A Structured Comparison of Social Engineering Intelligence Gathering Tools

Kristian Beckers[1], Daniel Schosser[1], Sebastian Pape[2], and Peter Schaab[1]

[1] Technische Universität München (TUM), Institute of Informatics
Boltzmannstr. 3, 85748 Garching, Germany
[2] Goethe University Frankfurt, Faculty of Economics and Business Administration
Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, Germany

Abstract. Social engineering is the clever manipulation of the human tendency to trust to acquire information assets. While technical security of most critical systems is high, the systems remain vulnerable to attacks from social engineers. Traditional penetration testing approaches often focus on vulnerabilities in network or software systems. Few approaches even consider the exploitation of humans via social engineering. While the amount of social engineering attacks and the damage they cause rise every year, the defences against social engineering do not evolve accordingly. However, tools exist for social engineering intelligence gathering, which means the gathering of information about possible victims that can be used in an attack. We survey these tools and present an overview of their capabilities. We concluded that attackers have a wide range of intelligence gathering tools at their disposal, which increases the likelihood of future attacks and allows even non-technical skilled users to apply these tools.

Keywords: social engineering; threat analysis; security awareness, security tools

1 Introduction

"The biggest threat to security of a company is not a computer virus, an unpatched hole in a key program or a badly installed firewall. In fact, the biggest threat could be you [...] What I found personally to be true was that it's easier to manipulate people rather than technology [...] Most of the time organizations overlook that human element". These words from Kevin Mitnick [7] were made over a decade ago and are still of utmost importance today.
As security technology improves the human user remains the weakest link in system security. It is widely accepted that the people of an organization are therefore both the main vulnerability of any organization's security as well as the most challenging aspect of system security [6, 27]. Chris Hadnagy [17] defines social engineering as "Any act that influences a person to take an action that may or may not be in their best interest". Numerous security consultants consider it a given for themselves as well as for genuine attackers to access critical information via social engineering [14, 43].
The harm of social engineering attacks has been discussed in various reports. In

2003 Gulati [15] reported that cyber attacks cost U.S. companies $266 million every year and that 80% of the attacks are a form of social engineering. Although not being very recent assessments of the situation, it seems that little has changed until today. A study of 2011 from Dimensional Research [9] shows that nearly half of the considered large companies and a third of small companies fell victim of 25 or more social engineering attacks in the two years before. The study further shows that costs per incident usually vary between $25 000 and over $100 000. Furthermore, surveys, like Verizon's *Data Breach Investigation Report* [41, 42] show the impact of social engineering. According to these studies the impact has grown from 7% of breaches in 2012 to 29% of breaches in 2013. These numbers should not be ignored and active support for mitigating these threats is needed. Even though companies are aware of the social engineering problem, they have little tools available to even assess the threat for themselves. Hiring penetration testing companies that *attack* their clients and show weaknesses in their defences is one available option. However, these tests have a number of inherent problems. Particularly, to address legal issues high effort has to be invested upfront [44]. In addition, the test outcome is closely related to the limited scope of the test. A tester may find that some employees are violating security policies. While this is an important finding that lets a company improve the education of their employees, the completeness of these kind of tests is an issue. Only few employees can be tested on only few occasions. Moreover, experiments indicate that this approach is difficult, due to humans' demotivation when confronted with these testing results [10].
A number of tools are available that enable intelligence gathering. On one side using these tools a social engineer can gather information that help him attack persons or organizations. On the other side, these tools provide an organization with an excellent alternative to pen testing or awareness trainings, as they allow to analyse possible vulnerabilities. However, a structured survey on the tools' capabilities is missing so far.
We believe to improve the current situation by conducting a structured survey of social engineering intelligence gathering tools and contribute the following:

- a classification of existing tools regarding categories such as proposed purpose, price, perceived usability, visualization of results etc.

