The Same, but Different: The Pentesting Study

Jan Roring, Dominik Sauer, Michael Massoth
Department of Computer Science
Hochschule Darmstadt — University of Applied Sciences Darmstadt
Darmstadt, Germany
E-mail: jan.roring@stud.h-da.de, {dominik.sauer,michael.massoth}@h-da.de

***Abstract*—When ordering a penetration test, customers assume that they will receive the same results regardless of who performs the testing. Although well-known standards are commonly used to ensure that results of penetration tests are consistent and reproducible, these results vary widely depending on the chosen service provider. To evaluate this, we had two penetration tests carried out on the same IT environment by independent service providers. While there was some overlap, the results show that the human component has a profound impact on the outcome of a penetration test.**

***Keywords*—penetration test; comparison; standards; human; soft skills.**

I. Introduction

As public reports of the German Federal Criminal Police Office show, cybersecurity incidents are on the rise [1]. To protect themselves, more and more companies have the security of their IT systems and applications checked by security experts. In order to identify security vulnerabilities in these technologies, it is common practice to carry out penetration tests [2]. In addition to the identification of threats, a penetration test also includes a risk analysis of each vulnerability, as well as remediation advice, which helps clients to address the most critical issues first [3][4]. In order to provide the best possible added value, as many security vulnerabilities as possible should be identified, so they can be fixed by the client to strengthen the company's security posture.

This paper aims to show how much the quality of penetration tests varies depending on the tester. To show the variability of outcome, two penetration tests are performed on the same IT environment by two independent service providers. To achieve the fairest comparison, the same general conditions apply to both penetration testers. Thereafter, we evaluate how much the results of the two penetration tests differ and what influence the penetration testers have on them.

This paper is structured as follows: Section II provides some background information on penetration testing. Section III describes the experimental setup. The results will be presented and discussed in Section IV. Section V contains our conclusion as well as an outlook on further research opportunities.

II. Background

The following section will be dedicated to the terminology relevant to the paper. In addition to a definition of the term 'Penetration testing', it also includes commonly used standards, as well as the skill set required of a penetration tester.

A. Penetration testing

Penetration tests are used to check the security of applications, individual systems or entire networks by simulating an attack by a hacker. The penetration tester uses the techniques and tools of a hacker to uncover security vulnerabilities in the IT environment under review. If possible, identified vulnerabilities are exploited by the penetration tester to prove their existence and investigate possible impacts to better assess the threat potential of a vulnerability. Upon completion of the penetration test, the customer receives a report listing all vulnerabilities found, including a risk assessment and recommendations for remediation. The aim is to find and fix security vulnerabilities before a potential attacker can exploit them [5]–[7].

B. Commonly used standards

While a hacker may only need a single vulnerability to gain access, the penetration tester always tries to uncover every possible vulnerability [8]. To ensure that no obvious vulnerabilities are overlooked and results are reproducible, a structured approach is required. Thus, most penetration testers rely on well-known standard approaches when performing penetration tests [9].

Several attempts have been made by governments and the IT security community to standardize the penetration testing process. Therefore, there is a wide choice of standards, each with its own advantages and disadvantages. There is no universal standard that is suitable for all types of penetration tests. Government contracts often require compliance with standards published by the respective national authorities, such as the National Institute of Standards and Technology (NIST) [5] in the US or the Federal Office for Information Security (BSI) [6] in Germany. In addition, there are established standards, such as the Open Source Security Testing Methodology Manual (OSSTMM) [10] or the Penetration Testing Execution Standard (PTES) [11] that are maintained by the IT security community.

Apart from the differing terminology, the process described in each of the previously mentioned standards always has a similar basic structure [7]. It can be divided into several phases, which can be seen in Figure 1.

