Two methodologies for physical penetration testing using social engineering

Trajce Dimkov, Wolter Pieters, Pieter Hartel
Distributed and Embedded Security Group
University of Twente, The Netherlands
{trajce.dimkov, wolter.pieters, pieter.hartel}@utwente.nl

ABSTRACT

Penetration tests on IT systems are sometimes coupled with physical penetration tests and social engineering. In physical penetration tests where social engineering is allowed, the penetration tester directly interacts with the employees. These interactions are usually based on deception and if not done properly can upset the employees, violate their privacy or damage their trust toward the organization and might lead to law suits and loss of productivity. We propose two methodologies for performing a physical penetration test where the goal is to gain an asset using social engineering. These methodologies aim to reduce the impact of the penetration test on the employees. The methodologies have been validated by a set of penetration tests performed over a period of two years.

Keywords: penetration testing, physical security, methodology, social engineering, research ethics

1. INTRODUCTION

A penetration test can assess both the IT security and the security of the facility where the IT systems are located. If the penetration tester assesses the IT security, the goal is to obtain or modify marked data located deep in the organizations network. Similarly, in testing the physical security of the location where the IT system is located, the goal of the penetration test is to obtain a specific asset, such as a laptop or a document. Physical and digital penetration tests can be complemented with social engineering techniques, where the tester is allowed to use knowledge and help from the employees to mount the attack.

In digital penetration tests the resilience of an employee is measured indirectly, by making phone queries or sending fake mail that lure the employee to disclose secret information. These tests can be designed in an ethical manner [1] and within the legal boundaries [2]. However, measuring the resilience of an employee against social engineering in a physical penetration test is *direct* and *personal*. When the tester enters the facility of the organization and directly interacts with the employees, she either deceives the employee, trying to obtain more information about the goal, or urges the employee to help her, by letting the tester inside a secure area or giving the tester a credential. The absence of any digital medium in the communication with the employees makes the interaction between the penetration tester and the employee intense, especially if the employee is asked to break company policies.

There are three main consequences from personal interaction between the tester and the employee. First, the employee might be stressed by having to choose between helping a colleague and breaking the company policies. Second, the tester might not treat the employee respectfully. Finally, when helping the penetration tester to enter a secure location, the employee loses the trust from the people who reside in the secure location. For example, employees might stop trusting the secretary when they find out she let an intruder into their office. To avoid ethical and legal implications, organizations may avoid physical penetration testing with social engineering, leaving themselves unaware of attacks where the attacker uses non-digital means to attack the system.

This paper tackles the problem how to perform a physical penetration test using social engineering in the most respectful manner, while still getting results that lead to improving the security of the organization. The contribution of this paper is two methodologies for physical penetration tests using social engineering where the goal is to gain possession of a physical asset from the premises of the organization. Both methodologies are designed to reduce the impact of the test on the employees. The methodologies have been validated by performing 14 live penetration tests over the last two years, where students tried to gain possession of marked laptops placed in buildings of two universities in The Netherlands.

The rest of the paper is structured as follows. In section 2 we present related work and in section 3 we set the requirements for the methodologies. Sections 4 and 5 outline the methodologies, section 6 provides an evaluation of the structure of the methodologies and section 7 concludes the paper.

This research is supported by the Sentinels program of the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs under projects number TIT.7628.

ACSAC '10 Dec. 6-10, 2010, Austin, Texas USA

2. RELATED WORK

In the computer science literature, there are isolated reports of physical penetration tests using social engineering [3, 4]. However, these approaches focus completely on the actions of the penetration tester and do not consider the impact of the test on the employees.

There are a few methodologies for penetration testing. The Open-Source Security Testing Methodology Manual (OSSTMM) [5] provides an extensive list of *what* needs to be checked during a physical penetration test. However, the methodology does not state *how* the testing should be carried out. OSSTMM also does not consider direct interaction between the penetration tester and the employees. Barret [6] provides an audit-based methodology for social engineering using direct interaction between the penetration tester and an employee. Since this is an audit-based methodology, the goal is to test *all* employees. Our methodologies are goal-based and focus on the security of a specific physical asset. Employees are considered as an additional mechanism which can be circumvented to achieve the goal, instead of being the goal. Türpe and Eichler [7] focus on safety precautions while testing production systems. Since a test can harm the production system, it can cause unforseeable damages to the organization. In our work the penetration test of the premises of an organization can be seen as a test of a production system.

In the crime science community, Cornish [8] provides mechanisms how to structure the prosecution of a crime into universal crime scripts and reasons about mechanisms how to prevent the crime. We adopt a similar reporting format to present the results from a penetration test. However, instead of using the crime script to structure multiple attacks, we use the script to identify security mechanisms that continuously fail or succeed in stopping an attack.

In social science research, the Bellman report [9] defines the ethical guidelines for the protection of humans in testing. The first guideline in the report states that all participants should be treated with respect during the test. Finn [10] provides four justifications that need to be satisfied to use deception in research. We use the same justifications to show that our methodology is ethically sound.

3. REQUIREMENTS

A penetration test should satisfy five requirements to be useful for the organization. First, the penetration test needs to be realistic, since it simulates an attack performed by a real adversary. Second, during the test all employees need to be treated with respect [9]. The employees should not be stressed, feel uncomfortable nor be at risk during the penetration test, because they might get disappointed with the organization, become disgruntled or even start legal action. Finally, the penetration test should be repeatable, reliable and reportable [6]. We call these the R* requirements:

Realistic - employees should act normally, as they would in everyday life.