- a survey of information types retrieved by the tools regarding information about company employees and their communication channels, as well as related information e.g. company policies;

- a discussion of how even simple attacker types can use these tools for sophisticated social engineering attacks

The remainder of our paper is organised as follows. Section 2 outlines the criteria for comparison, and Section 3 presents the results of our comparison. Section 4 concludes and provides directions for future research.

2 Social Engineering Basics and Tool Criteria

We acquire a basic understanding of social engineering and the general process attackers follow in Sect. 2.1. During the process various information is gathered about people, whom social engineers attack. Section 2.2 details our categorization of this *social engineering information* based on related work. Furthermore, we classify the tools on their *potential of applicability*, which describes the barriers that may or may not prevent an attacker from using them. For example, a tool that has a high price and poor usability will have little potential to be used by any attacker.

2.1 The Social Engineering Process

Various works report an underlying process to social engineering [21, 17, 27], which have recently been unified by Milosevic [26]. A social engineering attack consists of multiple phases as summarized in Table 1. In phase one the attacker conducts surveillance to identify a person within the inner circle of the targeted company. This person shall have access to the information the attacker desires. The next phase focuses on finding out as much about this person as possible. Every bit of information can help the attacker to manipulate the victim and her trust. During the pretexting phase the attacker starts building a relationship to the victim. Afterwards the attacker exploits the built up trust in the relationship and evaluates the gathered information in the post-exploitation phase.

Table 1: Overview of Social Engineering Phases by Milosevic [26]

Phase	Description
Pre-Engagement Interactions	Find targets with sufficient access to information/knowledge to perform an attack.
Intelligence Gathering	Gather information on each of the valid targets. Choose the ones to attack.
Pretexting	Use gathered information to build a relationship to the target. Gain victims' trust to access additional information.
Exploitation	Use the built up trust to get the desired information.
Post-Exploitation	Analyze the attack and the retrieved information. If necessary return to a previous phase to continue the chain of attack until the final information has been retrieved.

2.2 Social Engineering Information

This section focuses on types of information that can be gathered by a tool, in the following referred to as criteria. All criteria cover one or more essential

information for social engineering attackers. The more criteria a tool covers, the more interesting it is for a social engineer during information gathering.

Communication Channels. Communication channels are one of the most relevant information for a social engineer. This category will list which channels can be found by a certain tool. Possible channels are "Telephone Numbers", "Social Media Accounts", "E-Mails", "Instant Messengers", "Friends", "Personal Information" and "Private Locations" [23, 27].

User credentials. Some tools have access to databases which contain leaked user credentials. If a social engineer gets access to login information of a certain employee, it simplifies the conduction of an attack. Firstly, he can directly access a victim's accounts. Secondly, the attacker could pose as someone else, e.g. an administrator from the IT department, and by having access to the target's data convince his victim to act in a certain way [18, 27].

Locations. Some tools are especially designed to gather location data, while others provide them as a byproduct. Both, work addresses as well as an employee's private addresses can be useful for multiple purposes. Location data can be gathered from social media as it is embedded in photos and videos taken by cellphones. Also posts on social media can be tagged with a location. Other tools can convert IP addresses into physical locations and therefore find the physical locations of technical equipment [35, 18].

Job Positions. By retrieving the job position of an employee the social engineer can figure out what kind of information someone has access to. Based on job title, the attacker can draw conclusions about whether an employee is new to a company, what the hierarchy within the company looks like and much more. Based on the organization's structure, it is possible to use techniques such as name-dropping, using the name of someone higher in the company's hierarchy, to pressure the target into revealing information [18, 27].

Company Lingo. One of the easiest ways to convince someone of being authorized to access some information is by knowing the correct lingo [27]. Lingo means the words and abbreviations employees use within a company. Although this information is of great importance, it is very challenging to get access to. Knowledge about the lingo can be obtained by getting access to company manuals, internal reports or talking to employees.

