provide a reasonable framework for the integration of AI-powered attack tools deploying machine learning and other technologies.

Personal data is a sensitive asset in information systems. Unauthorized access to personal data is a severe security breach that can result in fines. Six protection goals are normally used when engineering privacy: confidentiality, integrity, availability, unlinkability, transparency, and intervenability. Automated testing can target the first three categories. Two areas are of special interest: the protection of personal data against unauthorized access and the protection of digital human identities and related identifiers against de-anonymization, linkage, or observation. Mass extraction of data from data stores is often prevented using heuristics in combination with machine learning. Intrusion detection systems will react to unusual access or transfer patterns to data stores. For AI attack simulation, both the simulated generation of data exfiltration traffic and the testing of detective controls must be part of the testing simulations. Encrypted network transfers have become under attack through machine learning, where classifiers successfully identify communication content and communication targets in encrypted data traffic. Concerning the protection of identities, a considerable attack surface has to be taken into account. Direct exploitation of available identity attributes, person-relateable attributes, and inferred attributes are feasible extraction attacks easily enhanced and automated by machine learning. Intrusion detection systems have long used user ID and user behavior in detecting data access and activity deviations.

4.2 Conclusion

We conclude therefore that the integration of artificial intelligence attack methods into the cyber-security testing is a mandatory path into the future of these tools. However, the field of simulated offensive, AI-powered cyber-attacks is not well-developed, and should be considered in future research.

References

1. Ahmed, K.: Canada's cyber security in a globalized environment: Challenges and opportunities. In: Routledge Companion to Global Cyber-Security Strategy (2021)
2. Badhwar, R.: Oem and third-party sourced application and services risk. In: The CISO's Next Frontier, pp. 335–344. Springer (2021)
3. Barra, M., Dahl, F.A., Vetvik, K.G., MacGregor, E.A.: A Markov chain method for counting and modelling migraine attacks. Scientific Reports **10**(1) (2020). https://doi.org/10.1038/s41598-020-60505-5
4. Fu, Y., O'Neill, Z., Wen, J., Adetola, V.: Evaluating the impact of cyber-attacks on grid-interactive efficient buildings. In: ASME International Mechanical Engineering Congress and Exposition. vol. 85642, p. V08BT08A047. American Society of Mechanical Engineers (2021)
5. Ho, G., Dhiman, M., Akhawe, D., Paxson, V., Savage, S., Voelker, G.M., Wagner, D.: Hopper: Modeling and detecting lateral movement. In: 30th USENIX Security Symposium (USENIX Security 21). pp. 3093–3110 (2021)
6. Jaber, A.N., Anwar, S., Khidzir, N.Z.B., Anbar, M.: The importance of ids and ips in cloud computing environment: Intensive review and future directions. In: International Conference on Advances in Cyber Security. pp. 479–491. Springer (2020)
7. Jaber, A.N., Anwar, S., Khidzir, N.Z.B., Anbar, M.: A Detailed Analysis on Intrusion Identification Mechanism in Cloud Computing and Datasets. In: Communications in Computer and Information Science. vol. 1347 (2021). https://doi.org/10.1007/978-981-33-6835-4_37

