An Investigation into Artificial Intelligence and Automation within Penetration Testing and Cybersecurity

Word Count: 3344

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Robert Mustacchio

Spring 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Richard Jacques, Department of Engineering and Society

An Investigation into Artificial Intelligence and Automation within Penetration Testing and Cybersecurity

Introduction

In recent years, the cybersecurity industry has seen a massive increase in the level of sophistication of cybercriminals and their attacks. As a result, organizations have started to turn to automation and artificial intelligence (AI) for their security solutions, and this has become a major development in the field of penetration testing. The goal of this paper is to investigate the new automated and AI-based penetration test solutions that have begun to surface in the penetration testing market. Specifically, it will explore the effectiveness of AI-based penetration tests in comparison to traditional, manual penetration tests including the technical advantages and drawbacks of them and the various potential concerns and questions pertaining to the security and ethical aspects of these tools and the organizations that use them.

The technical/capstone project associated with this topic starts with an internship I completed in 2022 for a southern U.S. cybersecurity firm that involved researching various cybersecurity firms to look for gaps in their penetration test offerings, modern technology and methods within the industry, and openings in the market for this firm to exploit. It was during this research when I discovered that cybersecurity firms were starting to unveil penetration test solutions that utilized AI to automate certain components of a penetration test to even a full test. Although not required for my internship, I decided to further investigate these automated penetration test solutions to find out if they were just a fad or if AI and automation are the future of penetration testing. This investigation would be centered around a comparison between traditional manual penetration tests and the new AI-based ones, seeing if either had any advantages or drawbacks when compared to the other.

The STS topic associated with this problem surrounds various potential security and ethical concerns that could be associated with the incorporation of AI and automation into penetration testing and the cybersecurity industry in general. Security concerns such as the potential for these automated tools to be manipulated for malicious use and the potential for over-reliance on AI and a lack of human oversight will be investigated and their validity will be examined, as will those of potential ethical concerns and implications such as the potential for massive job loss of security professionals and the privacy violations that could arise from the increase in AI and automation.

Literature Review

As the existence of AI within the field of penetration is relatively new, the collection of scholarly literature on the topic is relatively limited. However, there are still multiple scholarly papers and articles that investigate current or proposed automated penetration test tools, like the paper by Abu-Dabaseh and Alshammari (n.d.) that gives an overview of the current state of automated penetration testing, including penetration testing standards as well as comparisons between manual and automated penetration testing. Another example is a paper by Richard McKinnel et al (2019), which gives a snapshot and analysis of the current state of artificial intelligence in penetration testing and provides analysis of further work that needs to be done.

While sources like these were certainly helpful to this paper as they provided technical investigations of these automated penetration testing tools, there was no accessible literature that dealt with other aspects of this technology such as the ethical implications and security concerns that could be associated with it. This technology is still quite new, which is perhaps why these ethical and security concerns are not being given enough attention. There may not be enough of these tools to conduct in-depth research on their current or potential ethical implications or security concerns because many cybersecurity firms and organizations are still developing AI-based security solutions. As more of these products hit the market, there needs to be discussion about any possible effects or consequences before they're widely used and accepted in the industry.

The goals for the rest of this paper are to summarize what is known in the field about the technical side of this technology and to research and analyze any major potential issues or concerns with the security and ethics of these AI-based penetration test tools and solutions. The ultimate goal is that this paper will serve as a starting point for industry professionals and researchers to further explore and address some of the security and ethical topics that are discussed.

Methodology Used

A thorough research approach must be used to adequately answer the research question. To address the technical portion of the research question, the technical components of penetration testing and the existing differences in penetration test methodology and effectiveness between manual and AI-based penetration testing must be researched. To address the STS portion of the research question, any major potential concerns, questions, or implications that could be raised regarding the security and ethical sides of this technology must be researched and analyzed. The research approach will be qualitative with the goal of collecting and analyzing any data that can provide insight into the technical advantages and disadvantages, current and potential security concerns, and ethical implications that are associated with the current and future use of artificial intelligence in penetration testing. The data that is collected will also be secondary, meaning that it has already been collected by other researchers and organizations. The data will include scholarly papers, reports, and articles related to the use of artificial intelligence in penetration testing, as well as any relevant industry reports or articles. This data could also include works about artificial intelligence in general, with which certain aspects could be applied to the field of penetration testing or cybersecurity. The data will also be descriptive, offering a summary and analysis of the current state of AI in penetration testing.