Personal Information. The more personal information an attacker has on his target, the easier it is to find the correct angle and pressure points. One example would be well-defined spear-phishing e-mails using a person's interests. In case the e-mail contains enough private information to make it believable, the target is far more likely to open an attachment [35, 19].

2.3 Potential for Applicability

This section presents the evaluation criteria to generally classify the software.

Proposed Purpose. Some of the tools are actually designed to gather information on a person or company in the context of social engineering. However, a

user can also use tools for attacks which were designed for something completely different than social engineering.
Price. While some tools are free, others can be quite expensive and therefore might not be applicable for a quick self assessment. In some cases the tool itself is free, but for some features the user needs to have an API key that can be costly. This criteria focuses on the prices of each tool and its limitations coming with different price tiers.
Usability. Based on the user interface and the amount of documentation provided, this category assesses the ease of usage. The underlying question is if the usability of a tool allows a company to perform its own threat assessment.
Input Parameters. Some tools have a broad range of possible search arguments, but most tools need specific information to initiate a search. Depending on which specific piece of information is required by the tool, this might limit the social engineer in the decision what tools to use.
Visualize Output. Some tools print all information into tables while others have better ways of visualizing gathered information. For example location data can be illustrated by marking the positions on a map, instead of only providing GPS coordinates.
Ranking of results. As the amount of gathered information grows, the more valuable an adequate selection and sorting becomes. Therefore, filtering irrelevant information is helpful in focusing on more promising targets/information. We did not find significant support for filtering in the analysed tools and therefore do not list this criteria in the following.
Suggesting Counter-Measures. Most of the tools are only designed to gather information and do not inform how to protect this information. While this is not relevant for social engineers, it is highly relevant for those who want to protect themselves against attackers and against information gathering in general. Note that none of the tools suggest countermeasures, therefore we did not list the category in the following.

3 Comparing Social Engineering Tools and Webpages

In the following section, we introduce and analyze relevant tools and webpages. In a second step we provide an overview over the types of information that can be gathered by them.

3.1 Social Engineering Tools and Webpages

We compiled the following list of social engineering tools by using the following words "social engineering and tool or application or script or webpage" in a google[3] search and the list published by Hadnagy [17]. Three security researchers

[3] https://www.google.de

analysed the results independently and we included all tools and webpages that they agreed on having the potential to help a social engineering attacker conduct the process outlined in Sect. 2.1. We identified the following tools and webpages that met our criteria.

Maltego (Kali Linux Edition, Version 3.6.1) Maltego [32] is an intelligence and forensics application. Before starting a search, the user can choose between different machines. Every machine has its own purpose and is designed for a particular attack vector. Maltego offers 12 default machines within the software such as: *Company Stalker* This machine tries to get all e-mail addresses at a domain to resolve them on social networks. It also gets documents and extracts meta data. As an input, it needs a company's domain. *Find Wikipedia Edits* This machine takes a domain and looks for possible Wikipedia edits. *Footprint L1* This module performs a level 1 (fast, basic) footprint of a domain. *Person - E-Mail Address* This machine tries to obtain someone's e-mail address and checks where it's used on the internet.

Maltego combines multiple modules to gather information from various sources and represents them in an easy to understand way in form of a bubble diagram. The user can start of with a domain name, a username, an IP address or the name of a person depending on which module he wishes to use. The gained information can be used for further research e.g. as input for other modules.

Recon-ng (Version 4.8.0) Recon-ng [40] is a full-featured Web Reconnaissance framework. It is based on a large list of modules which can be used to gather information about a specific target. The modules range from host information to social media. The user is free to chain these modules after each other and by starting with a single domain name, the database can be filled with employee names, their e-mail addresses, usernames, passwords and geolocations of all involved servers. The final reports can be exported in json, csv, xml, html or as a pdf. Similar to the Social Engineering Toolkit and Metasploit its user interface is console based.

