Social Engineering 2.0: A Foundational Work

Invited Paper[†]

Davide Ariu
Università degli Studi di Cagliari
Piazza d'Armi, 09123 Cagliari
Italy

Pluribus One S.r.l.
Via Bellini 9, 09128 Cagliari
Italy
davide.ariu@pluribus-one.it

Enrico Frumento
CEFRIEL

Via Renato Fucini 2, 20133 Milano
Italy
enrico.frumento@cefriel.com

Giorgio Fumera

Università degli Studi di Cagliari
Piazza d'Armi, 09123 Cagliari
Italy
fumera@diee.unica.it

ABSTRACT

During the past few years, social engineering has rapidly evolved and has become a mainstream technique in cybercrime and terrorism. It is used especially in targeted attacks involving complex human and technological exploits, aimed at deceiving humans and IT systems. Building on the work carried out in the DOGANA project, funded by the European Union, this paper provides an overview of the evolution and of the current landscape of social engineering, and introduces as its main contribution a theoretical model of how human exploits are built, named the Victim Communication Stack.

CCS CONCEPTS

• **Security and privacy~Phishing** • Security and privacy~Malware and its mitigation • Security and privacy~Economics of security and privacy

KEYWORDS

Phishing, Social Engineering, Malware.

ACM Reference format:

D. Ariu, E. Frumento, G. Fumera, 2017. Social Engineering 2.0: A Foundational Work. In *Proceedings of ACM Computing Frontiers conference, Siena, ITALY, May 2017 (COMPUTING FRONTIERS'17)* 7 pages. DOI: 10.1145/3075564.3076260

1 INTRODUCTION

Until the end of the past century, Social Engineering (SE) was an advanced, but niche, way of attacking specific systems by exploiting humans involved in them. Since few years it is evolving at an incredible pace to what has been called SE 2.0, which is nowadays mainstream in cybercrime and terrorism. "Old-school" SE was an adaptation of the ageless art of deception to the modern communication media; it required personal talent and effort, and was limited to few attackers who focused only on valuable targets. Accordingly, traditional Information Security considered the "human factor" as a potential threat only in systems where «security-in-depth» was required. SE 2.0 is characterized instead by an incredibly higher complexity of attacks that actively exploit the human element to enable the subsequent technological step [1]. This evolution is motivated by the increasing relevance of Targeted Attacks (TAs) in today's attack strategy [2]. TAs are a specialized combination of complex "Human Attack Vectors" with technological exploits, aimed at deceiving humans and IT systems.

SE 2.0 includes and extends the old-school SE concepts into a wider vision: the key difference between them is the possibility to exploit the SE techniques on a larger scale, using automated attacks on a potentially large number of victims. The transition to SE 2.0 was triggered by the large amount of machine-readable data that is freely available today. This trend has been exponentially strengthened by the advent of Social Networks and the new social trends of information sharing. As a consequence, companies and public bodies, as well as persons, became tremendously exposed to SE 2.0, and thus prone to targeted cyber-attacks. Another important aspect is the involvement in the attack planning of competences such as psychologists, marketing experts and in general all the human sciences, that had never been previously seen in the cybercrime world, and are becoming requested by organized crime groups to better understand how to "exploit the humans".

Today we face an unseen and highly dynamic situation, where humans are increasingly the "system" under attack. Moreover, there is currently no solution available on the market that allows a comprehensive assessment of social vulnerabilities and the management and reduction of the associated risk. Filling this gap is the aim of the ongoing DOGANA project (aDvanced sOcial enGineering And vulNerability Assessment), funded by the European Union Horizon 2020 framework programme.[1] Its underlying concept is that Social Driven Vulnerabilities Assessments (SDVAs), when regularly performed within an efficient framework, supports the effective deployment of mitigation strategies and leads to reduce the risk created by SE 2.0 attack techniques. The framework proposed by DOGANA is centered around two main features: the presence of the "awareness" component as the cornerstone of mitigation activities, and the legal compliance by design. This paper builds on the work

[1] http://www.dogana-project.eu/index.php

CF'17, May 15-17, 2017, Siena, ITALY
 ISBN: 978-1-4503-4487-6/17/05. . . $15.00
DOI: http://dx.doi.org/10.1145/3075564.3076260

carried out by the DOGANA partners, and summarizes some of the achieved results. The main, original contribution is the introduction of a theoretical model of how Human Attack Vectors are built, named the Victim Communication Stack (VCS). In Sects. 2 and 3 we describe the evolution of SE and describe the main features of SE 2.0, discussing its relationship with modern cybercrime and cyberterrorism trends. In Sect. 4 we describe the "attack vectors" used to carry out SE 2.0 attacks, and present the VCS model. An example of a very recent SE 2.0 attack is reported in Sect. 6.

