

A novel Self-Organizing Network solution towards Crypto-ransomware Mitigation

Marco Antonio Sotelo Monge, Jorge Maestre Vidal, Luis Javier García Villalba

Department of Software Engineering and Artificial Intelligence

School of Computer Science, Complutense University of Madrid

C/ Prof. José García Santesmases, 9, Ciudad Universitaria, 28040, Madrid, Spain

{masotelo,jmaestre}@ucm.es,javiervg@fdi.ucm.es

ABSTRACT

In the last decade, crypto-ransomware evolved from a family of malicious software with scarce repercussion in the research community, to a sophisticated and highly effective intrusion method positioned in the spotlight of the main organizations for cyberdefense. Its *modus operandi* is characterized by fetching the assets to be blocked, their encryption, and triggering an extortion process that leads the victim to pay for the key that allows their recovery. This paper reviews the evolution of crypto-ransomware focusing on the implication of the different advances in communication technologies that empowered its popularization. In addition, a novel defensive approach based on the Self-Organizing Network paradigm and the emergent communication technologies (e.g. Software-Defined Networking, Network Function Virtualization, Cloud Computing, etc.) is proposed. They enhance the orchestration of smart defensive deployments that adapt to the status of the monitoring environment and facilitate the adoption of previously defined risk management policies. In this way it is possible to efficiently coordinate the efforts of sensors and actuators distributed throughout the protected environment without supervision by human operators, resulting in greater protection with increased viability

CCS CONCEPTS

• **Security and privacy** → Network Security; Malware and its mitigation; • **Networks** → Network management;

KEYWORDS

Cryptovirology, network function virtualization, ransomware, self-organizing networks, software-defined networking

1 INTRODUCTION

The communication technologies have evolved towards autonomous and intelligent management models aiming to address the challenges posed by the emerging landscape [24]. In this context, the new generation networks (5G) aspire to achieve Key Performance

Indicators (KPIs) much more ambitious to those of the predecessor mobile technologies [34]. They are grounded on a strong synergy of emerging technologies, among them Software-Defined Networking (SDN), Artificial Intelligence (AI), Cloud Computing, Network Function Virtualization (NFV), Self-Organizing Networks (SON), etc. which response and action capabilities adequate to deal with a vast variety of use cases [1], as is the case of the recent trends towards cybersecurity and the novel risks these technologies raise [2, 49]. But although the 5G networks still under development, several research projects have focused on the proposal of advanced management models applied to use cases that benefit from their self-organization capabilities. A clear example is illustrated in SELFNET [38], where an autonomous management framework that brings self-organization capabilities is proposed. SELFNET (H2020-ICT-2014-2/671672) undertakes three different use cases: Self-Protection, Self-Healing and Self-Optimization [25]; where Self-Protection faces the difficulties inherent to the management of defensive solutions and the auto-response to security threats, in particular, against those related with botnets detection and DDoS mitigation.

In parallel with the forthcoming communication solutions, the network threats and their propagation abilities have evolved from predictable and traceable processes to complex schemes of infection, propagation and asset deletion. From them, the object of study of our research is the malicious software that takes advantage of cryptographic methods, with nowadays embraces from basic cryptomalware specimens to standing out strains of crypto-ransomware, the latter being specially targeted for their impact by the different organization for information security. They mostly agreed that the principal target of crypto-ransomware is the human being that manages the compromised system, which usually gives in to blackmail in order to recover part of the compromised elements [5]. In the recent years, the research community explored possible solutions attempting to mitigate the emerging ransomware threat. This is the case of the project RAMSES (H2020-FCT-04-2015/700326) [37], which approaches the design and implementation of a platform that facilitates law enforcement digital forensic investigations related to malware with financial motivation, thus outlining a feasible defense against ransomware.

In view to encourage the definition of feasible defensive solutions, our research work describes a novel approach for crypto-ransomware mitigation. It brings together some of self-organizing capabilities supported by SELFNET and some of the defensive solutions framed in the project RAMSES. The proposal drives the smart management of network-based countermeasures against crypto-ransomware without human supervision, as well as the adaptation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ARES 2018, August 27–30, 2018, Hamburg, Germany

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6448-5/18/08...\$15.00

<https://doi.org/10.1145/3230833.3233249>

of the defensive deployment to the state of the monitored environment in a context of new generation networks. This is possible by instantiating/removing sensors and actuators according to their effectiveness and the risk level of the region they operate; and by establishing regions of quarantine for stronger monitoring and actuation. This restricts the deployment of the costliest countermeasures where they are strictly necessary, in this manner enhancing the user quality of service/experience and their operational/capital expenditures.

The rest of the paper is divided into four sections, being the first of the present introduction. Section II reviews the landscape of crypto-ransomware and the most promising solutions proposed by the research community. Section III discusses the design principles of the performed research. Section IV introduces a novel self-organizing approach for crypto-ransomware mitigation, and Section V summarizes the achieved conclusion and new research steps.

2 BACKGROUND

The following describes the origins, characterization and evolution of cryptomalware; and the different countermeasures posed by the research community, highlighting among them the role of both local-based and network-based detectors as triggering of most of the mitigation approaches.

2.1 Evolution of crypto-ransomware

The research area of cryptovirology was introduced by Young et al. in 1996 [56], where it is defined as the study of cryptoviruses, which *modus operandi* aims on encrypting information resident within the victim system, so that if there are no updated backups, only the attacker is able to recover it (nowadays the software with similar features is referred as cryptomalware). This definition excludes encryption mechanisms with another purpose, for example, those frequently used for malware obfuscation [47]. In the 1980s the first cryptomalware specimens were registered, highlighting among them the AIDS Trojan, being for several researchers the first known cryptoviral specimen [35]. This was a logical bomb that after 90 reboots, encrypted the contents of the victim system demanding a payment of \$189 prior to their release. For many authors, the AIDS Trojan was also the first example of ransomware (acronym composed of ransom and malware). Note that the term ransomware popularized much later to refer cryptomalware characterized by blocking assets or functionalities of the victim system, and only returning them after a ransom is paid. This is a particular type of Denial of Resource (DoR) attack that requires a ransom in exchange for recovering the compromised resources [4], hence included into the scareware family. The ransomware typically intends to intimidate the victims in order to give them to the attacker extortion [6]. Therefore, and as Andronino et al. [5] indicated, unlike conventional attacks, the ransomware is aimed at the human being behind the victim system, which must surrender to the blackmail. Consequently, it has evolved towards the exploitation of the human factor [32] through the adaptation of the rescue payment to local circumstances (e.g. varying the prize based on geographical region, value of denied assets, etc.), the elaboration of more threatening extortion notifications, and even providing tutorials

and guidelines for payment making [31]. For some researchers the most challenging part of this intrusion process is to be able to anonymously retrieve the ransoms paid, for which the ransomware has evolved from conventional bank transfers, to the implementation of cryptocurrencies and anonymous networks, which make its traceback almost impossible [57]. This threat is usually classified based on functionality, where a pair of major groups of strains is differentiated: the crypto-ransomware family, which is focused on the encryption and rescue of assets (as a subfamily of cryptomware with scareware behavior); and the locker-ransomware family, focused on the denial of access to the system, for which it is not necessary to resort to cryptographic processes [7].