Cree.py (Version 1.4) Cree.py [20] is a geolocation Open Source Intelligence (OSINT) tool. It is designed to gather geolocation related information from online sources like social networks. This information can be filtered by location or date and is presented on a map. Therefore, Cree.py is useful to follow the trace of where a person has been over the time of using certain social media platforms. Examples would be Instagram, Twitter or Tumblr which gather location data on where photos or posts have been created. These information can be displayed on a map and recreate a trace of places where a person has been.

Spokeo Spokeo [38] is a search engine for people in the United States of America. There exist equivalent versions for other countries e.g. Pipl.com and PeekYou.com index people from all over the world. By entering the name, e-mail address, phone number, address or username of a person all related people matching the provided criteria are reported back. Depending on the wanted detail of the provided report, the price varies.

Social Engineering Toolkit (SET) SET [16] does not focus on finding information about a person. SET rather uses information on persons to e.g. send them phishing e-mails or gather information about company networks. The SET allows integration with other tools such as Metasploit that contain various scripts for vulnerability testing.

The Wayback Machine The Wayback Machine [39] is an archive of the internet. The vendor claims to provide the history of more then 427 billion web pages (as of July 2015). The platform creates snapshots of websites and allows a user to go back to older versions of a website that have been replaced by newer ones.

theHarvester (Version 2.7) The Harvester [12] is designed to gather e-mail addresses, subdomains, hosts, and open ports from public sources. These sources contain search engines, PGP key servers and the SHODAN [36] computer database for internet-connected devices.

Whitepages The *Whitepages* [5] website supports persons in finding people, their addresses and telephone numbers, private and from work. The service focuses on the U.S. and also provides reverse phone searches and similar means to identify a person based on technical information such as a phone number.

Background Checks The *freebackgroundcheck.com* [1] website provides information about people that has been collected by background checks on them for e.g. a telecommunication provider. The intention is that people can get informed what information is available about them and most likely checked in situations such as job interviews. The website *Instant Checkmate* [2] on the other hand focuses on providing information to the public about people's arrest records and criminal behaviour.

Tax Records Especially in the United States it is very easy to gain access to government information, as most data is publicly available [30]. Every person interested in the data can get access to arrest records, tax records and more for a small monetary fee per request. In addition, Ratsit in Sweden [34], Veroposi in Finland [4], Skatterlister in Norway [3] and recently the Federal Board of Revenue in Pakistan [31] also publish tax records online.

Company Related Information As social engineers thrive to know as much about the social surroundings of a target as possible, there are a lot of tools, that help gathering social related information about a target. Websites like *KnowEm* [22] and *Namechk* [29] allow to search on more then 600 social media networks, if a username is already allocated or still available. While this is not the primary purpose of the website, an attacker can use this to track down social media networks, which a target is using. *SocialMention* [37] is a platform, that searches for user-generated content like posts, blogs, videos, etc. from a specific user. By gathering this kind of information the attacker learns a lot about the target and his behavior.

In most cases a social engineer is not after private information about a target, but work related information. This is due to an attacker generally trying to get access

to work related sensitive information. Websites such as *Monster* [28], *LinkedIn* [24] and *Xing* [45] are good sources for collecting CVs and current job positions of people related to the target. In addition platforms like *careerbuilder* [8] and *glassdoor* [13] provide information about open job offers and expected earnings. *Hoovers* [11], *MarketVisual* [25] and *LittleSis* [33] are useful to gain knowledge about the social networks of employees. Especially for larger companies, these websites offer information about who is connected to whom.

3.2 Analyzing the Social Engineering Attack Potential

After having established each tool's characteristics, it is important to know, what tool is able to retrieve which kind of information. Some tools are able to collect more information than others and some information can only be found with a specific tool. Table 2 provides an overview of the tools survey. Furthermore, Table 3 provides a refinement of the previous table considering the potential for applicability categories introduced in Sect. 2.3 for selected tools and webpages. For space reasons we do not show the information for all tools and websites.