The research for this paper will focus on finding sources that address any of the technical, security, and ethical aspects and topics that relate to the research question. This will ensure that the research is supported by substantial evidence, offers a thorough understanding of the topic and aspects of the research question, and also draws attention to the need for additional research to be conducted by security professionals to assess the validity of the concerns and questions raised by the findings of this paper.

Penetration Testing + AI Overview

Penetration testing is among the most popular security measures that organizations around the world take to protect their data, information, and systems. It refers to the process of identifying vulnerabilities within a system and exploiting them to understand the level of threat they pose and the damages that could be caused by an attack (Keshri, 2022). There are several types of penetration tests, or pen tests, the most prevalent being network, web and mobile application, security system, and social engineering pen tests. Penetration testing is a lengthy and arduous process that requires a great deal of attention and effort from security professionals. Because of this, cybersecurity firms have started to develop pen tests that use artificial intelligence (AI) to automate several parts of the pen test process to even fully automated tests. Firms are using AI to create automated vulnerability scanners, which can continuously run and search for vulnerabilities in an organization’s systems with little to no oversight and report any vulnerabilities that were found to security professionals who would then go in and attempt to exploit and patch them. In fact, there are currently more than 90 of these automated vulnerability scanners being offered by firms and organizations in the market (OWASP, n.d.). A more recent development is that firms are creating fully autonomous penetration tests that use AI to perform every component of a full penetration test. A notable instance of this is the tool Deep Exploit, an open-source fully automated penetration testing tool that uses AI and Machine Learning (Son, 2020). While the emergence of these tools is a recent development, the direction of the penetration testing and cybersecurity field is moving towards a more automated landscape, making it imperative to thoroughly investigate these new tools and methods.

Technical Comparisons Between Manual and Automated Pen Tests

One critical area of focus when exploring these new forms of penetration tests is the technical comparisons between them and the fully manual tests performed by security professionals. While both methods share the same goal of identifying and exploiting vulnerabilities in an organization's security infrastructure, they each have their own benefits and drawbacks. The technical portion of the research question and this paper surrounds an investigation into the various advantages and disadvantages of each type of pen test and will ultimately assess the prospects of automated pen tests fully taking over the penetration testing market.

As mentioned above, manual penetration tests are those that are performed entirely by security professionals while automated penetration tests are those that have either partially automated or fully automated the penetration test process. As mentioned in the paper by Abu-Dabaseh and Alshammari, when compared through a technical lens, there are notable differences between these two types of tests (n.d.). Firstly, the overall testing process is significantly different for each type, with the manual testing process being very labor, cost-intensive, and difficult to repeat, while the automated testing process is much faster, cheaper, and easier to repeat on an organization's system. Another major difference can be found in the reporting and cleanup components of the testing process, with the manual tests requiring time and effort from security professionals to thoroughly report all vulnerabilities that were found and exploits that were completed and then more time and effort to manually undo all the system changes that occurred during this, while automated solutions offer customized and automated reporting and cleanup. Additionally, fully manual penetration tests make significant modifications to an organization's systems and networks that need to be undone, while automated penetration tests work with those systems and networks being unchanged.

While the above differences certainly make the automated penetration tests seem far more favorable than the manual ones, there are differences between the two types of tests that explain why organizations still opt for manual tests. Firstly, automated solutions lack the ability to predict how a human hacker will attempt to access a network or system, which ultimately results in tests conducted by human security professionals to be better equipped at identifying certain types of vulnerabilities that automated solutions routinely overlook. Automated tests are also much less effective at identifying vulnerabilities related to social engineering, as these tend to deal with factors like psychology and human error. Another drawback of these automated tests is that they currently have a habit of producing false results in their tests. These tests have been known to report false positives, or warnings about vulnerabilities that either do not exist or are not even a vulnerability, and occasionally false negatives, or simply overlooking vulnerabilities that are actually there. A final key difference between manual and automated tests is that while automated test solutions can generate reports on the vulnerabilities that were found and exploits that were conducted, they lack the ability to provide advice to an organization on next steps after penetration tests are performed (Nesbo, 2022).

While automated penetration testing tools and solutions have made significant strides in recent years and have already begun helping various organizations and businesses strengthen and maintain their cybersecurity, manual penetration testing remains a valuable part of the penetration testing landscape. The outlook for a primarily automated penetration testing market is bright, with automated tests being much less costly and labor-intensive while also rivaling manual tests in effectiveness. However, it is important to remember there are still many advantages of manual testing that automated tools cannot yet replicate. As such, organizations should continue to invest in both manual and automated penetration testing methods to ensure holistic security coverage.