The most relevant ransomware specimens registered until 2008 were reviewed in [35], and the strains observed between 2006-2014 were studied in [12]. It should be noted that at the date these publications were released, the ransomware was not considered a large-scale menace by their authors nor most of the research community, since the analyzed encryption processes were reversible, easily detectable and the payments were traceable. Brewer described their distribution strategy as massive [6], where the attacker attempted to indiscriminately infect victims with the purpose of maximizing the chances of exploiting vulnerabilities and obtaining ransoms. But with the discovery of the first specimens that applied asymmetric encryption mechanisms (e.g. Cryptolocker, TorrentLocker, Cryptowall, etc.), their effectiveness drastically improved [22]. In recent strains the private keys are typically generated in remote Command and Control (C&C) servers, and the victims only receive the private key upon payment, which makes the ciphering of the asset difficult to reverse (most of the researchers directly adopt the assumption that it is impossible). According to Gazet [35], their *modus operandi* is similar, distinguishing three fundamental stages: 1) identification of the assets to be blocked; 2) elimination of the victim access to the compromised assets (in crypto, usually by encryption and deletion of backups), and 3) extortion, through which the victim pays the attacker in exchange for their recovery.

2.2 Countermeasures

The defenses against crypto-ransomware have traditionally been approached from the perspective of the conventional malware countermeasures, where in the last decades an enormous amount of contributions was published [27]. But despite their wide adaptation, some experts are skeptical about their effectiveness against current specimens. For example, Continella et al. [11] warned that purely detection-based approaches were not enough, as they usually address the problem of discovering encryption processes once initiated; but by then, part of the assets are already compromised. Bearing this in mind, in publications like [4] the need for their early detection is reviewed, as well as strategies for mitigating new specimens prior to their launch are introduced. This fact has led Gharib et al. [36] to separate the defensive deployments into two approaches, that entail very different paradigms: those based on offline detection, which is carried out in safe and isolated environments, and not at real time; and those based on online detection, which focuses on identifying the first stages of the attacks by monitoring the real scenarios on which they are acting, prior to the

deployment of mitigation measures. They can act at local-level and network-level as described below.

When the defense operates at local-level, the identification of threats typically lays on analyzing features of protected end-point, like memory, processor consumption or write/read operations on possible assets of the protected device [41, 48]. These tasks are supported by strong machine learning solutions [26] and sandboxing [40]. The entropy-based malware analysis is frequent in the bibliography, which allows to distinguish files with encrypted content from the original assets [44]. In [14] the crypto-ransomware is detected by recognizing of strategies for asset discovery, which try to enumerate specific file extensions (e.g. jpeg, xcl., docx, etc.). The reversion of the ciphering processes is studied in [15]. In [5] both software and network traffic, are analyzed looking for extortion messages, hence requiring adopting computer vision and natural language processing algorithms. Proposals like [16, 45] applied Model Checking techniques for harmful code discovery. It is important to emphasize that as pointed out in [43], for the sensors/actuators being effective, they must consider indicators as varied as possible. On the other hand, if the detection is not conducted on isolated environments, it could involve the loss of the first compromised files, which leads to previously stating how many initial losses are acceptable [23], and under what terms or conditions. Therefore, and in order to offer a higher level of protection, active defenses may be deployed, as is the case of of honeypot [46] or decoys [12].

Recently, and in parallel with the development of local-level solutions, the research community studied the impact of the crypto-ransomware on communication processes. This prompted the publication of the first proposals based on analyzing network features in emerging scenarios, as is the case of [10, 58] at Internet of Things (IoT) or [42] at Cloud Computing. As indicated by Cabaj et al. [8], the list of IP addresses with which each ransomware specimen tries to communicate with C&C server tends to be similar with those of previous detected threats. Because of this, blocking its access and identifying which potential compromised systems are attempting to communicate with them, are effective methods to recognize its spreading. Their research also led to another important finding: each ransomware family often lies on certain custom handshaking protocol that allows exchanging information between the compromise nodes and the C&C server [7]. During this trading, the public key of the asymmetric encryption algorithms is sent to the ciphering agents; therefore, if this process is interrupted in time, the initialization of the asset encryption stage is avoided.

3 DESIGN PRINCIPLES

This section describes a framework for self-organized defensive deployments that aims on mitigating the impact of cryptoviral-based ransomware, as well as preventing their spreading. With the purpose of facilitating its understanding, the adopted objectives, design principles and architecture are described.

3.1 Objectives

As reviewed in the preview section, the dissection and study of cryptomalware can be addressed from different perspectives. The same applies to the development of detection, mitigation and prevention

measures, which can be adapted from classic schemes inspired by traditional countermeasures [27], to strategies aiming on its behavior inherent to diverse contexts, for example local [48] or network [7] monitoring environments. Therefore, the following three main goals were raised:

- The recognition of crypto-ransomware activities at a network scenario with a view to attempt their interruption while minimizing their impact.
- The partial or total prevention of their replication and propagation throughout the protected environment.
- The introduction of the Self-Protection capabilities provided by Self-Organizing Networking as a promising ransomware mitigation paradigm.

On this basis it is easy to deduce that the object of study was the defense against cryptoviruses for extortion purposes at network environments by implementing Self-Organizing capabilities; in particular those related with Self-Protection. The rest of this subsection delves into the scope of the performed research, outlining the problem to be solved, and establishing the requirements and limitations of the proposed solution.

3.2 Assumptions

The following assumptions define the characteristics of the threat to be mitigated and its impact on the protected system.

Crypto-ransomware life-cycle. The crypto-ransomware life-cycle typically consists of five stages: infection, spreading, asset fetching, blocking and extortion [35]. At the infection stage the malware is settled into the victim system. It is assumed that the self-replication and propagation tasks may occur anytime (even simultaneously with other stages).

Asset fetching unpredictability. The asset fetching stage traverses the victim system looking for files of certain predefined extensions (e.g. jpeg, tex, java, docx, etc.) [14]. It is assumed that the search strategies behave unpredictably [46], as is observed in recent specimens, so it is not possible to know if decoy files would be processed before the rest (i.e. tricks such as renaming directories with expressions like "AAA" are not effective anymore).

Asset fetching complexity. It is assumed that the seek for assets and backup files conducted by crypto-ransomware entails a period of time ranging from few minutes/seconds in local sceneries to hours in heterogeneous network environments [38]. The computational complexity of these tasks varies depending on the search algorithms/libraries, the amount of information stored, the depth of the directory trees and the density of terminal nodes [18]. Additionally, the encryption time depends, among other factors, on the size of the files that host the assets.

Asymmetric encryption. It is assumed that the cryptoviruses to be analyzed implement asymmetric encryption methods, where the private keys are generated remotely by C&C servers, hence only being transferred public keys to the compromised systems. Note that alternative approaches tend to be easily reversible, and therefore do not often attract the research community attention [9][12].