Table 2: Social Engineering Tool Comparison

	SET	Maltego	Recon-ng	Cree.py	Spokeo	Wayback Machine	theHarvester	knowem.com	Whitepages	Instant Checkmate	freebackgroundcheck.org
Search by Person/ Company	o	++	++	++	++	++	++	+	++	++	++
Retrieve E-Mail Address	o	++	++	o	o	o	++	o	o	o	o
Retrieve Username/ Password	o	o	++	o	o	o	o	o	o	o	o
Retrieve Job-Title	o	o	++	o	o	o	o	o	o	++	++
Retrieve Locations	o	+	+	++	+	o	o	o	++	++	++
Retrieve Personal Data	o	o	o	o	++	o	o	+	+	++	++
Usability	+	+	+	++	++	++	+	++	++	++	++
Visualize Output	+	++	+	++	++	++	+	++	++	++	++
Retrieve Company Lingo	o	o	o	o	o	o	o	o	o	o	o
Free to use	++	++	++	++	o	++	++	++	++	o	o

o Does not apply or cannot be used in this case
\+ Does apply in some cases, does collect limited information
++ Does fully apply, does gather the amount/quality of information needed

Table 3: Potential for Applicability

Category	Maltego	Recon-ng	Cree.py	Spokeo	The Wayback Machine	The Harvester
Proposed Purpose	Delivery of a threat picture of an organization's environment.	Enables conduction of web-based reconnaissance.	Provision of geolocation related information from social media.	Provision of personal information.	Archive for webpages and other media	Gather e-mails, subdomains, hosts and open ports from different public sources.
Price	Free community edition, Full license $760first year, $320additional year.	Free, API Keys up to $60,000.	Free.	Free basic information, $4.95month for detailed reports, $9.95 for court records.	Free.	Free.
Usability	Easy to understand UI. Basic knowledge about structure and connection of information and available machines required.	Terminal based tool. Basic knowledge about structure and connection of information and available modules required.	Easy to use due to UI and step by step guidance.	Easy to use due to step by step guidance.	Easy to use due to centralization in single search field.	Terminal based tool. Simple execution.
Input Parameters	Depending on the machine name, web domain, username, company name.	Depending on the module domain name, URL, name.	Username.	Name, e-mail, phone, username, address.	Web domain.	Company name, web domain.
Visualize Parameters	Bubble diagram. Color coded data categories. Bubble sizes according to data amount.	Local database exportable to various formats.	Data listed, pins on map.	Pins on map.	Calendar based data entries. Available snapshots highlighted.	Data tables.
Relevant Phases	Phase 1 - Pre-Engagement Interactions, Phase 2 - Intelligence Gathering	Phase 2 - Intelligence Gathering	Phase 2 - Intelligence Gathering	Phase 1 - Pre-Engagement Interactions, Phase 2 - Intelligence Gathering	Phase 1 - Pre-Engagement Interactions, Phase 2 - Intelligence Gathering	Phase 2 - Intelligence Gathering

Our goal is to show the utility of these tools for attackers. Therefore, we selected three attack types mentioned repeatedly [27, 17, 23]: *Phishing*, *Baiting*, and *Impersonation*. We describe these below including their needs of two essential information categories: *communication channels* and *company knowledge*. An attacker requires communication channels since the attacker has to communicate with a victim to exploit her trust. In addition, an attacker requires knowledge about the company to know whom to attack and how to get the companies employees' trust. The more details an attacker knows, the more likely people believe he has a relation to the company. We detail these information needs for the attack types below and refine them in Table 4.

Phishing refers to masquerading as a trustworthy entity and using this trust to acquire information or manipulating somebody to perform an action. This often appears in an unguided way via email to thousands of possible victims. Recently, spear-phishing attacks happen, which aim for a specific target instead of the broader mass. The social engineer gathers as much intelligence about the target as he can or needs and then prepares a tailored message for the victim.