2 THE EVOLUTION OF SOCIAL ENGINEERING

SE is a well-known method of deception. It has been used for a very long time, and has evolved today into a significantly novel threat. In this section we summarize the early SE and the driving forces behind its evolution into modern SE.

2.1 Old-School Social Engineering

Early SE was an adaptation of the ageless art of deception to the modern communication media, mainly phone and early usage of email. Its main feature was the high level of ability required to the attackers, as they had to be directly involved in all the attack phases. Therefore, early social engineers were all IT experts or very talented people well prepared in hacking logics, who concentrated on very valuable targets [3-6]. This phase is called "old-school" SE, mostly because the above assumptions are not true anymore: SE is becoming increasingly simpler and requires less knowledge to the attackers. They are thus increasing in number, and can achieve results not comparable to old-school SE due to the involvement of professionals such as psychologists, marketing experts and cognitive scientists. The driving forces behind the evolution of SE are summarized below.

2.2 Forces Driving the Evolution of Social Engineering

Modern SE is the result of the evolution of technology and society, as well as of cybercrime.

Technology evolution has enabled new kinds of SE attacks due to several factors: the widespread use of mobile, wearable and smart devices at home, in public spaces and in company premises; the large bandwidth available from communication service providers; the widespread availability of social networking platforms and of the associated applications, to which a wealth of information is provided by users and their smart devices; the new methods constantly offered by the market to access a user's own dataspace, and the proliferation of virtual and augmented reality devices, such as Google's Cardboard and Microsoft HoloLens, that enable new kinds of interactions with online social networks. Another enabling technology is the spreading of online payment systems in many different environments.

At the same time, **society** evolved into a blended life world where physical and virtual experiences seamlessly merge. An increasing amount of human activities (e.g., working, communicating, online banking, information sharing, travelling, and entertaining) is being integrated in the online social networks. Individuals are becoming accustomed to a complete dematerialization of the personal dataspace on centralized cloud services, and workers want to complete a task in any possible place; this is enabled by several easy-to-use, context-sensitive tools to access one's own dataspace (see [7]). These factors are producing an always-available culture with no separation between personal and professional life, and an economic climate where human actions as projected in the social network can be profitable for data aggregation and service providers.

Another key driver of modern SE is the evolution of **cybercrime** into a fully-fledged industry, named "cybercrime as a service" (CaaS), which includes suppliers, service providers, markets, financing, trading systems, and many different business models [8] . CaaS has been enabled in turn by the use of cryptocurrencies, and by the exploitation of the "Dark Web" and of encryption technologies by cybercriminals. The consequence is that even unskilled or entry level cybercriminals can launch large-scale attacks in terms of risks, costs and profits [9], capable of causing significant damage to a large number of victims, thanks to the easy access offered by CaaS to criminal products and services. Beside technology (e.g., exploiting newly discovered defects in computer software), cyber-attacks often rely on human error (e.g., employee deception). CaaS is difficult to fight by law enforcement agencies due both to limitations of trans-borders investigations and to the use of the Dark Web.

3 SOCIAL ENGINEERING 2.0

The factors summarized in Sect. 2.2, together with the naive behaviour of users of online services (mainly social networks), contributed to the evolution of SE into a new multifaceted, complex phenomenon that we call SE 2.0. The amount of potential victims directly exposed on the internet has enormously increased, and attackers can use advanced, automatic techniques to gather and process the information needed to carefully select their targets. SE 2.0 involves the following heterogeneous technological and scientific fields.

Malware Ecosystem 2.0: SE became an important part of malware, and its main infection strategy, implying a change in the infection strategies and in the development process of new malware [20].