Respectful - the test is done ethically, by respecting the employees and the mutual trust between employees.

Reliable - the penetration test does not cause productivity loss of employees.

Repeatable - the same test can be performed several times and if the environment does not change, the results should be the same.

Reportable - all actions during the test should be logged and the outcome of the test should be in a form that permits a meaningful and actionable documentation of findings and recommendations.

These are conflicting requirements. For example:

1. In a realistic penetration test, it might be necessary to deceive an employee, which is not respectful.

2. In a realistic test, arbitrary employees might be social engineered to achieve the goal, which is unreliable.

3. In a reportable test, all actions of the penetration tester need to be logged, which is unrealistic.

Orchestrating a penetration test is striking the best balance between the conflicting requirements. If the balance is not achieved, the test might either not fully assess the security of the organization or might harm the employees.

We propose two methodologies for conducting a penetration test using social engineering. Both methodologies strike a different balance between the R* requirements, and their usage is for different scenarios. Both methodologies assess the security of an organization by testing how difficult it is to gain possession of a pre-defined asset.

The methodologies can be used to assess the security of the organization, by revealing two types of security weaknesses: errors in implementation of procedural and physical policies by employees and lack of defined security policies from the management. In the first case, the tests should focus on how well the employees follow the security policies of the organization and how effective the existing physical security controls are. In the second case, the primary goal of the tests is to find and exploit gaps in the existing policies rather than in their implementation. For example, a test can focus on how well the credential sharing policy is enforced by employees or can focus on exploiting the absence of a credential sharing policy to obtain the target asset.

In this paper we present the two methodologies which reduce the impact of these tests. The environment-focused (EF) methodology, measures the security of the environment where the asset is located. The methodology is suitable for tests where the custodian (person who controls the asset) is not subject of social engineering and is aware of the execution of the test. One example of such test is evaluating the security of the assets residing in the office of the CEO, but not the awareness of the CEO herself. The custodian-focused (CF) methodology is more general, and includes the asset owner in the scope of the test. In this methodology, the owner is not aware of the test. The CF methodology is more realistic, but it is less reliable and respectful to the employees.

4. ENVIRONMENT-FOCUSED METHOD

First, we define the actors in the environment-focused methodology. Then, we introduce all events that take place during the setup, execution and aftermath of the penetration test. Finally, we validate the methodology by conducting three penetration tests and present some insights from the experience.

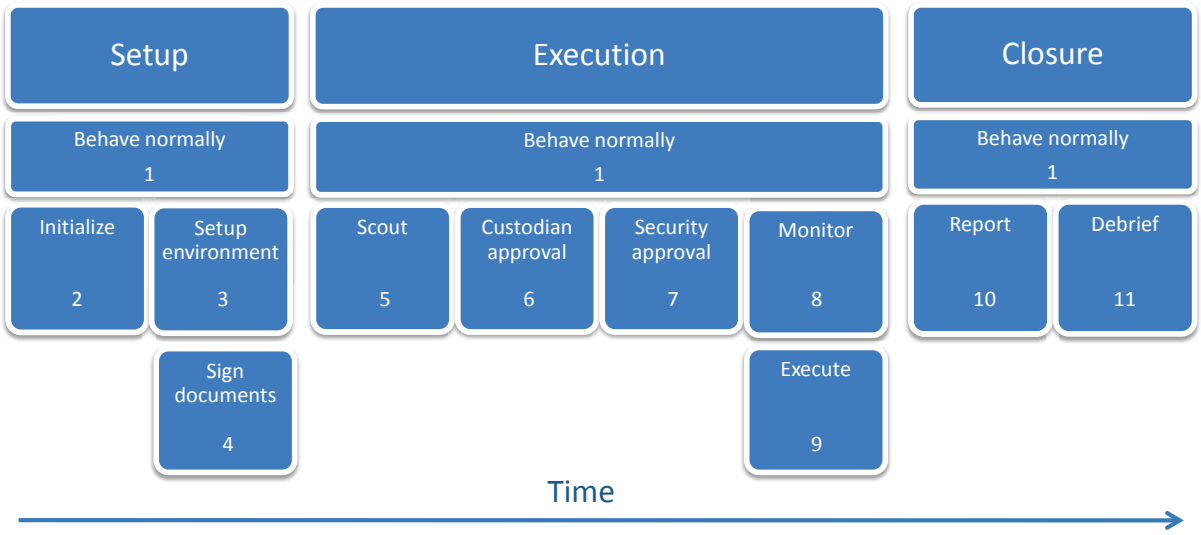

Figure 2: Sequence of events in the environment-focused methodology. Each box represents an event which happens in sequence or parallel with other events. For example, event 3 happens after event 2 and in parallel with events 1 and 4.

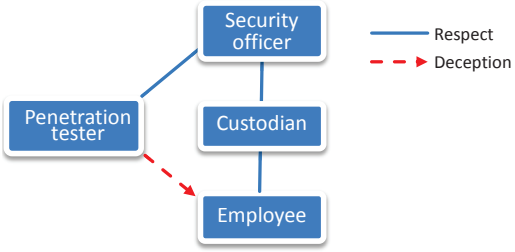

Figure 1: Actors in the EF methodology

4.1 Actors

The penetration test involves four different actors.