Custom handshaking protocols. As stated by Cabaj et al. [7, 8], it is assumed that the linkages between compromised systems and their C&C servers are conducted through custom, characteristic

and detectable handshaking protocols. In addition, it is assumed that the lists of network addresses for contacting C&C servers are practically the same between instances of the same strain.

Infection vector payload. Ransomware, like other types of malicious software, tend to reuse code from previous versions or similar families. To improve their effectiveness, developers often include additional functionalities, for example those related with creating secondary backdoors, wiping capabilities or extend the default remote control features. In the performed research, the impact of this additional payload is ignored. Therefore, the study focuses on actions related to cryptographic processes. Henceforth it is assumed that if the public key is prevented from reaching the victim, the initializations of the file locking processes are avoided [7].

Asset lost. Once the encryption stage began, it is assumed that there will be losses, regardless of whether it has been completed successfully or not. It is also assumed that the blocked files are irrecuperable.

Ransom messages. The main target of crypto-ransomware are the human beings that manage the compromised systems. Consequently, it is a fact that cryptography-based infections become visible once the encryption process is completed, at the extortion stage [5]. Note that for hindering detection, some specimens do not obtain the extortion messages/images from the C&C server until asset lockout is complete.

On this basis, four essential facts have been posed, which lay the ground of the proposed mitigation/prevention strategy: 1) the search for assets and backup files is a noisy characteristic of the crypto-ransomware; 2) if the arrival of the public key is prevented, the asset encryption stage is interrupted; 3) once the cyphering started, there will be losses; and 4) the handshake between the victim system and the C&C server is discernible.

3.3 Requirements

The requirements that have modeled the scope, design and implementation of the proposed Self-Organizing scheme, are detailed below.

Accomplice main goals. The proposed system must be able to contribute to the mitigation of crypto-ransomware infections, as well as preventing its propagation on communication networks. The proposed solution must adapt the foundations of Self-Organized Networks.

Legitimate activity prioritization. The deployed methods should minimize the impact of the countermeasures in the legitimate usage of the protected environment, hence avoiding actions like completely isolating the compromised nodes.

Impact minimization. The Self-Organizational defense should aim to regulate the impact of sensors and actuators in terms of Quality of Service (QoS), Quality of Experience (QoE), CAPital EXpenditures (Capex) and OPerating EXpense (Opex). It is assumed that the common is that the higher the cost, the more effective the deployment of actuators. This is because it allows a deeper analysis of the protected environment.

Perimeters and quarantine. The deployment of actuators must facilitate the definition of security perimeters and quarantine regions. In this way it is possible to dynamically adapt the cost of

the defensive mechanisms to the risk level at each region of the protected environment.

Reports. In order to facilitate forensics and detect misconfigurations, the results of both analysis and countermeasures, may be stored. Additionally, it is advisable to use standards that facilitate the incident notification to information security authorities and organizations.

3.4 Constraints

In order to facilitate the understanding of the performed research, the following topics have been overlooked or discussed in less detail.

Communication channel protection. Despite the existence of eavesdropping and identity theft methods capable of exploiting SON vulnerabilities [19], the performed research does not delve into the security of the communication channels between defensive agents.

Alert correlation. Methods capable of correlating incidents observed at different regions of the protected environment are not proposed nor deployed [54]. Neither have their report to information security authorities has been considered.

Adversarial attacks. At present, there is a large number of evasion methods and malware obfuscation strategies [17] (most of them based on imitating legitimate behaviors [52, 53]), able to thwart intrusion detection systems. The performed research does not deepen into their identification nor mitigation.

Privacy and data protection. Respecting privacy and data protection is essential to safeguard the information society. However, the use of Privacy-Enhancing Technologies (PET) [3] often makes difficult monitoring, analyzing and identifying the source of threats. At the performed research, neither the adoption of these technologies nor their impact on the defensive deployment have been studied. On the other hand, the proposal deploys active security capabilities. The drawbacks of these measures in terms of privacy and data protection are well-known, but the performed research does not delve into their impact nor mitigation [9, 13].

Data persistence. Although the situational knowledge acquired by the various actors of the defensive deployment is conveniently stored for facilitating forensic, it is beyond the scope of this paper to go into detail about the data models and storage technologies considered.

3.5 Architecture

The architecture of the proposal is illustrated in Fig. 1. It is aligned with the design principles established in 5G [38] and the ETSI NFV architecture [30], where the separation of the data and control planes and the orchestration capabilities for Virtual Network Function (VNFs) management play an essential role in the automatic deployment of defensive countermeasures. This self-organized implementation is strongly supported by sensing and actuation elements that are deployed in different levels of the network architecture to conduct a system based on the closed-loop paradigm.

The proposed framework relies on a distributed scheme where the monitoring and enforcing of actions embraces the user-endpoints and the instantiated VNFs. It leads to the design of a defensive strategy that starts at the protected devices since they incorporate lightweight sensors and actuators for applying countermeasures

whenever suspicious activities are detected at operating-system level. Solely local analysis is limited due to the lack of network-wide visibility, hence arising the importance of the upper levels of the architecture.

The NFV Infrastructure (NFVI) poses an important enabler towards the deploying of VNFs at different locations of the network. The NFVI is built upon the combination of compute, storage and networking resources managed by the virtualized Infrastructure Manager (VIM). Both components expose high-level programming interfaces (API), thus being easily configurable. Since this level of virtual abstraction is inherent in 5G architectures, orchestration processes are feasible through software entities that can manage VNF life-cycle from creation to removal. Hence, in the proposed mitigation framework sensors and actuators are deployed as VNFs by the Orchestration Layer, and their automatic instantiation guarantees the reinforcement of defensive measures in potential risk situations. This is triggered by the VNF manager whenever defensive actions demand the creation of new VNFs or, on the contrary, when their removal is decided. The latter case occurs when the protection level has decreased because the threats that trigger their deployment are considered extinct/mitigated, or they became counterproductive. The Orchestration Layer includes also the Host-App Manager, capable to enforce countermeasures on the protected devices, for which is important that host-level sensors and actuators implement the proper interfaces to allow the exchange of information.

Note that both VNF sensors and actuators are controlled by elements capable of configuring their behavior according to the defensive scheme deployed. With this purpose the Control Layer brings together both SON and SDN controllers that are responsible for exchanging their configuration directions through software interfaces. SON controllers can accommodate the behavior of VNFs according to the observed operational context, for instance, by refining sensing attributes or by updating configuration rules in an actuator. Likewise, SDN controllers allow the centralized traffic management by their interaction with forwarding elements through southbound interfaces such as the OpenFlow protocol. Because of this, SDN applications can deploy traffic isolation, redirection policies or dynamic flow-based routing directives to conduct the defensive strategy.