Some approaches such as BSI [6], PTES [11] or OSSTMM [10] give actionable instructions on what checks to perform in each phase. A more detailed look at these checks reveals that these standards are primarily designed to investigate IT

ISBN: 978-1-61208-893-8

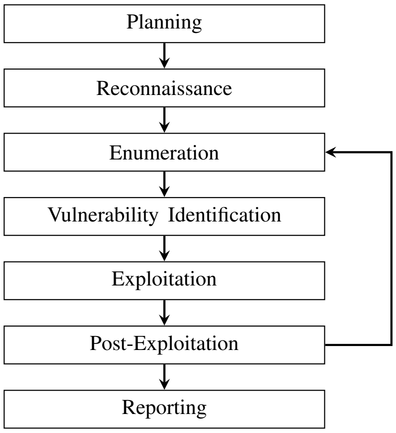


Figure 1. Penetration Testing Process

infrastructures [12]. When it comes to performing penetration testing for web applications, penetration testers typically refer to the OWASP Web Security Testing Guide, which is specifically tailored for this use case [2]. However, this is almost exclusively limited to the technical aspects of web application penetration tests, which is why some penetration testers tend to combine it with one of the aforementioned standards. In addition to such combinations, penetration testers also frequently use their own individual approaches based on the established standards [13][14].

The selection of an appropriate methodology is crucial for the success of a penetration test, as it determines what should be tested and how the penetration tester should proceed. Since the commissioned penetration testers relied on the BSI penetration testing model and the OWASP Web Security Testing Guide, these will be described in a bit more detail below:

a) BSI Penetration Testing Model: In 2003, the German Federal Office for Information Security presented a penetration testing model, which divides penetration tests into the following five phases [6]:

1) Preparation
2) Reconnaissance
3) Analyzing information and risks
4) Active intrusion attempts
5) Final analysis

Throughout the preparation phase, the objectives, scope and further general conditions of the penetration test, like time frame and target systems, are defined together with the customer. In addition, a suitable penetration test is classified and written approval is obtained from the client.

The reconnaissance phase is used to gather information on the target. This includes performing ping and port scans, as well as identifying operating systems and running services to determine possible entry points for an attacker. The tests to be performed are grouped into so-called I-modules. Suitable modules are selected based on the classification made previously.

In the subsequent phase ('Analyzing information and risk'), the previously gathered information is evaluated and potential risks are identified by looking for software versions with known vulnerabilities. In addition, the penetration tester manually searches for common types of vulnerabilities to identify new or more complex vulnerabilities within systems and applications.

To verify the actual existence of vulnerabilities, attempts are made to exploit them in the fourth phase. Through exploitation, the penetration tester aims to gain access to the affected system or read out sensitive data, which may help to escalate privileges or compromise additional systems. So-called E-modules comprise the tests that are carried out during this phase. E-modules, as well as I-modules, are based on the test points of the OSSTMM.

In the last phase ("Final Analysis"), the findings of the penetration test are reviewed and the resulting risks are assessed, depending on which sensitive data could be viewed and which systems could be accessed. Additionally, an action plan is developed with recommendations that can assist in addressing the identified vulnerabilities.

Each phase of the penetration testing process is documented to ensure the reproducibility of the test results and findings. Based on this progress log, a final report is then prepared for the customer, which contains a list of all identified vulnerabilities together with risk assessment and recommendations.

b) OWASP Web Security Testing Guide: The Open Web Application Security Project (OWASP) is a non-profit organization that aims to improve the security of web applications. To achieve this goal, OWASP works closely with the IT security community and provides valuable tools and information through open-source projects [15][16].

One of these projects is the OWASP Web Security Testing Guide, which was released in version 4.2 towards the end of 2020. In addition to the OWASP Testing Framework for developing secure web applications, this guide also includes the Web Application Security Testing Methodology, which can be used to perform web application penetration tests.