Security officer - an employee responsible for the security of the organization. The security officer orchestrates the penetration test.

Custodian - an employee in possession of the assets, sets up and monitors the penetration test.

Penetration tester - an employee or a contractor trying to gain possession of the asset without being caught.

Employee - person in the organization who has none of the roles above.

The actors and the relations between them are shown in Figure 1. The majority of actors treat each other with respect. No respect relation between two actors means either the actors do not interact during the penetration test (for example between the tester and the custodian) or do not have a working relationship (between the penetration tester and the employee). In this methodology, the tester deceives the employee during the penetration test, presented in the figure with a dashed line.

4.2 Setup

Figure 2 provides the sequence of events that take place during the setup, execution and closure of the penetration test. During all three stages of the penetration test, employees should behave normally (1 in Figure 2).

As in other penetration testing methodologies, before the start of the test, the security officer sets the scope, the rules of engagement and the goal (2 in Figure 2). The *goal* is gaining physical possession of a marked asset. The *scope* of the testing provides the penetration tester with a set of locations she is allowed to enter, as well as business processes in the organization she can abuse, such as processes for issuing a new password, or processes for adding/removing an employee. The *rules of engagement* restrict the penetration tester to the tools and means she is allowed to use to reach the target. These rules, for example, define if the tester is allowed to force doors, to break windows or to use social engineering.

The custodian first signs an informed consent form and then sets up the environment, by marking an asset in her possession and installing monitoring equipment.

The asset should not be critical for the daily tasks of the custodian or anyone else, including the organization. Thus, when the penetration tester gains possession of the asset, the productivity of the custodian using the asset and the process flow of the company will not be affected. The custodian leaves the asset in her office or an area without people (storage area, closet). If the custodian shares an office with other employees, the monitoring equipment should be positioned in such a way that it records only the asset and not the nearby employees. The custodian knows when the test takes place, and has sufficient time to remove/obscure all sensitive and private assets in her room and around the marked asset (3 in Figure 2).

Meanwhile, the penetration tester needs to sign the rules of engagement (4 in Figure 2). The OSSTMM methodology [5] provides a comprehensive list of rules of engagement.

4.3 Execution

The security officer should choose a trustworthy penetration tester and monitor her actions during the execution stage.

Generic Script	Attack trace	Circumvented mechanisms	Recommendations
Prepare for the attack	Buy a bolt cutter and hide it in a bag. Scout the building and the office during working hours. Obtain an after working hours access card.	Access control of the building entrances during working hours. Credential sharing policy.	Keep entrance doors to the building locked at all time. Provide an awareness training concerning credential sharing.
Enter the building	Enter the building at 7:30 AM, before working hours. Hide the face from CCTV at the entrance using a hat.	CCTV pre-theft surveillance.	Increase the awareness of the security guards during non-working hours.
Enter the office	Wait for the cleaning lady. Pretend you are an employee who forgot the office key and ask the cleaning lady to open the office for you.	Challenge unknown people to provide ID. Credential sharing policy.	Reward employees for discovering intruders.
Identify and get the asset	Search for the specific laptop. Get the bolt cutter from the bag and cut the Kensington lock. Put the laptop and the bolt cutter in the bag.	Kensington lock.	Get stronger Kensington locks. Use alternative mechanism for protecting the laptop.
Leave the building with the laptop	Leave the building at 8:00, when external doors automatically unlock for employees.	CCTV surveillance. Access control of the building entrances during working hours.	The motion detection of the CCTV cameras needs to be more sensitive .

Figure 3: Reporting a successful attempt. The figure shows an example of a generic script instantiated with an attack trace. First we define the generic script, which encompasses the stages of all attacks. In the example, they are: enter the building, enter the office, identify and get the asset, and exit the building. For each step in a trace, we identify both the mechanisms (if any) that were circumvented and mechanisms that stopped an attack. For failed attacks, the table shows which mechanisms were circumvented up to the failed action, and the mechanism that successfully stopped the attempt.

1. Social engineer night pass from an employee.
2. Enter the building early in the morning.
3. Social engineer the cleaning lady to access the office.
4. Cut any protection on the laptop using a bolt cutter.
5. Leave the building during office hours.

Figure 4: Example of an attack scenario

When the penetration test starts, the tester first scouts the area and proposes a set of attack scenarios (5 in Figure 2). An example of an attack scenario is presented in Figure 4. The proposed attack scenarios need to be approved first by the custodian (6 in Figure 2) and then by the security officer (7 in Figure 2). The custodian is directly involved in the test and can correctly judge the effect of the scenario on her daily tasks and the tasks of her colleagues. The security officer needs to approve the scenarios because she is aware of the general security of the organization and can better predict the far-reaching consequences of the actions of the tester.

If the custodian or the security officer disapprove an attack scenario, they need to evaluate the scenario and estimate the success. The tester puts in the report that the scenario was proposed, the reasons why the scenario was turned down and the opinion of all three roles on the success of the scenario. In this way the scenario although not executed, it is documented including the judgment on the effectiveness of the attack by the security officer, the custodian and the tester.