At management level, the Autonomous Detection and Mitigation Layer is composed of two main modules: Analysis and Incident Management. The Analysis component processes the data extracted by host and network-level sensors/actuators. With this purpose, it performs feature extraction and pattern recognition looking for discordant traits that point out anomalous behaviors that trigger more complex knowledge inference capabilities [50]. The analytic tasks allow the proposed framework to deduce potential risk situations, which are represented as alerts to be processed by the Incident Management module. This component coordinates the effective deployment of defensive countermeasures, which is sustained by three stages: event correlation, decision-making and NFV orchestration. The first of them aims on aggregating and filtering similar alerts, and then conducting their correlation with other security incidents registered thorough the protected network. At the decision-making step the acquired situational awareness is taken into account for determining countermeasures according to the risk level, thus ensuring that their computational cost is proportional to the expected problems. Finally, the decided defensive strategy is

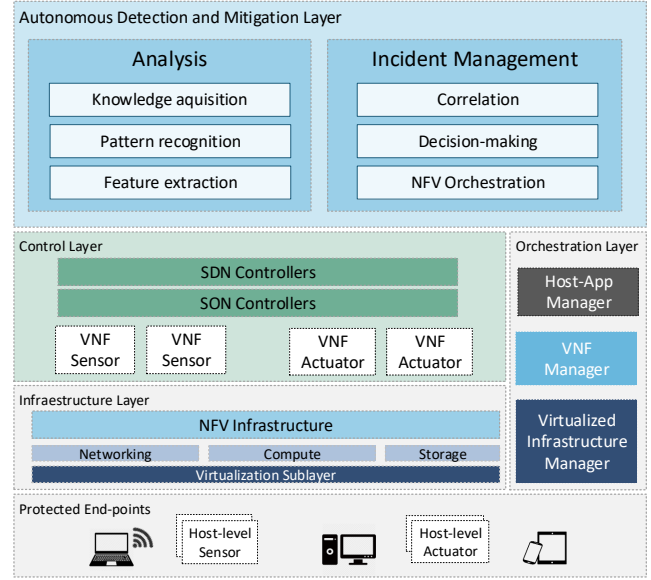


Figure 1: Architecture

automatically deployed driven by the NFV capabilities that contemplate the management of the life-cycle of the instanced elements, from their creation to their elimination. The self-organized behavior is therefore achieved through the interaction of the defensive framework components at each architectural level, hence supporting the strategy for ransomware mitigation described in the forthcoming section.

4 SON FRAMEWORK FOR CRYPTOMALWARE MITIGATION

Based on the design principles described above, the proposed self-organizing scheme introduces a closed-loop model that facilitates the adaptation of the defensive deployment to the monitoring environment without the direct intervention of operators. According to the standardized framework for SON 3GPP networks, this allows reducing the operational costs [29], hence streamlining decision-making and counteracting. As highlighted in [39], even the simplest self-organized networks govern their behavior based on information monitored by sensors. Therefore, the proposed defense against crypto-ransomware has as starting point data collected by sensors scattered throughout the protected environment. They pose different nature and display diversified targets, as is explained in detail at the following subsections. In order to generate high level metrics and identify traits of suspicious behaviors, the monitored information is aggregated and analyzed at a different data processing plane. Note that with the purpose of minimizing the impact of the sensor at the end-points (where it is not always possible to guarantee vertical/horizontal scaling capabilities), both aggregation and analysis tasks are launch on dedicated servers. From the acquired situational knowledge it is decided when to deploy/recall the additional security measures, which are executed by actuators defined as agents in charge of closing the loop by enforcing security policies.

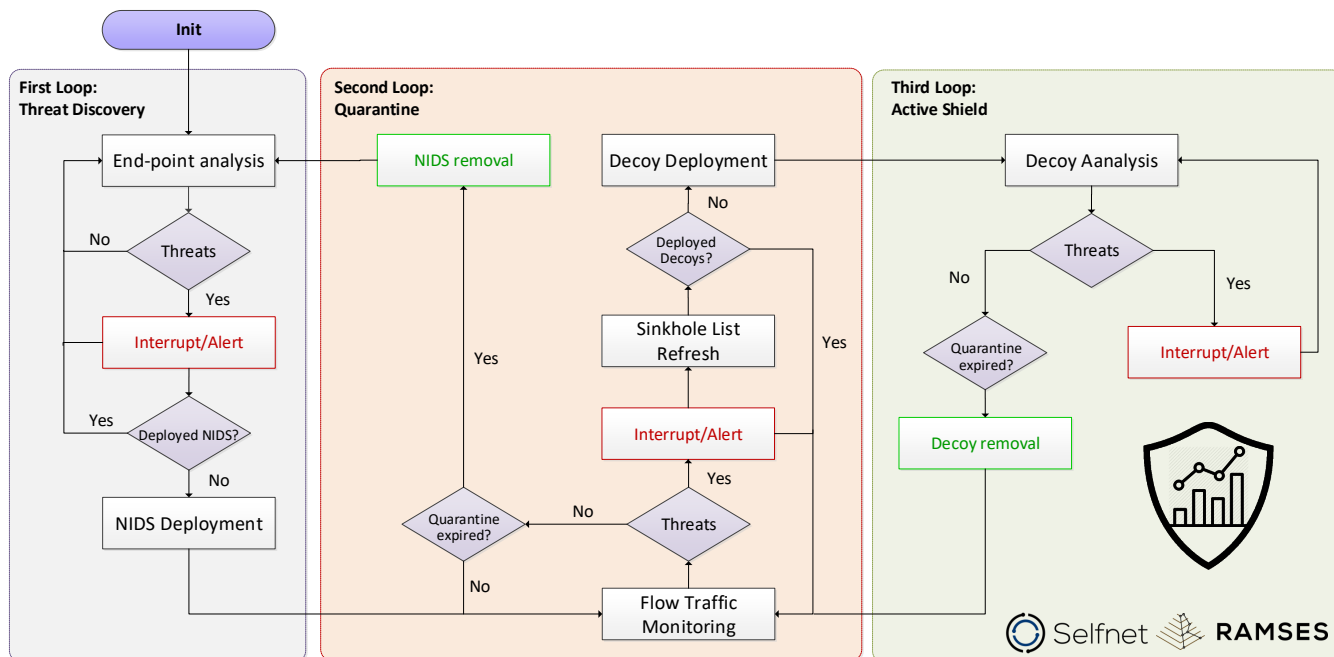


Table 1: Closed-loops for crypto-ransomware defense

The algorithm that manages the situational awareness and dictates the immune responses per network region is illustrated in Fig. 2. There three intelligent loops are described, each of them having a different purpose (see Table 1). The first closed loop is termed Threat Discovery and focuses on recognizing discordant behaviors typical of cryptomalware at the end-points [40]. This process runs continuously and serves as triggering for the most basic network-level actuations, which entail the deployment of additional monitoring elements at regions suspected of having been compromised [7]. During the second closed loop, referred as Quarantine, they perform as sensors that aim to manage the passive security countermeasures deployed at the affected regions (areas in quarantine), hence deciding how long they should remain operational, their configuration, and if it is required to increase their restriction level. Therefore, each iteration of the second loop has three possible consequences: leaving the defensive deployment as

In addition to satisfying the principal objectives of the performed research, through the proposed self-organization strategy the requirements defined in previous sections are met. For example, the impact on the legitimate activities is minimized by deploying defensive measures proportional to the risk level per region. In this

way, the resource consumption is enhanced, and the computational costs related with maintaining the security elements are reduced. On the other hand, the security actuators may vary from simple monitorization tools at the end-points, to active security measures dispersed along the network, the latter being much more efficient, but implying higher costs and administrative repercussions [13]. The action of the most expensive defensive agents is restricted by quarantine periods, which define the quarantine regions. Finally, the detection of suspicious events is carried out at different network layers, which are not only reported to the Network Function Virtualization Management and Orchestration (NFV M&O) services [28] in charge of executing the Self-Organizing algorithm, but also are also stored in a database with the purpose of facilitating forensics. The following defines in detail each defensive closed-loop.