The Web Application Security Testing Methodology is divided into a passive and an active phase. During the passive phase, the penetration tester explores the web applications from a user's point of view and tries to gain an understanding of the application's functionality and features. Throughout the active phase, the penetration tester performs the actual tests. For this purpose, the OWASP Web Security Testing Guide offers a comprehensive collection of test points, which are distributed across a total of twelve categories covering different areas of a web application [2]:

1) Information Gathering
2) Configuration and Deployment Management Testing
3) Identity Management Testing
4) Authentication Testing
5) Authorization Testing
6) Session Management Testing

7) Input Validation Testing
8) Testing for Error Handling
9) Testing for Weak Cryptography
10) Business Logic Testing
11) Client-side Testing
12) API Testing

C. Penetration Testing Skill Sets

The selection of a suitable approach by itself is no guarantee for a successful penetration test [17]. In order to maintain adaptability to the customer's needs and new types of technologies, standards should not be too restrictive [18][6]. While a standardized approach can give guidance to the penetration tester and point him in the right direction, at some point the tester may need to deviate from this predefined path. Thereafter, the testing is reliant on the abilities of the tester, which can not be covered by a standard.

Several of the previously mentioned standards describe the required skill sets to successfully perform penetration tests. In order to find vulnerabilities in a system, the penetration tester must understand how it works and how it can be abused. This can require extensive technical knowledge. Furthermore, performing penetration tests can have a negative impact on the customer's systems and networks. To prevent any damage, they should only be carried out by people with experience in IT security. According to BSI [6], penetration testers typically need the following hard skills:

- Knowledge of system administration/operating systems
- Knowledge of TCP/IP and, if applicable, other network protocols
- Knowledge of programming languages
- Knowledge of IT security products such as firewalls, intrusion detection systems
- Knowledge of how to handle hacker tools and vulnerability scanners
- Knowledge of applications/application systems

NIST [5] specifies similar technical know-how as a prerequisite, however, BSI also names 'creativity' as an essential soft skill. According to BSI [6], the creativity of the penetration tester is decisive for the success of the penetration test. Often, breaking into a system is only possible through creative combination of received information, discovered vulnerabilities, along with known tools and methodologies. OWASP [2] also claims that creativity allows for better results in finding vulnerabilities than fully automated tools. Creative penetration testers would therefore be expected to achieve better results than penetration testers who rely solely on the results of their tools [2][6].

According to OSSTMM [18], it is also important that a standardized approach does not interfere with the creativity of the penetration tester and thus negatively affects the quality of the outcome. However, OSSTMM [10] and BSI [6] also agree that creativity should not lead to unsystematic and untraceable penetration testing. Although intuition allows creativity to be applied to penetration testing, it can also lead to mistakes when a penetration tester relies solely on intuition by skipping checks that seem unnecessary [10].

Certificates usually serve as proof of a penetration tester's skills. There is a wide range of certification authorities that offer IT security and, in particular, penetration testing certificates. To obtain such a certificate, participants must pass an examination. Some of them are purely theoretical exams that solely test knowledge and thus only assess hard skills. However, others are more practical and require the successful completion of a penetration test as an exam and thereby also take into account soft skills [6][19].

III. Approach

A research project at Darmstadt University of Applied Sciences called fast electronic identification (SEIN) aims to provide an identification solution that enables fully automated identity verification via the account holder's online banking credentials [20]. To perform this type of identification, several web applications were implemented, which needed to be analyzed for their security through a penetration test.

We took this opportunity to commission two independent service providers to each conduct a penetration test of the SEIN web applications. To make the results of both penetration tests comparable, we made sure that the same conditions and terms applied to both contractors.

Since the SEIN research funds were not intended for cybersecurity research, but only to ensure that required standards such as ISO 27001 were met, the sample size was limited to these two service providers. Both service providers are local companies that have ties to the university through graduates and lecturers. One provider was initially contracted to assist with the implementation of an Information Security Management System (ISMS), and penetration testing was already included in their proposal. The other service provider offered a free initial penetration test as a promotional activity.

A. General conditions

Both contractors were given four days to perform the penetration test, plus an additional day to create the final report. In addition, both parties have been provided with the same technical documentation, including a list of the systems to be tested with short descriptions, and a sequence diagram to illustrate the identification process. SEIN assured that no changes have been made during the penetration tests, so that the same conditions applied to both penetration tests.