Modern Open Source Intelligence (OSINT): modern SE uses data mining techniques to collect information before the attack, exploiting the large amount of data that people and enterprises share intentionally or inadvertently on the internet, especially through social networks. Beside digital shadows and footprints, OSINT increasingly exploits the Web 3.0 (web-of-data) as a data source [25]. Abuse of publicly available information is a huge opportunity to improve the efficiency of information gathering in a SE attack. Therefore, social networks, the Web, etc. became enablers of SE 2.0, as they provide structured data that is easy to access and to automatically process.

(Ab)use of psychology, personality profiling systems, cognitive science models and human related sciences. SE consists in hacking humans using the most efficient ways available; therefore, psychology and other human sciences are frequently used to gain knowledge of the "vulnerabilities" of human targets. Cybercriminals are becoming more professional, and increasingly use memetics[2] [12] and personality models of victims [13], especially models from cognitive sciences, marketing and cyber-sociology theories [14]. Psychological profiling (e.g., identifying the most vulnerable victims) [15], use of memetics [12] and sentiment analysis [16] are used to rapidly contextualize and tailor attacks around selected victims with a localized approach.

Evolution of the attack vectors. Understanding victim's psychology has led to changing the way hooks are crafted and delivered. Massive usage of spam is no more the main technique; nowadays spam is mainly used to collect the so-called "low hanging fruits" to supply the cybercrime world with a low but constant flow of incidents. Advanced Persistent Attacks (APTs) are the most effective ones, instead: they massively use social networks and novel forms of phishing (spear- and context-aware phishing, collectively called *-phishing). As result, attack vectors have multiplied, and the modern *-phishing are no more tied to specific channels.

Automatic Social Engineering Attacks (ASE). One of the main features of modern SE is the possibility to automate most of the attack phases, increasing the efficiency of mass SE-based attacks. This is enabled by the automation of information collection and data mining from social networks, e.g., thanks to the improvement of sentiment analysis algorithms [17].

Economic Drivers. Whereas malware can be created just for fun and to prove the technical skills of the author (which were in fact the main motivations of early generations of malware), using SE for fun makes less sense: its only goal has always been to deceive people, often to make a profit. This led SE 2.0 to become an effective tool to carry out serious attacks, as well as a fruitful investment. The growth of identity thefts, industrial spying, on-demand attacks (Deny-of-Service on demand), commoditization of SE services in cybercrime and cyberterrorism are all consequences of the evolution of SE [8].

Most of the above methods and techniques used in SE 2.0 have been originally developed and used legitimately in different contexts, and have been abused by social engineers to collect information for performing highly contextualized attacks. For instance, some of the above techniques come from social marketing, and were originally used to catch and to influence social trends; also in SE they are used to influence people's way of thinking, but with malicious intentions.

Summing up, the real issue of SE 2.0 is the *abuse* versus the *use* of known methods. In particular, methods exploited in SE 2.0 are taken not only from technology, but also from human and social sciences like psychology and cyber-sociology. Fig. 1 summarizes the main features of SE 2.0.

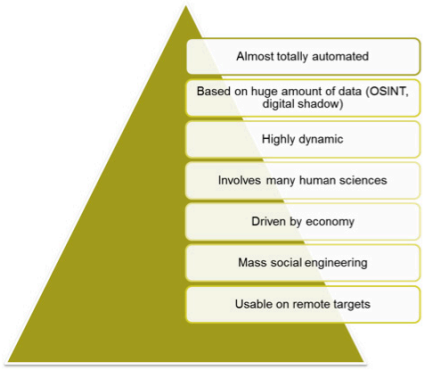

Figure 1. Main features of SE 2.0.

4 HUMAN VS TECHNICAL ATTACK VECTORS

In general, the ultimate target of an attack is the legal holder of an asset, or a human or ICT system that is involved in its handling. The aim of an attack is therefore to steal an asset from one of its handlers, which amounts to abuse a trust chain between systems and/or humans. An "attack vector" (AV) is the path, or the mean, or the set of operations by which this purpose is achieved; this includes, e.g., viruses, e-mail attachments, Web pages, pop-up windows, instant messages, chat rooms, and deception. The AV enables hackers to exploit vulnerabilities of systems and/or humans; quoting from [7], it "should contain all the elements of a social engineering attack". It has therefore a goal, a target and a social engineer; in addition, the attack plan must identify a medium, compliance principles and techniques.