4.1 First closed-loop: Threat Discovery

Threat Discovery is the frontline of the proposed self-organizing scheme. Therefore, it continuously operates looking for traits of suspicious activities. Fig. 3 illustrates the different stages of this first defensive loop, where the main source of knowledge is the data reported by sensors deployed at the end-points, which monitor and extract local activities as dynamic features on the user behavior. In particular, they aim on gathering information directly related with infections, asset fetching, propagation and asset ciphering as sequences of system calls registered in the protected systems, and report them to a dedicated server for analysis (step 1).

The sensors of the first defensive loop operate at two different layers: first, the new applications/updates that demand being installed at the monitored end-points are executed on isolated and safe environments (Sandboxes), in a way that in case of disguising malicious contents, they are prevented of jeopardizing the victim system [40]; hence the infection risk decreases. On the other hand, the sensors continuously collect information about the operative system processes in execution. This is directly performed at the protected end-points by Host-based Intrusion Detection Systems (HIDS), so when potential threats are detected, they are already acting at the victim system [23]. Note that depending on the maturity level of the discovered intrusions, it is possible to mitigate the infections, avoid their propagation or interrupt partially/totally asset hijacking.

In step 2, the detected situations are reported to the Incident Manager, which, according to the security policies, and in accordance with the algorithm described in Fig. 2, decides the countermeasures to be applied. Those are reported to the Orchestration Layer and vary depending on the level of data processing (step 3): at local-level the interruption of the processes related to possible intrusions are notified from the Host-App Manager; in the case of being detected inside the sandboxes, this action directly prevents the malware from reaching the protected end-points. At network-level, if no previous actions have been taken, additional network-based security elements are instantiated by the VNF Manager (step 4). They provide Deep Packet Inspection (DPI) solutions that are capable of gathering the flow-level information required for identifying custom handshakes between potentially compromised end-points and C&C servers. The life span of the reinforcements is regulated

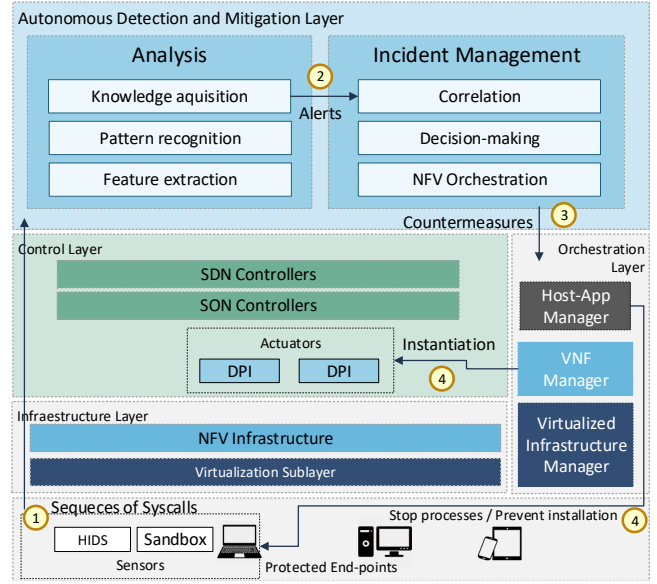


Figure 3: Threat Discovery closed-loop

by the second defensive loop, and the areas on which they act are refereed as quarantine regions.

The main advantages of the Threat Discovery loop are its local prevention and mitigation capabilities. Since the sensors operate at the end-points, their management is simple, and poses very little impact on the network, which is only related with the communications between sensors and a dedicated server with analytic purpose. However, since network metrics are not taken into account, it is difficult to detect activities involved in malware spreading or preventing the initiation of asset encryption without losses at the systems that have already been infected. Therefore, in order to improve these disadvantages, when signs of intrusive behaviors are detected a second defensive loop is activated.

4.2 Second closed-loop: Quarantine

Quarantine orchestrates the passive defensive deployment at network-level (see Fig. 4). It is activated by the first closed loop, which is responsible for triggering the instantiation of the NFV with DPI capabilities that act as sensors in the second defensive loop. The purpose of these technologies is to monitor information that allows detecting public/private key exchanges in custom handshakes between compromised end-points and malicious C&C servers. Hence metrics related with traffic flows are transferred to the analytic dedicated server (step 1), which executes the required pattern recognition tasks. In this way it is possible to detect the begin of processes involved in asset blocking and proceed with their interruption. This intervention lies on the thesis formulated by Cabaj et al. [7] where it was proven that at most of the crypto-ransomware specimens that implement asymmetric encryption, the ciphering procedures do not begin until the malicious agent that infected the end-point associates a unique identifier to the victim system, to which the malicious C&C servers generate a private/public pair of keys. The public key is necessary for initiate the encryption stage, and it is

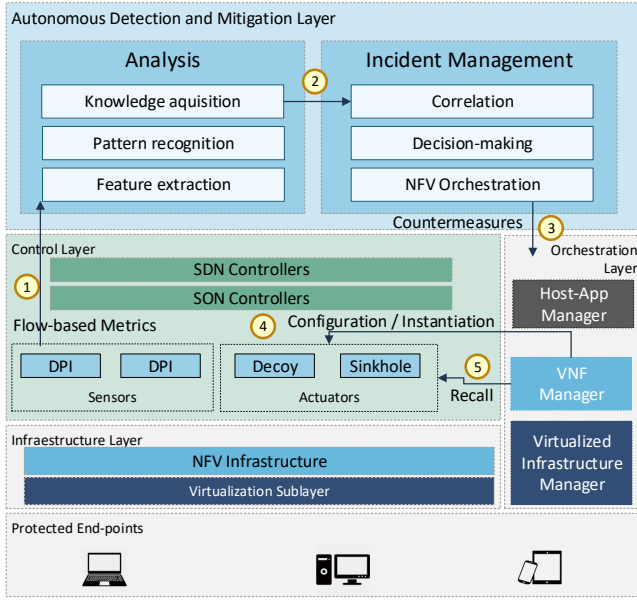


Figure 4: Quarantine closed-loop

transmitted during a custom handshake; the private key is necessary to recover the stolen information, being the element that the attacker tries to "sell" to the victims in exchange for the ransom. Therefore, if the packets with the payload that carries the public keys are discarded, the start of the blocking of assets is interrupted (although the malware retains self-replication and propagation features). Note that as indicated in [12, 35], the crypto-ransomware strains that do not proceed in a similar way, are not considered massive threats, since their ciphering typically can be jeopardized by reverse engineering, exploitation of vulnerabilities at encryption algorithms/libraries or brute force. In this case, the intrusion is usually locally mitigated.