Discussion of Potential Security Concerns and Ethical Implications

As stated above, there has been virtually no scholarly literature that addresses the topics of security and ethics associated with AI-based penetration testing. While analyzing the effectiveness of these tools is certainly important, it is also incredibly important to consider the security and ethical aspects of this technology. The goal of this paper is to draw more attention to these subjects and kickstart the research process for what is surely to become a critical research component of the penetration testing and overall cybersecurity field. This section will explore some of the potential security risks and ethical consequences and issues that are associated with AI-based penetration testing and will highlight the need for further investigation in these areas.

While the outlook on AI-based penetration testing is bright, the current state of the technology raises many questions about the overall security of these tools and that of the organizations that have begun to use them. The first topic of concern regards the potential for these tools to be manipulated and weaponized by cybercriminals. In her article about cybercriminals using AI to commit cybercrime, Jennifer Gregory writes about how cybercriminals are starting to develop their own AI-based tools to conduct attacks on an organization’s systems and networks. These attack tools could continually launch attacks on systems that are protected with AI-based tools, learn about how those tools are protecting the system, and potentially create attacks that take advantage of their behavior and develop ways to work around them (Gregory, 2021). This raises the level of urgency for these security tools to also be able to learn from the attack AI and be able to quickly predict their next attack behavior. Another potential security concern of these AI-based tools is that their widespread use could lead to an over-reliance on them and a lack of human oversight in the penetration testing process. While an increased adoption of automated penetration testing tools would certainly lead to less time, money, and effort consumed, a decrease or eventual lack of human oversight regarding penetration tests are a massive point of concern to security professionals. As mentioned above, these AI-based tools currently have a habit of producing false positives or false negatives in penetration test results. With fewer human eyes on the entire penetration test process, these false results could go unnoticed, which could potentially lead to either catastrophic security breaches or massive wastes of an organization’s time and capital. Additionally, a lack of human oversight in the penetration test process could potentially lead to decreased creativity and adaptability in the testing process, meaning that organizations would be increasingly reliant on the ability of automated tools to keep up with the evolving nature and behavior of cyber threats.

There are also several ethical questions and concerns with the impact of artificial intelligence in penetration testing and the cybersecurity field that remain unexplored. The first of which is the potential impact that AI-based penetration testing tools will have on employment within the cybersecurity field. As automated and AI-driven cybersecurity tools become increasingly sophisticated, they are expected to take over job roles such as Cybersecurity Engineers, Vulnerability Assessments, and SIEM Engineers, among others (Ijlal, 2023). AI has already started to replace tens of thousands of different jobs in several industries all over the world, but this is a new concern within the cybersecurity industry, which should be investigated further. Additionally, if AI starts to take over cybersecurity jobs, it could lead to a decrease in human knowledge and expertise in the field of cybersecurity because these tools might not be able to match the same level of human insight and judgement of cybersecurity professionals. This raises questions about the balance between technological advancement within the cybersecurity industry and job security for human professionals in the industry. Another ethical concern is also one that has been regularly discussed when talking about artificial intelligence but not specifically within penetration testing or cybersecurity, which is the potential for biases to work their way into the penetration testing process. Since these AI-based penetration testing tools are trained on datasets and security practices and patterns, if said data are biased in some way, the AI may inadvertently introduce these biases into its penetration test process. For example, the AI could only have been trained using certain types of vulnerabilities, which may leave it unable or less effective at identifying distinct types of vulnerabilities or vulnerabilities in several types of systems, networks, or applications. It is important to be aware of these potential biases and to take steps to ensure that they do not work their way into these AI-based penetration testing tools so that they are as effective and accurate as possible. A final major ethical question when it comes to AI-based penetration testing tools is about potential privacy implications and violations. A big issue with AI in all industries is that AI systems require large amounts of personal data, and if this data falls into the wrong hands it can be used for nefarious purposes, such as identity theft (Van Rijmenam, 2023). This poses a huge potential concern for penetration testing, because as AI-based penetration testing tools become more advanced, they may be able to gather and analyze significant amounts of data about people and organizations, such as personal or confidential organizational information, which could result in potential violations of privacy rights.

While AI-based penetration testing is a very promising development in the field of cybersecurity, there are a number of security and ethical issues and concerns that could arise and must be carefully considered before fully embracing its use. Although this paper does examine some of the main concerns that could potentially arise with AI-based penetration testing, there will undoubtedly be many more concerns, questions, and topics that will require equal attention by professionals within the field. As cybersecurity firms and organizations continue to develop these tools and explore their potential benefits and advantages, it is critical that they also address the potential risks and concerns associated with their deployment to ensure that they are used in a responsible manner.