Two kinds of AVs can be distinguished: the human AV (HAV) and the technical AV (TAV), which target respectively humans and ICT systems. For instance, old-school SE attacks (e.g., non-ICT-enabled frauds) are purely based on HAV, whereas attacks like automated infections are purely based on TAV.

4.1 Human Attack Vectors

HAVs are complex entities made up of different interacting components, that we propose to represent together as the "Victim Communication Stack" (VCS) model of Figure 2.

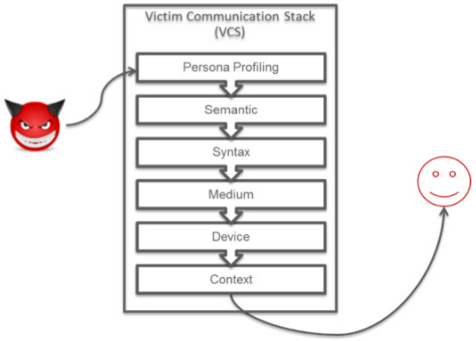

Figure 2. Victim Communication Stack (VCS) model.

[2] Memetics is the science that describes the spread and diffusion of ideas (https://en.wikipedia.org/wiki/Meme). It finds an interesting application in the area of SE and human manipulation. Engaging the emotional side of people and bringing them to adopt memetic behavior is the "philosopher's Stone" of online marketing, as well as of SE.

The VCS is a theoretical model that can be used to create HAVs for specific attacks. Given a target asset, the attacker needs to build a victim model in order to select and develop the HAV which is most likely to enter the victims trust zone and perform the attack. The interdependent layers of a VCS comprise the "Victim Model". They consist of psychological and behavioral aspects which need to be harmonized in order to deliver the correct HAV. As Figure 2 shows, the attacker starts assembling the HAV from the upper layer "persona profiling", and then defines the activities in the lower layers. The victim receives an entire HAV instance, which is shaped around his/her specific habits, communication register, psychological profile and so on.

Persona Profiling. This layer refers to methods and theories used to identify the profile that can provide the target asset, and is possibly the weakest link toward it. The corresponding actions consist in collecting information from different sources to identify initial elements required to proceed in building the HAV, e.g., age, sex, cultural and professional background, habits, etc. This layer also includes passive psychological profiling of the persona, which is performed by matching its characteristics with the common characteristics (persona template) of a specific category of persons he/she belongs to (e.g., common habits towards social networks or smartphones of younger people, if the persona is under 30 years, vs aged persons if he/she is above 50 years).

Semantic. This layer involves the psychological aspects on which the HAV will leverage: it comprises classic and advanced persuasion/manipulation techniques chosen according to the persona profiling. This layer defines the right communication semantic register to be used by the attacker to leverage the common habits of the victim, that are in turn defined by the Persona Profiling layer. This phase can involve classic persuasion techniques, or more sophisticated instruments like memetics.

Syntax. This layer defines the elements selected as the content of the message, for example wording, tone, graphics, etc. These elements depend on, and need to be consistent with, the other layers. For instance, thanks to the services of the upper layers, the syntax layer shapes the design of the hook (look, graphic, type of phishing) and the linguistic register. The term "syntax" is usually associated to the structure of a language (rules used to deliver a message), and is used here since the design of the hook is a way to deliver the message (i.e., the semantic) of the VCS.

Medium. This layer defines the channel used to deliver the message, e.g., voice, email, chat, social network, etc. The choice of the medium greatly influences the type of interaction that the attacker wants to establish with the victim, e.g., social network vs email or voice chat, as well as its characteristics, e.g., real-time for chats, asynchronous for emails, direct for voice calls.

Device. This layer refers to the device on which the victim will receive the message. This choice could be critical due to the different scenarios originated by the way people interact and evaluate the medium on which contents reach them. The same physical channel may trigger different reactions according to the medium used, e.g., reading and email on a laptop or on a smartphone correspond to different user experiences, and thus produce different sets of possible outcomes.