After approval from the custodian and the security officer, the tester starts with the execution of the attack scenarios (8 in Figure 2). The custodian and the security officer remotely monitor the execution (9 in Figure 2) through CCTV and the monitoring equipment installed by the custodian.

The penetration tester needs to install wearable monitoring equipment to log her actions. The logs serve three purposes. First, they ensure that if an employee is treated with disrespect there is objective evidence. Second, the logs prove that the penetration tester has followed the attack scenarios, and finally, the logs provide information how the mechanisms were circumvented, helping the organization repeat the scenario if needed.

4.4 Closure

After the end of the test, the penetration tester prepares a report containing a list of attack traces. Each attack trace contains information of successful or unsuccessful attacks (10 in Figure 2). Based on the report, the security officer debriefs both the custodians and any deceived employees during the test (11 in Figure 2).

Reporting. The attack traces are structured in a report that emphasizes the weak and the strong security mechanisms encountered during the penetration test, structured following 25 techniques for situational crime prevention [11]. For different domains there are extensive lists of security mechanisms to enforce the 25 techniques (for example, [12]). The combination of the attack traces together with the situational crime prevention techniques gives an overview of the circumvented mechanisms [13] (Figure 3)

Debriefing the employees and the custodian. After finding they were deceived by the same organization they work for, the employees might get disappointed or disgruntled. At the end of the test the security officer fully debriefs the custodian and the employees. The debriefing should be done carefully, to maintain or restore the trust between custodian and the employees who helped the tester to gain the asset.

4.5 Validation

To test the usability of the physical penetration tests using social engineering on the employees, we executed a series of penetration tests following the EF methodology. These pilots allowed us to gain a clear, first-hand picture of each execution stage of the methodology, and draw observations from the experience.

To avoid bias in the execution of the tests, we did not perform the tests ourselves, but recruited three teams of students who were in their first year of master studies to steal three laptops from the custodian (the first author). We locked the laptops with Kensington locks and hid the keys in an office desk. To monitor the laptops, we installed motion detection web cameras which streamed live feeds to an Internet server. Since the custodian shares the office with four other colleagues, the cameras were positioned in such a way to preserve the privacy of the colleagues. We told the colleagues we are doing an experiment, but we did not reveal the nature nor the goal of the experiment.

Since we knew about the penetration test, we did not allow the students to gain possession of the laptops in our presence. During the experiment, we carried on the normal work, thus the students were forced to carry on the attacks after working hours or during the lunch break.

The three teams scouted the building and wrote a list of attack scenarios they want to execute. Eventually, all three teams successfully obtained the target laptop and wrote the successful and unsuccessful attempts in the format shown in Figure 3. After the penetration test, we individually debriefed the security officer, the security guard, the secretary and the colleagues.

4.6 Lessons learned from the penetration tests

The observations are result of our experience with the penetration tests using qualitative social research and might not generalize to other social environments. However, the observations provide an insight of the issues that arose while using the methodology in practice.

The attack scenarios should be flexible. Although the students provided scenarios prior to all attacks, in all cases they were forced to deviate from them, because the target employee was either not present or was not behaving as expected. Attack scenarios assure the custodian and the security officer that the actions of the penetration tester are in the scope of the test, but at the same time there should be some freedom in adapting the script to the circumstances.

The methodology does not respect the trust relationship between the custodian and the employees. After the penetration test, the custodian knows which employees were deceived, and the trust relationship between them is disturbed. For example, if the secretary lets the penetration tester into the office of the custodian, the custodian might not be able to trust her again.

During the penetration test, separating the custodian from the employees is hard. Whenever the students approached a colleague from the office, the first reaction of the colleague was to call the custodian and ask for guidance. This led to uncomfortable situations where we were forced to shut down our phones and ignore e-mails while outside the office.

Debriefing proved to be difficult. After the test, we fully disclosed the test to all involved employees. Debriefing the security guard who opened the office for the penetration testers three times was the hardest. During the debriefing we focused on the benefits of the penetration test to the university and their help setting up the test. After the debriefing, we concluded that we caused more stress to the guard during the debriefing than the students had caused during the penetration test.

5. CUSTODIAN-FOCUSED METHOD

In the EF methodology, the custodian is aware of the penetration test. The knowledge of the penetration test changes her normal behavior and thus influences the results of the test. Since the asset belongs to the custodian, and the asset is in the office of the custodian, in many environments it is desirable to include the custodian's resistance to social engineering as part of the test.

After performing the first series of penetration tests, we revisited and expanded the environment-focused methodology. The CF methodology can be seen as a refinement of the EF methodology, based on the experience from the first set of penetration tests. In the CF methodology the custodian is not aware of the test, making the methodology suitable for penetration tests where the goal is to check the overall security of an area *including* the level of security awareness of the custodian.

5.1 Actors

There are six actors in the CF methodology.

Security officer - an employee responsible for the security of the organization.

Coordinator - an employee or contractor responsible for the experiment and the behavior of the penetration tester. The coordinator orchestrates the whole penetration test.

Penetration tester - an employee or contractor who attempts to gain possession of the asset without being caught.

Contact person - an employee who provides logistic support in the organization and a person to be contacted in case of an emergency.

Custodian - an employee at whose office the asset resides. The custodian should not be aware of the penetration test (1 in Figure 5).

Employee - person in the organization who has none of the roles above. The employee should not be aware of the penetration test (2 in Figure 5).