8. Jaber, A.N., Fritsch, L.: Covid-19 and global increases in cybersecurity attacks: Review of possible adverse artificial intelligence attacks. In: 2021 25th International Computer Science and Engineering Conference (ICSEC). pp. 434–442 (2021). https://doi.org/10.1109/ICSEC53205.2021.9684603
9. Jaber, A.N., Fritsch, L., Haugerud, H.: Improving phishing detection with the grey wolf optimizer. In: 2022 International Conference on Electronics, Information, and Communication (ICEIC). pp. 1–6. IEEE (2022)
10. Jaswal, N.: Mastering Metasploit: Take your penetration testing and IT security skills to a whole new level with the secrets of Metasploit. Packt Publishing Ltd (2018)
11. Jayasuryapal, G., Pranay, P.M., Kaur, H., et al.: A survey on network penetration testing. In: 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM). pp. 373–378. IEEE (2021)
12. Kour, R., Thaduri, A., Karim, R.: Predictive model for multistage cyber-attack simulation. International Journal of Systems Assurance Engineering and Management **11**(3) (2020). https://doi.org/10.1007/s13198-020-00952-5
13. Lallie, H.S., Debattista, K., Bal, J.: A review of attack graph and attack tree visual syntax in cyber security. Computer Science Review **35**, 100219 (2020)
14. Lu, K.C., Liu, I.H., Li, J.S.: A survey of the offensive and defensive in industrial control system. Bulletin of Networking, Computing, Systems, and Software **11**(1), 1–6 (2022)
15. Macak, M., Daubner, L., Sani, M.F., Buhnova, B.: Cybersecurity analysis via process mining: A systematic literature review. In: International Conference on Advanced Data Mining and Applications. pp. 393–407. Springer (2022)
16. Moyal, M.: Home page (Jan 2022), https://cymulate.com/
17. Naderifar, M., Goli, H., Ghaljaie, F.: Snowball sampling: A purposeful method of sampling in qualitative research. Strides in Development of Medical Education **14**(3) (2017)
18. Qian, K., Zhang, D., Zhang, P., Zhou, Z., Chen, X., Duan, S.: Ontology and Reinforcement Learning Based Intelligent Agent Automatic Penetration Test. In: 2021 IEEE International Conference on Artificial Intelligence and Computer Applications, ICAICA 2021 (2021). https://doi.org/10.1109/ICAICA52286.2021.9497911
19. Refat, R.U.D., Elkhail, A.A., Hafeez, A., Malik, H.: Detecting CAN Bus Intrusion by Applying Machine Learning Method to Graph Based Features. In: Lecture Notes in Networks and Systems. vol. 296 (2022). https://doi.org/10.1007/978-3-030-82199-9_49
20. Sarker, I.H.: Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective. SN Computer Science **2**(3), 1–16 (2021)
21. Scells, H., Zuccon, G., Koopman, B.: Automatic boolean query refinement for systematic review literature search. In: The world wide web conference. pp. 1646–1656 (2019)
22. Selçuk, A.A.: A guide for systematic reviews: Prisma. Turkish archives of otorhinolaryngology **57**(1), 57 (2019)
23. Shakir, H.A., Jaber, A.N.: A short review for ransomware: pros and cons. In: international conference on P2P, parallel, grid, cloud and internet computing. pp. 401–411. Springer (2017)
24. Singh, H., Jangra, S., Verma, P.K.: Penetration testing: analyzing the security of the network by hacker's mind. Volume V IJLTEMAS pp. 56–60 (2016)
25. Snider, K.L.G., Shandler, R., Zandani, S., Canetti, D.: Cyberattacks, cyber threats, and attitudes toward cybersecurity policies. Journal of Cybersecurity **7**(1) (2021). https://doi.org/10.1093/cybsec/tyab019

26. Yaacoub, J.P.A., Noura, H.N., Salman, O., Chehab, A.: A survey on ethical hacking: Issues and challenges. arXiv preprint arXiv:2103.15072 (2021)

A Boolean Search string

Web of science
S1= ”(ALL=((cybersecurity modeling AND Simulations in Cyber-Security OR cybersecurity simulation) OR (cyber threat emulations AND attack Detection and Response (ADR)) (simulated attacks AND open source attack toolkit OR evolving attack methodology) AND (attacker OR attacking tool) AND (Modeling and Simulation of Behavioral Cybersecurity)))”

ACM
S2 =,”query”: AllField:(” cybersecurity modeling ” AND ”Simulations in Cyber-Security” OR ”cybersecurity simulation ” OR ”cyber threat emulations” OR ”attack Detection and Response (ADR) ” AND ” simulated attacks” AND ”open source attack toolkit ” OR ”evolving attack methodology” AND ”attacker OR attacking tool” AND ”Modeling and Simulation of Behavioral Cybersecurity”) ”filter”: Publication Date: (01/01/2016 TO 12/31/2021),ACM Content: DL Google syntax
S3 = cybersecurity modeling AND Simulations in Cyber-Security OR cybersecurity simulation OR cyber threat emulations OR attack Detection and Response (ADR) AND simulated attacks AND open source attack toolkit OR evolving attack methodology AND attacker OR attacking tool) AND (Modeling and Simulation of Behavioral Cybersecurity .