Context. This layer describes where and when the attack is delivered. This is an important aspect, as it affects the overall credibility of a hook: the same hook delivered in two different contexts usually gives different results. This layer answers the following questions: when is the attack launched, and where is the victim supposed to be at that time, in order to maximize the attack effectiveness? The two most important aspects to consider are therefore the timing (e.g., at what time of the day, or in which days of the week), and the location (e.g., on a train, in an airport, when the victim is driving, or when the victim is abroad). This layer then focuses on the chosen delivery context of the HAV.

4.2 Technical Attack Vectors

Spam and its evolution into phishing and its variants are among the main TAVs. **Spam** is a generic mail sent identical to millions of victims with a flat approach. The revenue model is simply tied to the probability to hit a vulnerable person (someone who falls into the hook). It could be graphic or not, but the discriminant is always that the hooks are generic being applicable possibly to anyone. It is a blind massive form of attack. **Phishing** is a more sophisticated form of spam that, thanks to graphics, delivers a more sophisticated hook, specialized for a subsample of users belonging to the targeted company (e.g., customers of a targeted Bank are more likely to fall into this hook than non-customers). The business model is not flat. It is usually sent to less people than spam but also to people not belonging to the chosen user category, supposing for them the hook just does not work. **Spear phishing** is a specialized form of phishing which is sent only to the customers of the company which the mail pretends to come from. Its return is greater, because actual customers of the targeted organization are selected as victims. Victims are selected on the Social Networks using OSINT techniques or setting up customers' assistance un-official pages on the social networks. Spear phishing is the most common attack on internet today, accounting 95.22% of the attacks [18]. In **context aware phishing** (aka "whaling"[19]) the semantic distance to a real email is minimal. These emails are crafted around the few selected victims of the attack, which are found using OSINT operations. It is usually used in APT or Targeted attacks. **Vishing** is a combination of "voice" and "phishing": the telephone is used to acquire information or to attempt to influence actions, possibly exploiting Voice over Internet Protocol (VoIP) technology to "spoof" the attacker's outgoing number. Vishing attacks are riskier for the attackers and require them an extra effort.

In general, AVs used in the different kinds of phishing attacks can be categorized into behavioural and non-behavioural. Non-behavioural AVs consist of non-real-time and non-interactive media like social networks and emails, and real-time interactive media like chat and instant messaging systems. In particular, email-based attacks have to be crafted completely offline, and the victim must be convinced in "one shot" just by looking the hook (i.e., the email); social networks do not allow "one-shot" attacks, instead: the hook can be adjusted according to the target reactions.

Chat- and instant messaging-based attacks use a virtual communication channel which is controllable and is not able to transmit non-verbal messages. Behavioural AVs consist of voice and physical presence. The former is a real-time communication channel that conveys also some non-verbal behaviour (e.g., the tone of the voice), and requires some special skills to control them. The latter is the most complex AV, since beside the hook the attacker must control all the non-verbal elements (also the unconscious ones), to avoid revealing his/her own final intentions through them.

Other TAVs are the following. **Pop-up windows** are an older kind of attack: its TAV consists of software delivered to the end user's terminal, to steal username and password of some protected resource. **Interesting software** is a kind of attack constantly present on the Internet, where users are persuaded to voluntarily download and install a very useful program or application, such as CPU performance enhancer, a great system utility or a crack to an expensive software package. When they do that, malicious software (malware) is installed. This kind of attack is commonly used by hackers, as it does not require their active participation in the scam; they only have to develop the malicious software, design its interface to make it look like a legitimate program, place it in a server and wait for people to download it. Many users download these programs because they are not aware that they may be fake, or would like to avoid paying for the official ones. Hackers are aware of this behaviour, and fill the Internet with malicious programs to leverage it. **Malware** has been affected by the improved efficiency of SE 2.0 and has in turn evolved into what is called **malware 2.0**, which differs from malware identified in recent past. Its main characteristics are the following [20]: lack of a single control centre and ability to adapt the infection to the attacked machine; extensive use of methods to fight anti-virus systems; victim machines take the role of servants and attacks get more discrete; intense production on syntactic - not logical - variations; short and targeted attacks from many directions; intense and advanced use of SE techniques; modularity and complexity of infections; malwares and SE follow the market laws governed by supply and demand (MaaS)[21]. In other words, the exploit starts with an HAV and continues with a TAV. Counted as 100% the overall vulnerability abused by malware, resulting by a sum of HAV and TAV, what differentiates the malware today is the relative complexity of the human exploit, which simplifies the technological one. Previous generation malware was like a Rover on Mars: it needed to explore and unknown environment and survive, because the target machine context was, most of the times, unknown. Today, the selection and fingerprinting of the victims allows for a different strategy: "I (the attacker) am smarter than you are", whereas the malware "only" needs to be smarter than the selected target.