Although DPI analysis provides a significant level of defense, its maintenance entails a high computational cost, that in the worst cases forces to reduce the network bandwidth in order to process packets at real time [33]. Therefore, it should be limited to situations where relevant risk is observed. Consequently, when within a reasonable quarantine period there are not visible replications of the intrusion at the end-points nor in the flow-level data reported by the DPI sensors, the network defenses are dismantled by the VNF Manager (step 5). This retreating action poses analogy with the programmed death processes (apoptosis) typical of the immune systems in nature, where after removing the infection, most of the immune agents deployed to cope with it are eliminated [55].

On the other hand, if the information reported by DPI leads to discover the presence of custom handshakes (step 2), the hypothesis that the protected environment has been compromised is reaffirmed. Therefore, spreading attempts and new infections are expected. In this situation the system acts transmitting data about nodes involved in the suspicious handshakes to the Orchestration component (step 3), which invokes the VNF Manager for managing the traffic that flows from the aforementioned nodes to a sinkhole

server (step 4). In the meantime, the VNF Manager proceeds to the instantiation of a decoy server (if it has not been previously instantiated) [12], which aims on confirming the presence of the intruders, delaying their progress and gathering information from their modus operandi (step 4). Since both solutions represent active security actuations [9], their deployment implies a cost greater than the passive DPI monitorization. It is justified because at this point, the certainty that the threat is real is high. Their actuation is managed by the third closed loop, which is automatically initiated once they are deployed.

In general terms the second defensive loop has a higher cost than the first, because it requires the instantiation of additional NFVs in potentially compromised regions, that implies a greater impact on QoE/QoS. On the other hand, it is not able to avoid the internal propagation of the infections, for which the activation of the third defensive loop is required. The main advantages of Quarantine are that it facilitates the identification of replicas of the detected intrusions and that it is able to prevent asset ciphering from its beginning, even without knowing a priori what end-points have been compromised, in this way constituting an important passive security solution.

4.3 Third closed-loop: Active Shield

The main goals of Active Shield are the management, deployment and configuration of active security countermeasures previously triggered by the second defensive loop (see Fig.5). With this purpose, the information provided by two kinds of sensors is analyzed (step 1), which also behave as actuators. They are sinkhole and decoy servers. Each time they accomplish a defensive action, the results are reported to the Incident Management component with the information necessary for situational awareness, forensics and generating redirection directives for sinkhole server configuration (step 2). The range of action of these countermeasures is the second quarantine region, which is initiated/revalidated upon detection of custom handshakes exchanging identifiers and public keys. As in the second closed loop, this quarantine level is maintained until the expiration of certain period of time where network monitors do not collect signs of intrusive behaviors. If deactivated, the first quarantine region remains operative. The apoptosis task involves removal of the instantiated decoy servers and discarding the rules for redirecting traffic from suspicious nodes to sinkhole servers that were added during quarantine (step 3). Therefore, only NFVs with decoy capabilities are instantiated, because the sinkhole servers remain operational throughout the protection process (even if there are not quarantines); in the latter case, default redirection rules are considered.

At this defensive stage, the decoy servers play two essential roles aiming on prevent the infection spreading. Firstly, they allow detecting malicious behaviors at local level, hence incorporating the same monitoring techniques applied at Threat Discovery, and additionally providing advanced functionalities toward determining the modus operandi of the intruder (e.g. analysis of file accesses/modifications, permission requests, etc.), which involve a major computational cost at the original end-points. Note that the new NFV may implement their own analytical engine, so the use of a dedicated server for this purpose is optional. To

maximize effectiveness, they must be deployed near potentially compromised end-points [21] that have been previously identified by the first/second defensive loop; or they may receive redirected traffic from a sinkhole server. Therefore, they must implement the same propagation via (usually refereed as lateral movement) than the protected systems: Server Message Block (SMB), SMTP and HTTP services, notification groups, remote backup services, etc. [51]. On the other hand, decoy servers slow down malware progression. With this purpose, they present a huge directory tree with a large number of files with extensions typically searched by crypto-ransomware at its asset fetching step. Due to the high computational cost involved in browsing their contents, asset fetching is delayed, hence bringing additional time for analyzing and discovering the incidences before real asset ciphering [14].

The sinkhole servers (also refereed in the bibliography as DNS sinkholes or Blackhole DNS) redirect requests from network addresses associated with suspicious C&C servers, to decoy servers not reachable by the protected systems. This action prevents custom handshakes involved at public key exchange from being accomplished. Note that as demonstrated in [8], crypto-ransomware based on asymmetric encryption usually includes a list of network addresses for contacting the C&C server with elements commonly shared between instances propagated by the same attacker. Therefore, these network addresses are recycled between different versions and families of malware. This allows the sinkhole server to manage a black list with all the possible threat sources, which is updated each time the DPI sensors from the second closed loop or the decoy servers discover novel intrusions.

Concluding, the third defensive loop allows to prevent crypto-ransomware propagation, slows down the asset fetching stages, interrupts ciphering and collects information about the modus operandi of the attacker. This is achieved through the instantiation of VNFs that act as decoy servers and by updating traffic redirection policies. However, the computational cost is higher than in actions orchestrated by the rest of defensive loops, and it adds complexity to the management tasks directly related with self-organization. Due to this, and in analogy with the second loop, they are temporary measures that after a quarantine period expire.

5 CONCLUSION

Throughout this paper the problem of cryptomalware has been reviewed, deepening into the evolution of crypto-ransomware and the reasons that places it in the line of fire of the current organizations for cyberdefense. In order to contribute to its mitigation, a novel self-organizing defensive approach has been proposed, which allows the smart coordination and calibration of countermeasures as sensors and actuators, and orchestrates their instantiation considering the acquired situational awareness of the protected environment and risk level. This has empowered by the adoption of emerging communication technologies inherent to the progress towards the development of new generation networks (5G). In particular, the proposal described a SON solution based on the closed-loop paradigm, grounded in three defensive actions: Threat Discovery, Quarantine and Active Shield. The first of them is the frontline of the proposed self-organizing scheme, that continuously operates looking for traits of suspicious activities at the end-points.