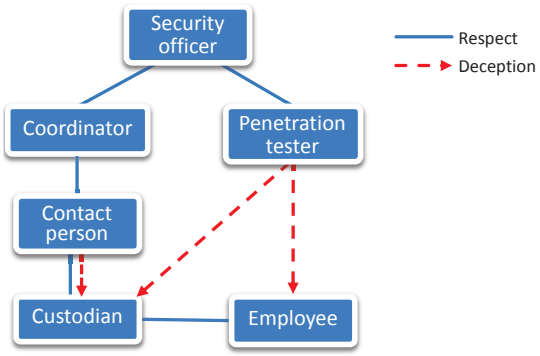

Figure 6: Actors in the CF methodology

Figure 6 shows the actors and the relations between them. In this methodology, the penetration tester deceives both, the employees and the custodian. Moreover, the contact person also needs to deceive the custodian. These relations

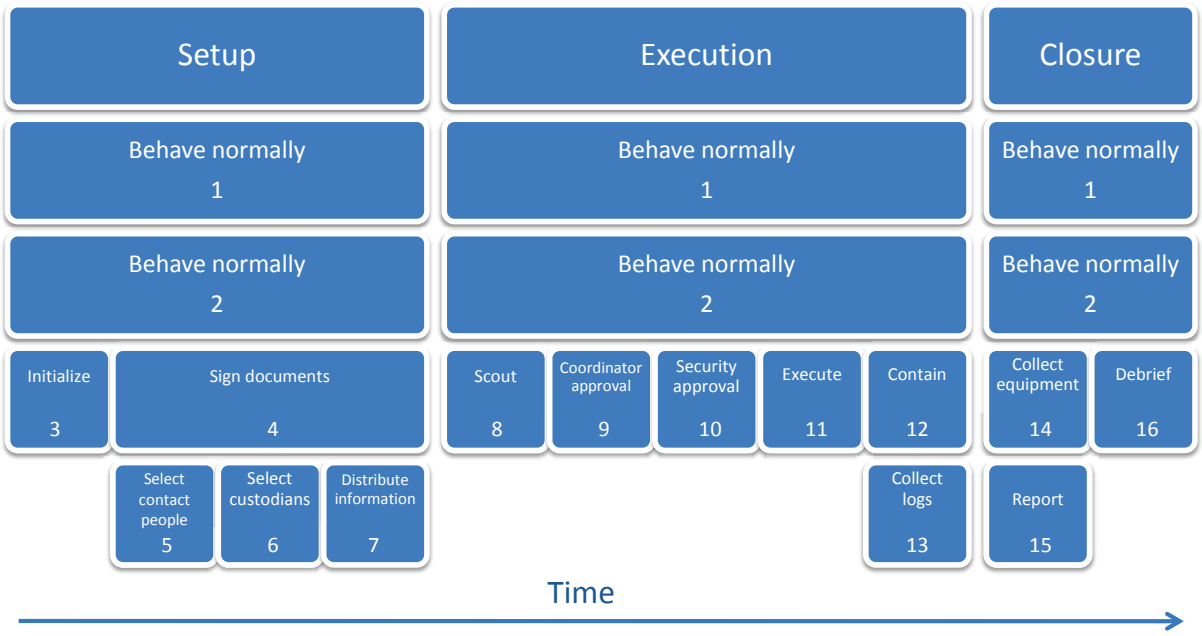

Figure 5: Sequence of events in the custodian-focused methodology

are discussed in greater depth in section 6.

5.2 Setup

At the beginning, similar to the EF methodology, the security officer initializes the test by defining the target, scope and the rules of engagement. The security officer at this point assigns a coordinator for the penetration test and provides the coordinator with marked assets and equipment for monitoring the assets (3 in Figure 5). The marked assets should be similar to the asset of interest for which the security is measured. The monitoring equipment should be non-intrusive and its purpose is to have additional information on the activities of the penetration tester.

The penetration tester should sign the rules of engagement (Appendix A) before the start of the execution stage (4 in Figure 5). The coordinator selects a number of contact people and provides them with the marked assets and the monitoring equipment (5 in Figure 5). Furthermore, the coordinator provides a cover story which explains why the custodian is given the asset. The contact person selects a number of custodians based on the requirements from the security officer (random, specific roles, specific characteristics) and distributes the marked assets and the monitoring equipment to the custodians. After giving the monitoring equipment, the contact person should get a signed informed consent (Appendix B) from the custodians (6 in Figure 5). If the asset can store data, the document must clearly state that the custodian should not store any sensitive nor private data in the asset. Before the penetration test starts, the coordinator distributes a list of penetration testers to the security officer, and a list of asset locations to the penetration tester (7 in Figure 5).

5.3 Execution

The first steps of the execution stage are similar to the previous methodology. The penetration tester scouts the area and proposes attack scenarios (8 in Figure 5). The coordinator and later the security officer should agree with these scenarios before the tester starts executing them (9 and 10 in Figure 5). After approval from both actors, the tester starts executing the attack scenarios. If a penetration tester is caught or a termination condition is reached, the penetration tester immediately informs the contact person. Thus, if the custodian stored sensitive data in the asset, the data is not exposed.