Information needs: Phishing attacks are mainly based on communicating with the victim, therefore the amount of information on communication channels is critical. The more channels an attacker has, the easier it is, to find one that can help bridge the gap between the engineer and the victim. In addition, the more company knowledge exist, the more targeted the attack can be.

Baiting is to leave a storage medium (e.g., a USB stick) inside a company location that contains malicious software (e.g., a key logger). The malicious software is executed automatically when the stick is inserted in a computer.

Information needs: Baiting is a passive attack vector, which does not need direct interaction with the victim. Therefore, the focus lies on gathering company knowledge. In particular, locations and walking routes of employees for placing the storage medium are essential.

Impersonation is to play the role of someone a victim is likely to trust or obey, e.g. an authority figure. The attacker fools the victim into allowing him access to the desired location or information. Usually, attackers prepare well for an impersonation and leverage vast amount of information.

Information needs: For a successful impersonation attack company knowledge is a priority. The social engineer needs knowledge of numerous areas of the company. The more information he has on the persona he is playing, the more convincing he can be. Communication channels are of less importance, since the victim is approached in person.

We illustrate the degree to which the information needs of a social engineer can be covered for the discussed attack types. Tables 5 and 6 match tools with communication channels and company knowledge. Table 6 reveals that numerous tools cover information gathering for locations, websites, new employees etc. of companies. However, the *Company Lingo* is not covered at all. Company lingo contains all abbreviations and specific terms used in a company and has been used by social engineers to bypass authentication mechanisms, e.g. personnel often thinks everyone knowing the company lingo belongs to the company [27].

Table 4: Mapping of Social Engineering Characteristics to Attack Types

		Attack Type		
		Phishing	Baiting	Impersonation
Communication	Telephone Number	x		
	Friends	x		x
	Personal Information	x		x
	Private Locations	x		x
	EMail	x		
	Instant Messenger	x		
	Co-Workers: Communication			x
Company Knowledge	Co-Workers: New Employee			x
	Co-Workers: Hierarchies			x
	Lingo	x		x
	Facilities: Security-Measures		x	x
	Facilities: Company Location		x	x
	Websites	x		
	Policies: Software		x	
	Policies: Network		x	
	Policies: Organization		x	

Table 5: Tool Coverage for Communication Channels

	Cree.py	Gitrob	KnowEm	LinkedIn	Maltego	Namechk	Recon-ng	Spokeo	theHarvester	Wayback Machine	Wireshark	Xing
Telephone Number							x					x
EMail				x	x		x		x			x
Instant Messenger			x		x	x		x				x
Friends			x	x	x	x						x
Personal Information	x		x	x		x		x				x
Private Locations	x							x				x

For “Facility Security Measures”, “Security Policies” and “Software Policies” there is a similar result. Besides *theHarvester* and *Recon-ng*, which can both only gather information concerning web-security like open ports or SSL-Encryption, all other tools are not directly suitable for social engineers. *Wireshark* needs physical access, which is not exactly what a social engineer prefers and *Gitrob* is one of the tools, with very slim chances of success. If the company has any security policies or hosts their sourcecode within the company, then *Gitrob* will most likely not be able to access it and therefore not gain any information.

To sum up, modern social engineers have a variety of tools at their disposal for information gathering, which they can use in numerous attacks. We provide an exemplary overview for phishing, baiting, and impersonation attacks and summarize in Table 7. The empty fields mean that three security researchers could not identify a use for that tool for any of the attacks above. Note that there are still some types of information that are difficult to gather for an attacker such as company lingo, but we have little doubt that in the future further tools and social media offers will fill this gap. Furthermore, our comparison showed that all tools have a good or great usability and provide easy to understand output.