The two service providers stated that the penetration tests would be performed by certified professionals. Further, they claimed that their methodologies are based on the BSI Penetration Testing Model and the OWASP Web Security Testing Guide.

The penetration tests were performed sequentially to ensure that the penetration testers did not interfere with each other. On completion of both penetration tests, the findings of the two reports were reviewed and compared.

B. Investigated web applications

Four web applications of the SEIN research project were examined. These included a business portal, the business portal API, a demo application, and the API of the SEIN backend server.

- **Business Portal**
  The Business Portal is a Javascript-based single-page application that is connected to the Business Portal API. Companies can register via the Business Portal to receive an API key for the use of the SEIN Backend API.
- **Business Portal API**
  The Business Portal API is based on Strapi, a headless content management system. Strapi does not provide its own web frontend, but solely provides a REST API that allows content to be retrieved or edited. Using this API the essential functions of the business portal are made available.
- **Demo Application**
  The demo application simulates a webshop where an identity check of customers is performed via SEIN as part of the ordering process. In the web store shopping cart view, users are asked to enter their personal data. The identification process then starts. For this purpose, the demo application communicates as a client with the SEIN backend via the API provided. After the verification is completed, the results are displayed in the demo application.
- **SEIN Backend**
  The SEIN backend provides a REST API that can be used to confirm a person's identity. After a client, such as the demo application, has sent the data to be verified to the API, the user is asked to select a bank. In the next step, the user logs into the selected bank's online banking portal and grants SEIN access to the account holder's personal data for verification purposes. The SEIN backend then queries the required information and performs a comparison with the previously provided data. Finally, the result of this data comparison is sent back to the client.

IV. Results

The penetration testing reports of both service providers were reviewed and the included findings were extracted. A comparison of the aggregated results of both penetration tests can be seen in Table I. The two right-hand columns indicate whether the respective finding was listed in the corresponding report of penetration test A or B.

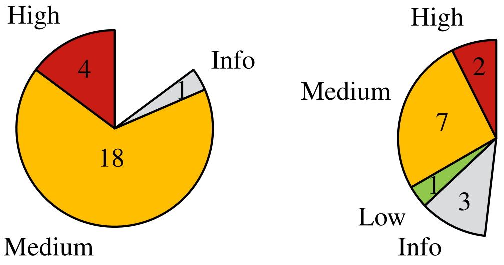
Due to the fact that the contractors used similar approaches, the results show some overlap. However, the direct comparison illustrates that one service provider was able to identify significantly more vulnerabilities, especially more with high or medium criticality. Most of these are among the OWASP Top 10, a collection of the ten most common and critical vulnerabilities in web applications, which is maintained by OWASP to create awareness for web application security. The

TABLE I
Comparison of the identified vulnerabilities

Vulnerability	Risk	P.T. A	P.T. B
Stored Cross-Site-Scripting	High	✓	✗
Error-handling enables denial of service	High	✓	✗
Plain text transmission of authentication data	High	✓	✓
Support for TLS 1.0 and TLS 1.1 and cryptographically weak cipher suites	High	✓	✓
Use of outdated software	Medium	✓	✓
Missing attributes in HTTP headers	Medium	✓	✓
SSH service allows login by password	Medium	✓	✗
Incomplete implementation of two-factor authentication	Medium	✓	✗
Publicly available API documentation	Medium	✓	✗
Meaningful error messages allow user enumeration	Medium	✓	✓
Internal services exposed	Medium	✓	✓
Bypass of the reverse proxy possible	Medium	✓	✗
Use of self-signed certificates	Medium	✓	✗
Missing access control	Medium	✓	✗
Disclosure of software versions and components	Medium	✓	✓
Long-lived access tokens	Medium	✓	✓
No deactivation of access tokens after a user logout	Medium	✓	✗
Link to registration confirmation contains valid access token	Medium	✓	✗
Disclosure of internal error messages	Medium	✓	✓
Lack of rate limiting in the APIs	Medium	✓	✗
Sensitive data in URLs of the demo application and the back-end API	Medium	✓	✗
Cross-Origin Resource Sharing for any origin	Medium	✓	✗
SSH weak MAC algorithms	Low	✗	✓
JSON Web Tokens use a symmetric algorithm for the signature	Info	✓	✗
Web server delivers default files	Info	✗	✓
Responding to ICMP timestamp requests	Info	✗	✓
Responding to TCP timestamp requests	Info	✗	✓