When the tester gains possession of the target asset, she informs the contact person and the coordinator and returns the asset to the contact person (11 in Figure 5). The contact person collects the monitoring equipment and informs the security officer (12 in Figure 5). If the tester gains possession of the asset without the knowledge of the custodian, the contact person needs to reach the custodian before the custodian reaches the office and explain to the custodian that the test is terminated. The security officer obtains surveillance videos from the CCTV and access logs and gives them to the coordinator (13 in Figure 5).

5.4 Closure

After the execution stage, the penetration tester writes a report of all attempts, both failed and successful, in the form of attack traces and gives them to the coordinator (14 in Figure 5). The coordinator has two tasks. First, she collects the marked assets and monitoring equipment from the contact person (15 in Figure 5) and returns them to the security officer. Second, the coordinator debriefs the security officer and the custodians and provides the custodian a form of reward for helping in the assessment (16 in Figure 5).

Not all employees that were social engineered should be debriefed. Employees who were treated with respect and to whom the penetration tester did not cause discomfort during the interaction should not be debriefed, because the debriefing can cause more stress than the interaction with

the penetration tester. The decision which employees need to be debriefed lies with the security officer, and is based on the logs from the penetration tester and the monitoring equipment. The criteria on which employees need to be debriefed are presented in greater detail in Section 6.

Debriefing the custodians. All custodians should be debriefed, because they sign an informed consent at the beginning of the test. However, to preserve the trust between the custodian and the employees, the custodian should not know which employee contributed to the attack.

Three elements should be considered before the debriefing. First, the custodians were deceived by the organization they work for (more specifically, by the contact person). Second, in case of direct interaction, their privacy might be violated by the logging equipment from the tester. Third, they might be stressed from the penetration test either directly, through interaction with the penetration tester, or indirectly, by finding their asset is gone before the contact person reaches them.

The debriefing should focus on the contribution of the custodian in finding the security vulnerabilities in the organization, and the custodian should be rewarded for the participation.

5.5 Validation

We orchestrated eleven penetration tests with the custodian-focused methodology. As targets we used eleven marked laptops that just came out on the market. We selected three PhD students as contact persons, who chose eleven friends as custodians (snowball sampling [14]). The custodians were ten PhD students and one professor spread across two universities, in nine different buildings.

We gave the laptops to the custodians for two weeks usage. The cover story was that the universities are interested in the usability of the new generation of laptops. Thus, we needed to monitor the behavior of the custodian and measure their satisfaction level. The contact persons explained that they cannot tell the custodians exactly which behavior we measure, since it might change the results of the experiment. We instructed the custodians to lock the laptops with a Kensington lock and to keep them in the office. To reduce the risk of data leakage and loss of productivity, we asked the custodians not to store any private or work data on the laptops. Later, thirty-three master students in computer security, divided in eleven teams of three, tried to gain possession of the laptops by using physical and social means.

Figure 7: Recording from the validation of the EF methodology. The student provided to the janitor a fake email stating he needs to collect a laptop from the custodian office. The janitor let the student into the office and helped him find the key from the Kensington lock.

Figure 8: Recording from the validation of the CF methodology. The student went to the office early in the morning, disguised as an employee who forgot his key. The cleaning lady let the student in. The student used a bolt cutter to remove the Kensington lock.

The students took roles as service desk employees, students that urgently needed a laptop for a few hours or coordinator representatives. The students used mobile phones and pocket video cameras to record the conversation with the employees. In one case they took a professional camera and a cameraman, and told the custodian the recording is part of a study to measure the service quality of the service desk.

The resistance of the employees varied. In five cases, the employees gave the laptop easily after being showed a fake email and being promised they will get the laptop back in a few hours. In two cases the custodian wanted a confirmation from a supervisor or the coordinator. In one case a colleague of the custodian got suspicious and sent an email to the campus security. Since only the main security officer knew about the penetration test, in few hours the security guards were all alerted and started searching for suspicious students.

However, in two cases the students were not able to social engineer the custodian directly and were forced to look for alternative approaches. For example, in one of the cases the students entered the building before working hours. At this time the cleaning lady cleans the offices, and under the assumption it is their office let the students inside. After entering the office, the students cut the Kensington lock and left the building before the custodian arrived.

We debriefed only the custodians through a group presentation, where we explained the penetration test and its goal.

5.6 Lessons learned from the validation

It should be specified in advance which information the penetration tester is allowed to use. For example, the penetration tester should not use knowledge about the cover story used by the contact person. During the validation, six penetration testers used knowledge of the cover story to convince the custodian to hand in the laptop. Thus, these tests were less realistic.

Panic situations need to be taken into consideration in the termination conditions. Several times the custodian or an employee got suspicious and raised an alarm. Since only the security officer knew about the experiment, and the other security personnel was excluded, news of people stealing laptops spread in a matter of hours. In these situations the coordinator should react quickly and explain to the employees that the suspicious activity is a test.

The penetration test cannot be repeated many times. If a custodian participated in the penetration test once, she knows what will happen. The same holds for the employees she told about the experiments and the employees that were socially engineered.

6. EVALUATION

In this section we compare both methodologies against the R* requirements. The satisfaction of the requirements is defined by the rules of engagement, which attack scenarios are approved for execution, and the structure of the methodologies. Less restrictive rules of engagement and approving more invasive attack scenarios make the penetration test more realistic, but make the test less reliable and respectful to the employees. The evaluation below assumes these two elements are tuned to the risk appetite of the organization and focuses only on the structure of the methodologies.