4.3 Social Driven Vulnerability Assessment

As discussed above, the complexity level of attacks based on non-technological exploits, and in particular on the human element, is incredibly high. Attacks have become narrower and involve less generic victims. This is a consequence of improved hiding tactics which aim at reducing the risk of being detected, but is also a sign of a better a-priori selection of the potential targets and thus a more aggressive usage of SE techniques. Since few years Cefriel is exploring the role and the evolution of SE in the attacks, to develop solutions to measure, mitigate and patch the human element of security. The first and more concrete approach was, since 2010, the development of an original methodology for performing vulnerability assessment of the employees of an organization, called Social Driven Vulnerability Assessment (SDVA) [22]. The main problems behind a European-wide adoption of the SDVA methodology were of moral and legal nature since, in particular, the SDVA would somehow require the monitoring activity carried out by the employer to interfere with the private life of the employees. As labor law of European countries typically do forbid such kind of interference, the DOGANA project is seeking for a trade off between the technical requirements and needs and moral and legal compliance.

The average results of SDVAs carried out so far provided relevant insights, for instance: on average, 3 emails are enough to obtain one click (34% average click-rate); 5 emails on average to obtain a valid credential; after 2 hours the attack is exhausted and almost 100% of the potentials victims is captured; high promptness of the attack: 4 minutes to capture 20% of the victims, 40% after 10 minutes, 50% in 20 minutes; slow reactions: 6 minutes to report the phishing attempt and 20 minutes to block the site (fastest reactions ever recorded) [23].

If the human is the "system" under attack, it follows that all the human sciences must be involved in a multidisciplinary effort to model the attacked target, i.e., to define the vulnerabilities which can be exploited, as pointed out in Sect. 4.1 (see also Fig. 3). This also implies that the same multidisciplinary approach involving all human sciences should be used to defend IT systems. Unfortunately, our experience shows this is not an easy task; on the other hand, it is clear that cybercrime already solved this dilemma, because the human sciences are known to be used to increase the efficiency of high-profile attacks. As an example, in its own SDVAs Cefriel understood that a wise use of memetic, marketing techniques and cognitive sciences enhance the effectiveness of penetration tests.

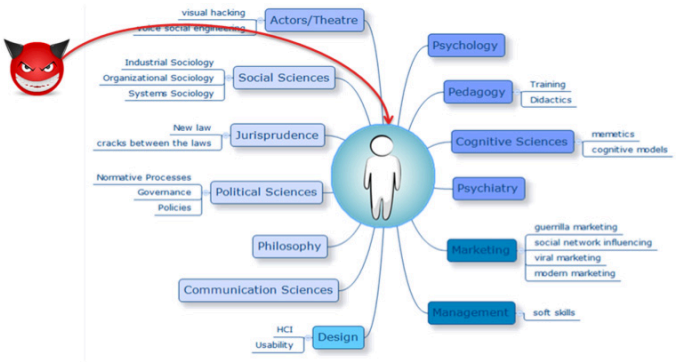

Figure 3. A non-exhaustive list of sciences involved in the definition of the human target in modern SE (Source: CEFRIEL).

6 A RECENT EXAMPLE: THE FREE iPHONE 6 SCAM CAMPAIGN

Scam campaigns advertised through social networks represent a very typical way of leveraging human weaknesses to deliver cyber-attacks. Such campaigns typically leverage hot topics which can quickly guarantee significant returns in terms of number of clicks.