document provides information about these vulnerabilities and references other documents, such as specific OWASP Cheat Sheets, that can assist in their investigation and remediation.

A closer look reveals that a large number of the vulnerabilities, that have been overlooked by one of the contractors, are actually covered by the OWASP Web Security Testing Guide [2]. A variety of these are authentication and authorization based vulnerabilities, which the OWASP Testing Guide addresses in detail in the categories 'Identity Management Testing', 'Authentication Testing' and 'Session Management Testing'. Furthermore, the overlooked high-risk vulnerabilities are covered by the chapters 'Input Validation Testing' and 'Testing for Error Handling' [2]. By fully applying the OWASP Testing Guide and including the referenced Cheat Sheets, these should have also been found by the second service provider. Therefore, it is essential that the checks described in the guide are carried out without exception. The impact that skipping or forgetting individual checks can have on the results of a penetration test can be seen in Figure 2.



(a) Penetration Test A (b) Penetration Test B

Figure 2. Overall vulnerabilities identified by the contractors

Further, the comparison also shows some vulnerabilities that are not addressed by the OWASP Testing Guide. This indicates that both penetration testers performed checks beyond the OWASP Testing Guide as part of their individual approach. This may be due to the fact that the OWASP Testing Guide focuses primarily on the web application itself. Although the chapter 'Configuration and Deployment Management Testing' also covers the configuration of the webserver used, other services that could run on the same system are not considered here. Yet these could also be potential entry points for an attacker, which is why they are also checked for obvious vulnerabilities by some penetration testers. In that case, the individual approach determines which checks are performed beyond the OWASP Testing Guide to analyze these additional services. Since the focus of a web application penetration test lies on checking web applications, it must be decided where the line is drawn to an external IT infrastructure penetration test.

It appears that service provider B invested more effort into performing these additional checks. This enabled them to uncover a few minor misconfigurations, although they do not add much value for the customer. This work might have been better spent on processing the checks of the OWASP Testing Guide.

V. Conclusion

Overlooking vulnerabilities has a direct impact on the quality of the penetration test and thus on the client's security. To prevent this, penetration testers usually rely on standards that define which areas a penetration test should cover. However, a standard is no guarantee for a successful penetration test. As our studies have shown, the results of two penetration tests conducted in the same environment under identical conditions can still differ significantly despite the use of established standards. The decisive factor here was the human component, precisely the penetration testers themselves, as one of them was able to find considerably more vulnerabilities and, above all, more valuable ones in terms of risk.

As both were certified penetration testers, it is safe to assume they have similar hard skills. However, it appears the decisive factor was how they dealt with their creativity and intuition. While it may enable penetration testers to archive better results, it may also cause problems when they solely rely on intuition and do not stick to the chosen approach. It is important that all checks of this approach are performed without exceptions.

Furthermore, it could be observed that the penetration testers or their companies can add their own touch by using individual approaches that are an extension of established standards. This allows them to add value to the customer by performing additional checks on top of the predefined ones. Still, it is important not to lose focus on the actual objectives of the penetration test.

Future research could further investigate the interaction between hard skills and soft skills of penetration testers and their impact on penetration tests results. A larger sample size could provide insight into how often major discrepancies between penetration tests occur. This could also indicate whether it makes sense to always have penetration tests performed by several independent service providers in order to achieve better coverage. In addition, individual penetration testing approaches seem to be widely used but little researched. Further research could compare individual approaches with standardized ones in terms of their effectiveness. It could also be investigated whether combinations of standards are useful and which combinations work well together.