Conclusion

The goal of this paper was to conduct research on AI-based penetration testing technology and solutions. This research included a technical comparison of AI-based penetration testing solutions and the traditional, manual penetration test methods and a discussion of various potential concerns, questions, and topics that relate to the security and ethical aspects of this technology. As this technology continues to advance and develop within the industry, special attention must be paid to it to prevent putting the security and privacy of people and organizations at risk. The research conducted for this paper and its results indicate that while these tools will certainly become more effective and common in the industry, there are still significant security and ethical concerns and questions that need to be addressed before these tools become more widespread. The technical comparison between the AI-based and manual penetration tests showed that while the AI-based tests do have several benefits over manual tests, they still have various limitations such as an inability to adapt to certain systems and a habit of producing false test results. Additionally, there are many security concerns with this technology and its use, such as the potential for these tools to be manipulated or the potential for the cybersecurity industry to become over-reliant on them. These tools also have a number of ethical concerns and implications that need to be addressed, such as their impact on the employment of cybersecurity professionals and the potential for them to violate privacy rights of individuals and organizations. The lack of scholarly and industry research that addresses the security and ethical aspects of this technology demonstrates the urgency for which research on these topics must be conducted by professionals within the field of cybersecurity.

It is recommended that future research and analysis of the security and ethical components of this technology be conducted by cybersecurity professionals, including the various potential concerns and questions that were discussed in the body of this paper. While the future of AI-based penetration testing is promising, it is important to acknowledge the potential risks and limitations of these tools. The security and ethical concerns and questions raised in this paper must be carefully considered and addressed to ensure the responsible and effective use of this technology in the future.

References

Abu-Dabaseh, F., & Alshammari, E. (n.d.). *AUTOMATED PENETRATION TESTING: AN OVERVIEW* . airccj.org. Retrieved March 1, 2023, from https://airccj.org/CSCP/vol8/csit88610.pdf

Gregory, J. (2021, May 15). *AI Security Threats: The Real Risk Behind Science Fiction Scenarios*. Security Intelligence. Retrieved March 1, 2023, from https://securityintelligence.com/articles/ai-security-threats-risk/

Ijlal, T. (2023, January 26). *Will AI replace cybersecurity jobs ?* Medium. Retrieved March 1, 2023, from https://pub.towardsai.net/will-ai-replace-cybersecurity-jobs-7f3abc7a987

Keshri, A. (2022, October 10). *What is Automated Penetration Testing? Difference Between Automatic & Manual Pentesting*. Astra Security Blog. Retrieved March 1, 2023, from https://www.getastra.com/blog/security-audit/automated-penetration-testing/#:~:text=Automated%20penetration%20testing%20or%20Vulnerability,performed%20by%20competent%20security%20researchers

McKinnel, D. R., Dargahi, T., Dehghantanha, A., & Choo, K.-K. R. (2019, March 4). *A systematic literature review and meta-analysis on artificial intelligence in penetration testing and Vulnerability Assessment*. Computers & Electrical Engineering. Retrieved March 1, 2023, from https://www.sciencedirect.com/science/article/abs/pii/S0045790618315489

Nesbo, E. (2022, June 27). *Manual vs. Automated Penetration Testing: What's the difference?* MUO. Retrieved March 1, 2023, from https://www.makeuseof.com/manual-vs-automated-pen-testing/

Son, D. (2020, May 9). *Deep Exploit: Fully automatic penetration test tool using Machine Learning*. Security Online. Retrieved March 1, 2023, from https://securityonline.info/deep-exploit/

Tjoa, S., & Kieseberg, P. (2020, October 1). *Penetration Testing Artificial Intelligence*. ResearchGate. Retrieved March 1, 2023, from https://www.researchgate.net/profile/Peter-Kieseberg/publication/349172682_Penetration_Testing_Artificial_Intelligence/links/6023ba3ca6fdcc37a8163a28/Penetration-Testing-Artificial-Intelligence.pdf

van Rijmenam, M. (2023, February 17). *Privacy in the age of AI: Risks, challenges and solutions*. The Digital Speaker. Retrieved March 1, 2023, from https://www.thedigitalspeaker.com/privacy-age-ai-risks-challenges-solutions/#:~:text=The%20Issue%20of%20Violation%20of%20Privacy&text=One%20of%20the%20primary%20challenges,as%20identity%20theft%20or%20cyberbullying\

Vulnerability Scanning Tools. Vulnerability Scanning Tools | OWASP Foundation. (n.d.). Retrieved March 1, 2023, from https://owasp.org/www-community/Vulnerability_Scanning_Tools