Table 6: Tool Coverage for Company Knowledge

	Cree.py	Gitrob	KnowEm	LinkedIn	Maltego	Namechk	Recon-ng	Spokeo	theHarvester	Wayback Machine	Wireshark	Xing
Company Locations	x			x			x	x	x			x
Company Lingo												
Special Knowledge				x	x		x					x
New Employees				x	x							x
Hierarchies				x	x							x
Websites					x		x		x	x		
Facility Security Measures		x									x	
Security Policies		x							x		x	
Software Policies		x					x				x	

Table 7: Tools vs. AttackType Knowledge with P for Phishing, I for Impersonation, and B for Baiting

	Cree.py	Gitrob	KnowEm	LinkedIn	Maltego	Namechk	Recon-ng	Spokeo	theHarvester	Wayback Machine	Wireshark	Xing
Telephone Number							P					P
Friends			P,I	P,I	P,I	P,I						P,I
Personal Information	P,I		P,I	P,I		P,I		P,I				P,I
Private Locations	P,I							P,I				P,I
E-Mail				P	P		P		P			P
InstantMessenger			P		P	P		P				P
Co-Workers: NewEmployee				I	I							I
Co-Workers: Hierarchies				I			I					I
Lingo												
Facilities: Security-Measures		B,I									B,I	
Facilities: Company Location	B,I			B,I			B,I	B,I	B,I			B,I
Websites					P		P		P	P		

This means intelligence gathering can be used by an attacker with little technical knowledge such as script kiddies. Therefore, we have to take the threats arising from increased and easily available knowledge for social engineering seriously.

4 Conclusions

We conducted a structured survey of social engineering tools, which ease the attacker's effort of finding information about victims. We mapped the information to their usefulness for phishing, impersonation or baiting attacks. Our analysis revealed that the social engineering threat is more dangerous than ever before,

due to the number of tools at an attacker's disposal and the significant amount of detail they provide. We propose the following.

Implications for possible Victims People in general, not only employees in companies, can fall victim to social engineering. Therefore, people should find out what is available about them in the web using the tools or websites listed here. Ideally, stories of new contacts and unusual requests to secret information should be checked and verified more carefully than in the past. Means of protection can include false information released such a bogus address or non-existing hobbies. Any requests using this information identify possible social engineers.

Implications for Security Practitioners Chief information officers and consultants should integrate a demonstration of the tools in this publication to raise awareness of the social engineering threat in companies. Just when people see the ease of collecting information with the tools and websites and how these are used e.g. in phishing, they can understand the need for strict security policies with regard to the release of data in the web.

Suggestions for Law Enforcement has to operate under the assumption that criminals will get all information about their victims without ever leaving their home or having mature computer skills. Everyone can be a social engineer and is a possible perpetrator. Countermeasures have to include network traffic analysis of how an attacker gathered the information for his attacks.

Limitations of the Tools The only information type that social engineering tools do not provide today is the so-called *company lingo*, the abbreviations and specific words used in a company or domain. However, we are certain that in the future, tools combining machine learning and big data analysis will fill this gap.

Limitations of our Study We conducted the study using a previous survey of tools and a web search engine. These sources can be extended in particular to including sites that are not indexed by web search engines e.g. in the dark web. This work will require a collaboration with a law enforcement agency.

Acknowledgements

This research has been partially supported by the Federal Ministry of Education and Research Germany (BMBF) with project grant number 16KIS0240.

References

1. Freebackgroundcheck. `https://mybackgroundcheck.preemploy.com`.
2. Instant checkmate. `https://www.instantcheckmate.com`.
3. Norwegian register. `http://skattelister.no/`.
4. Tax information. `http://www.veroporssi.com/`.