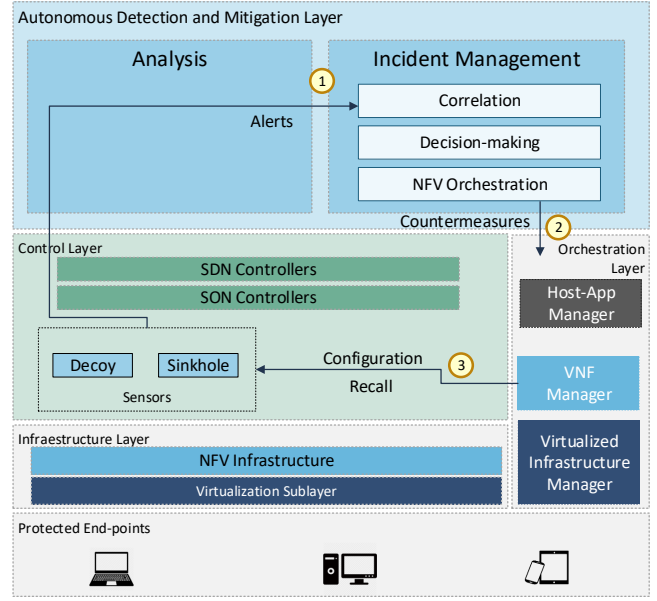


Figure 5: Active Shield closed-loop

Quarantine reinforces the default security deployment, and it is in charge of the network-level passive security. Finally, when the risk level is high, the third loop orchestrates the instantiation of active defensive solutions, as is the case of decoys or sinkhole servers.

However, although the proposal has taken into account all the assumptions and requirements described at its design principles, throughout the paper different aspects have been identified that, either due to the lack of maturity of the technologies involved, or in order to facilitate the understanding of the proposal, have not been described with sufficient detail. Because of this, to delve in them has been delegated to future research, which for example may include to implement advanced alert correlation techniques, protecting the communication channels between defensive elements, and accommodating policies that allow compliance with the recent EU General Data Protection Regulation (GDPR).

ACKNOWLEDGMENTS

This work was partially funded by the JSAN Travel Award 2018 bestowed by the MDPI Journal of Sensors and Actuator Networks (JSAN). In addition, the authors sincerely appreciate the support of the European Commission Horizon 2020 Programme under the Grant Agreements number H2020-ICT-2014-2/671672 (SELFNET: Framework for Self-Organized Network Management in Virtualized and Software Defined Networks) and H2020-FCT-04-2015/700326 (RAMSES: Internet Forensic platform for tracking the money flow of financially-motivated malware).

REFERENCES

- [1] 5G-PPP 2016. 5G PPP use cases and performance evaluation models. (2016). <https://5g-ppp.eu/white-papers/>.
- [2] 5G PPP Security Working Group 2017. 5G PPP Phase1 Security Landscape. (2017). <https://5g-ppp.eu/white-papers/>.
- [3] A. Acquisti, L. Brandimarte, and G. Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347 (2015), 509–514. Issue 6221.