Reliable: In the EF methodology, the penetration tester gains possession of a non-critical asset which the custodian is prepared to lose. Thus, the result of the penetration test will not affect the productivity of the custodian. In the CF methodology, the productivity of the custodian may be affected, since the custodian does not know the asset will be stolen. The informed consent is a mechanism to avoid productivity loss, since it explicitly states not to use the marked asset for daily tasks nor store sensitive information on the asset. In both methodologies, the productivity of other employees is not affected, since the penetration tester does not gain possession of any of their belongings without their approval.

Repeatable: The repeatability of any penetration test using social engineering is questionable, since human behavior is unpredictable. Checking if a penetration test is repeatable would require a larger set of tests on a single participant, and a larger number of participants in the test.

Reportable: The approach used in reporting the results of the penetration test completely covers all information needed to perform the attack in a real-life situation and provides an overview of what should be improved to thwart such attempts. The logs from the tester and the monitoring equipment installed by the custodians provide detailed information on all actions taken by the penetration tester, giving a clear overview of how the mechanisms are circumvented.

	EF methodology	CF methodology
Reliable	+++	++
Repeatable	-	-
Reportable	+++	+++
Respectful: actors	++	+
Respectful: trust relations	-	++
Realistic	+	+++

Figure 9: Evaluation of both methodologies

Respectful: Both methodologies should respect all the employees and the trust relationships between them.

In physical penetration testing, the social engineering element is more intense than in digital penetration testing because the interaction between the penetration tester and the employee is direct, without using any digital medium. Baumrind [15] considers deception of subjects in testing as unethical. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, also clearly states this in their first rule of ethical principles: "Respect for persons" [9].

However, some tests cannot be executed without deception. Finn [10] defines four justifications that need to be met do make deception acceptable: (1) The assessment cannot be performed without the use of deception. (2) The knowledge obtained from the assessment has important value. (3) The test involves no more than minimal risk and does not violate the rights and the welfare of the individual. Minimal risk is defined as: "the probability and magnitude of physical or psychological harm that is normally encountered in the daily lives" [16]. (4) Where appropriate, the subjects are provided with relevant information about the assessment after participating in the test. Physical penetration testing using social engineering can never be completely respectful because it is based on deception. However, the deception in both methodologies presented in this paper is justifiable.

The first two justifications are general for penetration testing and its benefits, and have been discussed earlier in the literature (for example, Barrett [6]). The third justification states that the risk induced by the test should be no greater than the risks we face in daily lives. In the EF methodology, the only actor at risk is the employee. The penetration tester cannot physically harm the employee because of the rules of engagement, thus only psychological harm is possible. If the employees help the penetration tester voluntarily, the risk of psychological harm is minimal. The logging equipment assures the interaction can be audited in a case of dispute. In the CF methodology, an additional actor at risk is the custodian. The only case when the risk is above minimal for the custodian is if the tester gains possession of the asset without custodian's knowledge. When the custodian finds the asset missing, her stress level might increase. Therefore it is crucial for the contact person to reach the custodian before custodian learns about the theft.

The fourth justification states that all actors should be debriefed after the exercise. In both methodologies, all actors except the employees are either fully aware of the exercise, or have signed an informed consent and are debriefed after the exercise. Similarly to Finn and Jakobsson [1], we argue that there should be selective debriefing of the employees. Debriefing can make the employee upset and disgruntled and is the only event where the risk is higher then minimal. Thus, an employee should be debriefed only if the security officer constitutes the tester did more than minimal harm.

Besides being respectful toward all the participants, the methodology needs to maintain the trust relations between the employees. The EF methodology affects the trust between the custodian and the employees and the employees and the organization. This is a consequence of the decision to fully debrief all participants in the test. The CF methodology looks at reducing these impacts. First, the custodians are not told who contributed to the attack. Only the coordinator and the security officer have this information, and they are not related to the custodian. Second, the employees are not informed about the penetration test unless it deemed necessary. However, the trust between the custodian and the contact person is shaken. Therefore, the contact person and the custodian should not know each other prior to the test.

In conclusion, the CF methodology is less respectful to the custodian than the EF methodology, because the custodian is deceived and might get stressed when she finds out

the asset is gone. The EF methodology does not preserve any trust between the employees, the organization and the custodian. The CF methodology preserves the trust bond between the custodian and the employees and between the employees and the organization. However, the trust bond between the custodian and the contact person may be affected.

Realistic: The EF methodology allows testing the resilience to social engineering of employees in the organization. Since the custodian knows about the penetration test, she is not directly involved during the execution of the test, making this methodology implementable in limited number of situations. In the CF methodology, neither the custodian nor any of the other employees know about the penetration test, making the test realistic.

One might argue that if the asset is not critical for the employee, the tests are not realistic. On the other hand, taking away "real" assets in the penetration tests will clearly cause loss of production. In the EF methodology, this issue does not exist, as the employees who may be social-engineered are not aware of the importance of the target asset. Therefore, they have no reason to behave differently toward the experimental asset than to a "real" asset. However, in the CF methodology, the value of the asset as perceived by the custodian might influence the result of the tests, as the employee may be more likely to give the asset away if she knows it is not critical. As future work, we plan to investigate the effect of the perceived importance of the asset on the results of such tests.