References

[1] Bundeskriminalamt [in English: Federal Criminal Police Office], "Bundeslagebild Cybercrime 2020," Jul. 2021. [Online]. Available: https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Cybercrime/cybercrimeBundeslagebild2020.pdf?__blob=publicationFile&v=4 [retrieved: Aug, 2021].

[2] E. Saad and R. Mitchell, *OWASP Web Security Testing Guide Version 4.2*. OWASP Foundation, Dec. 2020.

[3] S. Alavi, N. Bessler, and M. Massoth, "A Comparative Evaluation of Automated Vulnerability Scans versus Manual Penetration Tests on False-negative Errors," in *CYBER 2018: The Third International Conference on Cyber-Technologies and Cyber-Systems*, 2018, pp. 1–6.

[4] P. Engebretson, *The basics of hacking and penetration testing: ethical hacking and penetration testing made easy*, 2nd ed. Amsterdam ; Boston: Syngress, an imprint of Elsevier, 2013.

[5] K. Scarfone, M. Souppaya, A. Cody, and A. Orebaugh, *Technical guide to information security testing and assessment*. National Institute of Standards & Technology, 2008.

[6] Federal Office for Information Security, “A Penetration Testing Model,” 2003.

[7] J. Andress, *Foundations of information security: a straightforward introduction*, 1st ed. San Francisco: No Starch Press, 2019.

[8] H. C. A. v. Tilborg and S. Jajodia, Eds., *Encyclopedia of cryptography and security*, 2nd ed., ser. Springer reference. New York: Springer, 2011.

[9] K. M. Henry, *Penetration testing protecting networks and systems*. Ely, U.K: IT Governance Pub, 2012.

[10] P. Herzog, “OSSTMM 3-The Open Source Security Testing Methodology Manual: Contemporary Security Testing and Analysis,” *ISECOM-Institute for Security and Open Methodologies*, 2010.

[11] C. Nickerson *et al.*, “The Penetration Testing Execution Standard,” 2014. [Online]. Available: http://www.pentest-standard.org/index.php/Main_Page [retrieved: Aug, 2021].

[12] R. Baloch, *Ethical hacking and penetration testing guide*. Boca Raton: CRC Press, Taylor & Francis Group, 2015.

[13] T. P. Chiem, “A study of penetration testing tools and approaches,” PhD Thesis, Auckland University of Technology, 2014.

[14] K. Cardwell, *Building Virtual Pentesting Labs for Advanced Penetration Testing*, 2nd ed. Birmingham: Packt Publishing, Limited, 2016, oCLC: 963293305.

[15] A. Shanley and M. Johnstone, “Selection of penetration testing methodologies: A comparison and evaluation,” *13th Australian Information Security Management Conference*, pp. 65–72, 2015.

[16] OWASP Foundation, “OWASP - Main Page,” 2021. [Online]. Available: https://www.owasp.org/index.php/Main_Page [retrieved: Aug, 2021].

[17] E. Rey, M. Thumann, and D. Baier, *Mehr IT-Sicherheit durch Pen-Tests*. Wiesbaden: Vieweg+Teubner Verlag, 2005.

[18] P. Herzog, “OSSTMM 2.2–Open Source Security Testing Methodology Manual,” *ISECOM-Institute for Security and Open Methodologies*, 2006.

[19] D. Bhattacharyya, “Penetration Testing for Hire,” *International Journal of Advanced Science and Technology*, vol. 8, pp. 1–8, 2009.

[20] M. Massoth and S. L. Ahier, “Fast Electronic Identification at Trust Substantial Level using the Personal Online Bank Account,” in *CYBER 2020: The Fifth International Conference on Cyber-Technologies and Cyber-Systems*, 2020, pp. 94–99.