5. Whitepages. http://www.whitepages.com.
6. N. Barrett. Penetration testing and social engineering: hacking the weakest link. *Information Security Technical Report*, 8(4):56–64, 2003.
7. BBC News. How to hack people. news.bbc.co.uk/2/hi/technology/2320121.stm, October 2002.
8. CareerBuilder. Job search engine. http://careerbuilder.com/.
9. Dimensional Research. The risk of social engineering on information security. http://docplayer.net/11092603-The-risk-of-social-engineering-on-information-security.html, September 2011.
10. T. Dimkov, A. van Cleeff, W. Pieters, and P. Hartel. Two methodologies for physical penetration testing using social engineering. In *Proceedings of ACSAC*, ACSAC '10, pages 399–408. ACM, 2010.
11. Dun & Bradstreet. Sales acceleration platform. http://www.hoovers.com/.
12. Edge-Security. theharvester. http://www.edge-security.com/theharvester.php.
13. Glassdoor. Recruiting website. https://www.glassdoor.de/.
14. D. Gragg. A multi-level defense against social engineering. *SANS Reading Room, March*, 13, 2003.
15. R. Gulati. The threat of social engineering and your defense against it. *SANS Reading Room*, 2003.
16. Hadnagy. Social engineering toolkit (set). http://www.social-engineer.org/framework/se-tools/computer-based/social-engineer-toolkit-set/.
17. C. Hadnagy. *Social engineering: The art of human hacking*. John Wiley & Sons, Indianapolis, 2010.
18. C. Hadnagy. The Official Social Engineering Portal, 2015.
19. Internetsafety 101. *Social Media Statistics*, 2013. http://www.internetsafety101.org/Socialmediastats.htm.
20. Kakavas. Geolocation OSINT Tool. http://www.geocreepy.com/.
21. J. Kee. Social engineering: Manipulating the source. *GCIA Gold Certification*, 2008.
22. KnowEm LLC. Social media brand search engine. http://knowem.com/.
23. K. Krombholz, H. Hobel, M. Huber, and E. Weippl. Social engineering attacks on the knowledge worker. In *Proceedings of Security of Information and Networks*, SIN '13, pages 28–35. ACM, 2013.
24. LinkedIn. Business social networking service. http://linkedin.com/.

25. MarketVisual. Business search engine. http://www.marketvisual.com/.
26. N. Milosevic. Introduction to Social Engineering, 2013.
27. K. D. Mitnick and W. L. Simon. *The Art of Deception: Controlling the Human Element in Security*. 2003.
28. Monster Wolrdwide Inc. Job search engine. http://monster.com/.
29. Namechk. Username and domain search tool. https://namechk.com/.
30. National Association of Counties. http://www.naco.org/.
31. Pakistan Government. Federal board of revenue. http://www.fbr.gov.pk/.
32. Paterva. Maltego clients and servers. https://www.paterva.com/web6/products/maltego.php.
33. Public Accountability Initiative. http://littlesis.org/.
34. Ratsit & Invativa. Credit business website. http://www.ratsit.se/.
35. K. Regan. 10 Amazing Social Media Growth Stats From 2015, 2015.
36. Shodan. Search engine for the internet of things. https://www.shodan.io/.
37. socialmention. social media search platform. http://socialmention.com/.
38. Spokeo. People search website. http://www.spokeo.com/.
39. The Internet Archive. The wayback machine. https://archive.org/web/.
40. T. Tomes. Web reconnaissance framework. https://bitbucket.org/LaNMaSteR53/recon-ng.
41. Verizon. Data Breach Investigations Report, 2012. http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2012-ebk_en_xg.pdf.
42. Verizon. Data Breach Investigations Report, 2013. http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2013_en_xg.pdf.
43. M. Warkentin and R. Willison. Behavioral and policy issues in information systems security: the insider threat. *European Journal of Information Systems*, 18(2):101–105, 2009.
44. G. Watson, A. Mason, and R. Ackroyd. *Social Engineering Penetration Testing: Executing Social Engineering Pen Tests, Assessments and Defense*. Syngress, 2011.
45. Xing. Business social networking service. http://xing.com/.

All online references were last checked on 12.01.2017.