A widely-known campaign is the one associated with the release of Apple iPhone 6. In 2014, Apple had announced the release of the iPhone 6 for the second half of September (on 19th). Almost one month before the official release, attackers started a campaign on Social Networks (Figure 4 shows the Facebook page of the campaign) which promised to give away for free 500 iPhone 6 to 500 "Lucky Winners". The promise was quite appealing as winners would have the possibility to get the phone (which was expected to be on great demand and thus not so easy to get) only 10 days after its official release. This campaign, provides a clear example of all the layers of a *Victim Communication Stack*. **Persona profiling** is in fact easily carried out on Social Networks, both manually by friend requesters or eventually also using malware (which can spread through clickjacking). The **Semantic** layer leverages on the desire of the victim for social validation and liking, and stimulates its desire to possess a trending good. The same elements leveraged on by the semantic layer, also ensure that every victim shares the message with his contacts making the campaign progressively more and more effective. The **Syntax** is that typical of social media (Facebook is, in this case, the **Media)**, made primarily of short and immediate sentences and messages, also supported by graphical contents. In this specific case, the message is really short (less than 400 characters), provides very elementary instructions to the participants and requests to make really simple actions (which would require about 15 seconds to complete) thus ensuring the participation also of not so interested people. The Device is of course represented by the **Facebook** client, which might be either simply the web browser or an application installed on a mobile device. In the latter case of a mobile device, the **Context** might be less frequently a work context, as the mobile device is used in transportation or eventually during relax moments. Again, the simplicity of the actions requested allows them being taken almost anywhere and anytime a user has a few seconds of spare time.

Figure 4. The Facebook page of the Apple iPhone 6 scam campaign.

In the case of scam campaigns like this, the Technical Attack Vector consists of:

- A web page hosted under a domain name that whilst being an abused one (typically using *typo-squatting* [24]) is still a credible one;
- an infection agent assembled ad-hoc by a malware forgery, based on the fingerprint of the victim's browser and operating system. The agent usually exploits a vulnerability of the web browser (or of one of its plugins) to infect the victim's machine.

From the analysis it turns out that technically speaking the attack is really simple, and that the most of this effectiveness is due to the exploitation of the weaknesses on the human side.

7 CONCLUSIONS

The ultimate target of an SE attack is the legal holder of an asset, or a system that is somehow involved in its handling. It could be a human or an ICT system. The aim of an attack is hence always to follow the least resistant path among the handlers of an asset, abusing its trust, or in general a trust-chain among systems. In this paper we discussed the nature of SE 2.0 and its evolution as the main tool for modern attackers, to make money out of cyber-attacks. TAs are a type of attack which takes advantage of a complex HAV (usually performed via SE) tightly integrated with a TAV (usually performed via ad-hoc malware) to create a unique targeted and specialized ad hoc AV. The AV is used to exploit (deceive) both humans (through the HAV) and systems (through the TAV). For example, so-called pure SE attacks (e.g. non-ICT-enabled frauds) use only HAV, whereas attacks like automated infections are purely based on TAV.

Despite the role of SE in cybercrime, creating effective HAVs is still largely a challenging task which has some degrees of freedom, and requires the involvement of talented attackers. Organised crime groups and defenders are both seeking ways to automate as much as possible the creation of SE attacks, for respectively improving the attacks revenues and the efficiency of defences. In this context, the main contribution of this paper is the introduction of the Victim Communication Stack, a theoretic model of how to build HAVs, which is a topic not thoroughly investigated in the scientific literature so far.

ACKNOWLEDGMENTS

This work has been partially supported by the DOGANA project funded by the European Union Horizon 2020 framework programme, under grant agreement number 653618.

REFERENCES

[1] L. Kharouni et al., "Operation Pawn Storm Using Decoys to Evade Detection," Trendmicro, 2014. [Online]. Available: http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-operation-pawn-storm.pdf

[2] P. Paganini, "The differences between targeted attacks and advanced persistent threats," 2015. [Online]. Available: http://securityaffairs.co/wordpress/40228/cyber-crime/targeted-attacks-vs-advanced-persistent-threats.html.

[3] K. D. Mitnick, W. L. Simon, and S. Wozniak, *The art of deception: Controlling the human element of security*. Indianapolis, IN: Wiley, 2001.