- [4] B.A.S. Al-Rimy, M.A. Maarof, and S.Z.M. Shaid. 2017. A 0-Day Aware Crypto-Ransomware Early Behavioral Detection Framework. In *Proceedings of the 2nd International Conference of Reliable Information and Communication Technology: Recent Trends in Information and Communication Technology*. Johor, Malaysia, 758–766.
- [5] N. Andronino, S. Zanero, and F. Maggi. 2015. HelDroid: Dissecting and Detecting Mobile Ransomware. In *Proceedings of the 18th International Symposium on Research in Attacks, Intrusions, and Defenses*. Kyoto, Japan, 382–404.
- [6] R. Brewer. 2016. Ransomware attacks: detection, prevention and cure. *Network Security* 2016 (2016), 5–9. Issue 9.
- [7] K. Cabaj, M. Gregorczyk, and W. Mazurczyk. 2018. Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics. *Computers & Electrical Engineering* 66 (2018), 353–368.
- [8] K. Cabaj and W. Mazurczyk. 2016. Using Software-Defined Networking for Ransomware Mitigation: The Case of CryptoWall. *IEEE Network* 3 (2016), 14–20. Issue 6.
- [9] D.E. Denning. 2014. Framework and principles for active cyber defense. *Computers & Security* 40 (2014), 108–113.
- [10] A. Azmoodeh et al. 2017. Detecting crypto-ransomware in IoT networks based on energy consumption footprint. *Journal of Ambient Intelligence and Humanized Computing* (2017), 1–12. DOI: <http://dx.doi.org/10.1007/s1265>
- [11] A. Continella et al. 2016. ShieldFS: a self-healing, ransomware-aware filesystem. In *Proceedings of the 32nd Annual Conference on Computer Security Applications*. Los Angeles, CA, US, 336–347.
- [12] A. Kharraz et al. 2015. Cutting the Gordian Knot: A Look Under the Hood of Ransomware Attacks. In *Proceedings of the 12th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Milan, Italy, 3–24.
- [13] B.H. Chang et al. 2004. Active security management based on Secure Zone Cooperation. *Future Generation Computer Systems* 20 (2004), 283–293. Issue 2.
- [14] C. Zheng et al. 2016. GreatEatlon: Fast, Static Detection of Mobile Ransomware. In *Proceedings of the 12th International Conference on Security and Privacy in Communication System*. Guangzhou, China, 617–636.
- [15] E. Kolodenker et al. 2017. PayBreak: Defense Against Cryptographic Ransomware. In *Proceedings of the ACM on Asia Conference on Computer and Communications Security*. Abu Dhabi, United Arab Emirates, 599–911.
- [16] F. Mercaldo et al. 2016. Ransomware Steals Your Phone. Formal Methods Rescue It. In *Proceedings of the 36th International Federated Conference on Distributed Computing Techniques*. Heraklion, Greece.
- [17] F. Zhang et al. 2016. Adversarial Feature Selection Against Evasion Attacks. *IEEE Transactions on Cybernetics* 16 (2016), 766–777. Issue 3.
- [18] H.H. Huang et al. 2012. Just-in-Time Analytics on Large File Systems. *IEEE Trans. Comput.* 61 (2012), 1651–1664. Issue 11.
- [19] H.Y. Lateef et al. 2015. LTE-advanced self-organizing network conflicts and coordination algorithms. *IEEE Wireless Communications* 22 (2015), 108–117. Issue 3.
- [20] J. Kwon et al. 2016. PsyBoG: A scalable botnet detection method for large-scale DNS traffic. *Computer Networks* 97 (2016), 48–73.
- [21] K. Wang et al. 2017. Strategic Honeypot Game Model for Distributed Denial of Service Attacks in the Smart Grid. *IEEE Transactions on Smart Grid* 8 (2017), 2474–2482. Issue 5.
- [22] N. Kiss et al. 2016. Kharon Dataset: Android Malware under a Microscope. In *Proceedings of the LASER Workshop: Learning from Authoritative Security Experiment Results, and Vulnerability Assessment*. San Jose, CA, US, 1–12.
- [23] N. Scaife et al. 2016. CryptoLock (and Drop It): Stopping Ransomware Attacks on User Data. In *Proceedings of the 36th IEEE International Conference on Distributed Computing Systems*. Nara, Japan, 303–312.
- [24] P. Demestichas et al. 2013. 5G on the horizon: key challenges for the radio-access network. *IEEE Vehicular Technology Society* 8 (2013), 47–53.
- [25] P. Neves et al. 2016. The SELFNET Approach for Autonomic Management in an NFV/SDN Networking Paradigm. *International Journal of Distributed Sensor Networks* 12, 2 (2016), 1–14.
- [26] S. Homayoun et al. 2017. Know Abnormal, Find Evil: Frequent Pattern Mining for Ransomware Threat Hunting and Intelligence. *IEEE Transactions on Emerging Topics in Computing* (2017). DOI: <http://dx.doi.org/10.1109/TETC.2017.2756908>
- [27] Y. Ye et al. 2017. A Survey on Malware Detection Using Data Mining Techniques. *Comput. Surveys* 50, 41 (2017). Issue 3.
- [28] ETSI 2013. Network Functions Virtualisation – Update White Paper 2. (2013). https://portal.etsi.org/nfv/nfv_white_paper2.pdf.
- [29] ETSI 3GPP TS 32.500 2015. Self-Organising Networks (SON): Concepts and requirements. (2015). <https://www.3gpp.org/DynaReport/32500.htm>.
- [30] ETSI NFV ISG 2014. Network Functions Virtualization (NFV) Management and Orchestration. (2014). ETSI GS NFV-MAN 001 V1.1.1.
- [31] C. Everett. 2016. Ransomware: to pay or not to pay? *Computer Fraud & Security* 2016 (2016), 8–12. Issue 4.
- [32] M. Fimin. 2017. Are employees part of the ransomware problem? *Computer Fraud & Security* 2017 (2017), 15–17. Issue 8.
- [33] L.J. Garcia Villalba, A.L. Sandoval Orozco, and J. Maestre Vidal. 2017. Advanced Payload Analyzer Preprocessor. *Future Generation Computer Systems* 76 (2017), 474–485.
- [34] L. Gavrilovska, V. Rakovic, and V. Atanasovski. 2015. Visions Towards 5G: Technical Requirements and Potential Enablers. *Wireless Personal Communications* 87 (2015), 731–757.
- [35] A. Gazet. 2010. Comparative analysis of various ransomware virii. *Journal in Computer Virology* 6 (2010), 77–90. Issue 1.
- [36] A. Gharib and A. Ghorbani. 2017. DNA-Droid: A Real-Time Android Ransomware Detection Framework. In *Proceedings of the 11th International Conference on Network and System Security*. Helsinki, Finland, 184–198.
- [37] H2020-FCT-04-2015/700326 2016. RAMSES: Internet Forensic platform for tracking the money flow of financially-motivated malware. (2016). <https://ramses2020.eu>.
- [38] H2020-ICT-2014-2/671672 2015. SELFNET: Self-Organized Network Management in Virtualized and Software Defined Networks. (2015). <http://www.selfnet-5g.eu>.
- [39] S. Hamalainen, S. Sanneck, and C. Sartori. 2012. *LTE self-organising networks (SON): network management automation for operational efficiency*. John Wiley & Sons, Hoboken, NJ, US.
- [40] A. Kharaz and S. Arshad. 2016. UNVEIL: A Large-Scale, Automated Approach to Detecting Ransomware. In *Proceedings of the 20th USENIX Security Symposium*. Austin, TX, US, 757–772.
- [41] A. Kharraz and E. Kirda. 2017. Redemption: Real-time Protection Against Ransomware at End-Hosts. In *Proceedings of the 20th International Symposium on Research in Attacks, Intrusions and Defenses*. Atlanta, GA, US.
- [42] J.K. Lee, S.Y. Moon, and J.H. Park. 2017. CloudRPS: a cloud analysis based enhanced ransomware prevention system. *The Journal of Supercomputing* 73 (2017), 3065–6084. Issue 7.
- [43] Y. Lemmou and E.M. Souidi. 2017. An Overview on Spora Ransomware. In *Proceedings of the 5th International Symposium on Security in Computing and Communications*. Manipal, India, 259–275.
- [44] F. Mbol, J.C. Robert, and A. Sadighian. 2016. An Efficient Approach to Detect TorrentLocker Ransomware in Computer Systems. In *Proceedings of the 15th International Conference on Cryptology and Network Security*. Milan, Italy, 532–541.
- [45] F. Mercaldo, V. Nardone, and A. Santone. 2016. Ransomware Inside Out. In *Proceedings of the 11th International Conference on Availability, Reliability and Security*. Salzburg, Austria, 628–636.
- [46] C. Moore. 2016. Detecting Ransomware with Honeypot Techniques. In *Proceedings of the Cybersecurity and Cyberforensics Conference*. Amman, Jordan.
- [47] P. O’Kane, S. Sezer, and K. McLaughlin. 2011. Obfuscation: The Hidden Malware. *IEEE Security & Privacy* 9 (2011), 41–47. Issue 5.
- [48] S. Song, B. Kim, and S. Lee. 2016. The Effective Ransomware Prevention Technique Using Process Monitoring on Android Platform. *Mobile Information Systems* 2016, 2946735 (2016), 1–9.
- [49] M.A. Sotelo Monge, J. Maestre Vidal, and L.J. Garcia Villalba. 2017. Entropy-Based Economic Denial of Sustainability Detection. *Entropy* 19, 649 (2017). Issue 5.
- [50] M.A. Sotelo Monge, J. Maestre Vidal, and L.J. Garcia Villalba. 2017. Reasoning and Knowledge Acquisition Framework for 5G Network Analytics. *Sensors* 17, 2405 (2017). Issue 10.
- [51] Symantec 2017. Internet Security Threat Report: Ransomware 2017. (2017). <http://https://www.symantec.com/security-center/threat-report>.
- [52] J. Maestre Vidal, M. Sotelo Monge, and L.J. Garcia Villalba. 2018. A novel pattern recognition system for detecting Android malware by analyzing suspicious boot sequences. *Knowledge-Based Systems* (2018). DOI: <http://dx.doi.org/10.1016/j.knsys.2018.03.018>
- [53] J. Maestre Vidal, A.L. Sandoval Orozco, and L.J. Garcia Villalba. 2016. Online masquerade detection resistant to mimicry. *Expert Systems with Applications* 61 (2016), 162–180.
- [54] J. Maestre Vidal, A.L. Sandoval Orozco, and L.J. Garcia Villalba. 2017. Alert correlation framework for malware detection by anomaly-based packet payload analysis. *Journal of Network and Computer Applications* 97 (2017), 11–22.
- [55] J. Maestre Vidal, A.L. Sandoval Orozco, and L.J. Garcia Villalba. 2018. Adaptive artificial immune networks for mitigating DoS flooding attacks. *Swarm and Evolutionary Computation* 38 (2018), 94–108.
- [56] A. Young and M. Young. 1996. Cryptovirology: extortion-based security threats and countermeasures. In *Proceedings of the 1996 IEEE Symposium on Security and Privacy*. Oakland, CA, US, 1–12.
- [57] A.L. Young and M. Young. 2017. On Ransomware and Envisioning the Enemy of Tomorrow. *Computer* 50 (2017), 82–85. Issue 11.
- [58] A. Zahra and M.A. Shah. 2017. IoT based ransomware growth rate evaluation and detection using command and control blacklisting. In *Proceedings of the 23rd International Conference on Automation and Computing*. Huddersfield, UK.