7. CONCLUSION

Securing an organization requires penetration testing on the IT security, the physical security of the location where the IT systems are situated, as well as evaluating the security awareness of the employees who work with these systems. We presented two methodologies for penetration testing using social engineering. The custodian-focused methodology improves on the environment-focused methodology in many aspects. However, the environment-focused methodology is more reliable, does not deceive the custodian and fully debriefs all actors in the test. We provide criteria to help organizations decide which methodology is more appropriate for their environment. We evaluated both methodologies through analysis of their structure against a set of requirements and through qualitative research methods by performing a number of penetration tests ourselves. This paper shows that physical penetration tests using social engineering *can* reduce the impact on employees in the organization, and provide meaningful and useful information on the security posture of the organization.

In the future, we will focus on two topics. First, we want to investigate the effect of the perceived importance of the asset on the results of the test. We plan to separate the custodians in two groups and inform one of the groups that the laptop contains information critical for the organization. Second, we want to investigate the aspect of *safety* for both the employees and the testers. This research will help penetration testers perform tests in potentially hazardous environment, such as chemical or nuclear laboratories.

References

[1] P. Finn and M. Jakobsson. Designing ethical phishing experiments. *Technology and Society Magazine, IEEE*, 26(1):46–58, Spring 2007.

[2] C. Soghoian. Legal risks for phishing researchers. In *eCrime Researchers Summit, 2008*, pages 1–11. IEEE, 2008.

[3] C. Greenlees. An intruder's tale-[it security]. *Engineering & Technology*, 4(13):55–57, 2009.

[4] Wil Allsopp. *Unauthorised Access: Physical Penetration Testing For IT Security Teams*, chapter Planning your physical penetration test, pages 11–28. Wiley, 2009.

[5] P. Herzog. OSSTMM 2.2–Open Source Security Testing Methodology Manual. *Open source document, www.isecom.org/osstmm*, 2006.

[6] N. Barrett. Penetration testing and social engineering hacking the weakest link. *Information Security Technical Report*, 8(4):56–64, 2003.

[7] S. Türpe and J. Eichler. Testing production systems safely: Common precautions in penetration testing. In *Proceedings of Testing: Academic and Industrial Conference (TAIC PART 2009)*, pages 205–209. IEEE Computer Society, 2009.

[8] D. B. Cornish. The procedural analysis of offending and its relevance for situational prevention. In R. V. Clarke, editor, *Crime Prevention Studies*, volume 3, pages 151–196. Criminal Justice Press, Monsey, NY, 1994.

[9] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. pages 1–18, 1978.

[10] P.R. Finn. *Research Ethics: Cases and Materials*, chapter The ethics of deception in research, pages 87–118. Indiana University Press, 1995.

[11] D.B. Cornish and R.V. Clarke. Opportunities, precipitators and criminal decisions: A reply to Wortley's critique of situational crime prevention. *Crime Prevention Studies*, 16:41–96, 2003.

[12] G. Kitteringham. Lost laptops = lost data: Measuring costs, managing threats. Crisp report, ASIS International Foundation, 2008.

[13] R. Willison and M. Siponen. Overcoming the insider: reducing employee computer crime through situational crime prevention. *Communications of the ACM*, 52(9):133–137, 2009.

[14] B.L.A. Goodman. Snowball sampling. *The Annals of Mathematical Statistics*, 32(1):148–170, 1961.

[15] D. Baumrind. Research using intentional deception. Ethical issues revisited. *The American psychologist*, 40(2):165–174, 1985.

[16] Code of Federal Regulations. Title 45: Public welfare department of health and human services. part 46: Protection of human subjects. pages 1–12. 2005.

Appendix A:

Rules of engagement

I, _____ (name of student) agree to perform penetration tests for _____ (name of researcher)

I understand that the participation of is completely voluntary. At any time, I can stop my participation.
I fully oblige to the following rules of engagement:

1. I will only execute attacks that are pre-approved by the researcher and only to an assigned target.
2. I am not allowed to cause any physical damage to university property, except for Kensington locks.
3. I am not allowed to physically harm any person as part of the test.
4. I will video or audio record all my activities while interacting with people during the penetration test as a proof that no excessive stress or panic is caused to anyone.
5. If I am caught by a guard of a police officer, I will not show any physical resistance.

Signature of researcher:_____ Date: _____

Signature of student: _____ Date: _____

Appendix B:

Informed consent

I, _____ (name of employee) agree to participate in the study performed by _____ (name of the research group).
I understand that the participation of the study is completely voluntary. At any time, I can stop my participation and obtain the data gathered from the study, have it removed from the database or have it destroyed.
The following points have been explained to me:

1. The goal of this study is to gather information of laptop usage. Participation in this study will yield more information concerning the habits people have in using mobile devices.
2. I shall be asked to work for 5 min every day on a laptop for one month. The laptop will be monitored and recorded using a keynoter and a web-camera. At the end of the study, the researcher will explain the purpose of the study.
3. No stress or discomfort should result from participation in this study.
4. The data obtains from this study will be processed anonymously and can therefore not be made public in an individually identifiable manner.
5. The researcher will answer all further questions on this study, now or during the cause of the study.

Signature of researcher:_____Date:_____

Signature of employee:_____ Date:_____