[4] K. D. Mitnick and W. L. Simon, The art of intrusion: The real stories behind the exploits of hackers, intruders and Deceivers. New York: Wiley, John & Sons, 2005.

[5] lvxferis, "Hacking the mind for fun and profit," in *phrack.org*, 2010. [Online]. Available: http://phrack.org/issues/67/15.html. Accessed: Mar. 6, 2017.

[6] S. Granger, "Social Engineering Fundamentals, Part I: Hacker Tactics," 2001. [Online]. Available: http://www.symantec.com/connect/articles/social-engineering-fundamentals-part-i-hacker-tactics. Accessed: Mar. 6, 2017.

[7] E. Frumento, F. Freschi, "How the Evolution of Workforces Influences Cybercrime Strategies: The Example of Healthcare," in B. Akhgar, B. Brewster (Eds.): Combatting Cybercrime and Cyberterrorism -- Challenges, Trends and Priorities, Springer, 2015.

[8] K. Thomas, D. Huang, D. Wang, E. Bursztein, C. Grier, T. J. Holt, C. Kruegel, D. McCoy, S. Savage, G. Vigna, Framing Dependencies Introduced by Underground Commoditization, Workshop on the Economics of Information Security, 2015.

[9] European Cybercrime Center (EC3), The Internet Organized Crime Threat Assessment (iOCTA), 2014. [Online]. Available: https://www.europol.europa.eu/content/internet-organised-crime-threatassesment-iocta.

[12] S. Blackmore, "The meme machine". United Kingdom: Oxford University Press, 1999.

[13] I. Mann, "Hacking the human: Social engineering techniques and security countermeasures". Aldershot, Hants, England: Ashgate Publishing, 2009.

[14] A. Algarni, Y. Xu, T. Chan, and Y.-C. Tian, "Social engineering in social networking sites: Affect-based model," 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013), pp. 508–515, Dec. 2013

[15] G. Farrell, K. Clark, D. Ellingworth, and K. Pease "Of targets and supertargets: a routine activity theory of high crime rates", *Internet Journal of Criminology (IJC)*, Mar. 2005.

[16] A. Bermingham, M. Conway, L. McInerney, N. O'Hare, and A. F. Smeaton, "Combining social network analysis and sentiment analysis to explore the potential for online Radicalisation," *International Conference on Advances in Social Network Analysis and Mining*, Jul. 2009.

[17] M. Huber, S. Kowalski, M. Nohlberg, and S. Tjoa, "Towards automating social engineering using social networking sites," *International Conference on Computational Science and Engineering*, 2009.

[18] Anti-Phishing Working Group (APWG), "Phishing activity trends report unifying the global response to Cybercrime", Oct. 3, 2016. [Online]. Available: http://docs.apwg.org/reports/apwg_trends_report_q2_2016.pdf

[19] Y. Ilyin, "What is "whaling", and what's the difference from phishing", Kaspersky Lab, January 6, 2016. [Online]. Available: https://business.kaspersky.com/whaling/5009/

[20] S. Pontiroli, "Social Engineering, Hacking The Human OS," in Kaspersky Blog, 2013. [Online]. Available: https://blog.kaspersky.com/social-engineering-hacking-the-human-os.

[21] C. Nachreiner, "Signature antivirus' dirty little secret," in HelpNet Security, 2015. [Online]. Available: http://www.net-security.org/article.php?id=2239&p=2.

[22] M. Valori, G. Pravettoni, C. Lucchiari and E. Frumento, "Cognitive approach for social engineering," Wien, 2010 [Online]. Available: https://deepsec.net/docs/Slides/2010/DeepSec_2010_Cognitive_approach_for_Social_Engineering.pdf.

[23] E. Frumento and R. Puricelli, "An innovative and comprehensive framework for Social Vulnerability Assessment," Magdeburger Journal zur Sicherheitsforschung, Proceedings, 2014

[24] J. Spaulding, S. Upadhyaya, A. Mohaisen, The Landscape of Domain Name Typosquatting: Techniques and Countermeasures, arXiv Pre-Print, arXiv:1603.02767, 2016.

[25] T. Berners-Lee, "The next web," TED Talks, 2009. [Online]. Available: http://www.ted.com/talks/tim_berners_lee_on_the_next_web?nolanguage=us.