A Threat Computation Model using a Markov Chain and Common Vulnerability Scoring System and its Application to Cloud Security

Ngoc T. Le
University of Technology Sydney, Faculty of Engineering & IT

Doan B. Hoang
University of Technology Sydney, Faculty of Engineering & IT

Abstract: Securing cyber infrastructures has become critical because they are increasingly exposed to attackers while accommodating a huge number of IoT devices and supporting numerous sophisticated emerging applications. Security metrics are essential for assessing the security risks and making effective decisions concerning system security. Many security metrics rely on mathematical models, but are mainly based on empirical data, qualitative methods, or compliance checking, and this renders the outcome far from satisfactory. Computing the probability of an attack, or more precisely a threat that materialises into an attack, forms an essential basis for a quantitative security metric. This paper proposes a novel approach to compute the probability distribution of cloud security threats based on a Markov chain and Common Vulnerability Scoring System. Moreover, the paper introduces the method to estimate the probability of security attacks. The use of the new security threat model and its computation is demonstrated through their application to estimating the probabilities of cloud threats and types of attacks.

Keywords: Security threats, quantitative security metrics, cloud threats, Markov Chain, Common Vulnerability Scoring System.

Introduction

As cyber infrastructures and their interconnection are increasingly exposed to attackers while accommodating a massive number of IOT devices and provisioning numerous sophisticated emerging applications (Ghayvat et al., 2015; D. Hoang, 2015), security incidences occur more often with severe financial damages and disruption to essential services. Securing cyber systems thus becomes more critical than ever. A simplistic approach to addressing this problem would be to prevent security breaches directly or fix them if they are unavoidable. The approach appears simple and straightforward; however, the achieved solutions are far from

http://doi.org/10.18080/jtde.v7n1.181

satisfactory for several reasons. We have not developed effective predictive tools to anticipate what and where to launch preventive security actions. We may have developed a whole range of tools to deal with security breaches, but this constitutes only temporary and reactive solutions and we are still in the dark, not knowing what comes next!

We suggest a realistic and concrete approach: the goal is to determine the probability of a security threat materialised into an attack (a security breach) on a system, the cost consequences (what it hurts), and the distribution of the costs over the system's constituents or stakeholders (where it hurts) when the threat materialises. Knowing the probability that a threat materialised into an attack we are able to predict the chance that it will occur and take appropriate measures to reduce or prevent its occurrence. Knowing the consequences, we can make appropriate judgments whether the damages caused by the attack are significant enough to warrant a security response or it can be written off as one of the components of the operational costs. Knowing "where it hurts" allows us to use our security knowledge and tools to respond appropriately to the security attack. Clearly, the central issues are the probability of a threat materialised and the distribution of its consequences. In this paper, we only address the problem of determining the probability of a threat materialised into an attack.

The above discussion implies the need for a set of relevant security metrics that allows us to deal with security issues proactively and to set appropriate security goals for our systems and determine the performance of any solution for protecting the systems (both preventing potential incidences and tackling incidences head on). To ascertain the security of a system, it is necessary to develop meaningful metrics to measure appropriately the system's security level or status. Lord Kelvin stated that "when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind" (Thomson, 1889). To measure the security of a cyber space, standards organizations and researchers have proposed many security metrics. The Center for Internet Security (CIS) published a number of security metrics in management, operation, and technique (CIS, 2010). The National Institute of Standards and Technology (NIST) has developed security metrics in implementation, effectiveness, and impact (Aroms, 2012) Other metrics have been proposed for risk assessment and network security evaluation (Hu, Asghar, & Brownlee, 2017; Huang, Zhou, Tian, Tu, & Peng, 2017).

Recently, several security metrics related to the computation of probability of security threats have been developed. In (Patel & Zaveri, 2010), seven types of model-based metrics, which are created by integrating mathematical models and empirical measurements, are also used to calculate the probability of security threat. In (Almasizadeh & Azgomi, 2013), the study used a semi-Markov model to investigate the attack process to compute the transition probability between security states. Mean Failure Cost is one of the sound approaches to quantitative security metrics, taking into account various security components like stakeholders, security requirements, and security threats (Aissa, Abercrombie, Sheldon, & Mili, 2012). The probability distribution of security threats is central to this metric, but the computation is based largely on empirical or qualitative data. Several other security metrics relate to successful attacks, but they are specific to a particular type of attack and hence difficult to generalise.

With these considerations, we pose two questions: (1) how to model a security threat that involves three main security components: attackers, security vulnerabilities, and defenders? and (2) how to predict the probability that the threat materialises into an attack? Considering cloud systems, we address these challenges by proposing a security threat model based on Markov theory to calculate the probability distribution of security cloud threats. For this purpose, the Common Vulnerability Scoring System (CVSS) will be applied to compute the probability of an attack. For evaluating the proposed method, cloud security threats reported by the Cloud Security Alliance (CSA) will be investigated to calculate the probability of cloud threats materialising and the probability of various types of attack. These computation results will generate the quantitative metrics used to measure the security level of a cyber-system (Le & Hoang, 2017).

Major contributions of this paper are as follows:

1. It proposes a security threat model that takes known and major cloud security threats into account. For each security threat, security factors, like attackers, security vulnerabilities and defenders, are investigated to form attack paths for calculating the probability of security threat being materialised.
2. It proposes a method for computing the probability distribution of security threats based on a Markov chain application. The Common Vulnerability Scoring System (CVSS) is investigated to obtain the data for the computation.
3. It provides a method for determining the probability of materialised cloud threats and types of attack using relevant data for supporting security management.

The remainder of the paper is organised as follows. Section 2 provides the background related to security metrics to compute the probability of security threats and Markov theory in security metrics. Section 3 analyses the relationship between security threats and vulnerabilities. Section 4 proposes the security threat model based on a Markov chain. Section 5 describes the computation method for computing the probability distribution of security threats. Section 6 analyses the application of the proposed method in computing attack probabilities. Section 7 concludes the paper with suggestions for future work.

http://doi.org/10.18080/jtde.v7n1.181

Related work

This section discusses related work concerning security metrics related to probability of security threats, and Markov theory in security metrics.

Security metrics related to probability of security threats

For computing security threat probability based on empirical approach, Aissa *et al.* (2012) introduced a security metric named Mean Failure Cost (MFC) that measures the security of an IT system through quantifying variables including stakeholders and the loss resulting from security threats. It includes several desirable features: it identifies stakeholders and provides the cost for each as a result of a security failure; it measures the financial loss per unit of investigation time ($/h). Despite these appropriate considerations, MFC has a major drawback in that the probability distribution of security threats is based on simple empirical data, while security threats are changeable, dynamic, and specific to different IT systems. Due to the stochastic nature of threats, modelling their probability distributions has become a necessity for any security measuring and predicting system. Relevant and sound classification of threats in terms of deployed vulnerabilities, attack motivation perspectives, and likelihood of successful attacks are essential to facilitate the identification of potential security threats and the development of security countermeasures.

For computing security threat probability using a stochastic model, in (Almasizadeh & Azgomi, 2013), the authors used the attack path concept and time is used to calculate transition probabilities. The authors used probability distribution functions to define the transitions of the model for characterizing the temporal aspects of the attacker and the system behaviour. The stochastic model was recognised to be a semi-Markov chain that was analytically solved to calculate the desirable quantitative security metrics, such as mean time to security failure and steady-state security.

For Probability-Based Security Metrics related to security threat, probability-based security metrics usually express the likelihood of an adversary compromising the system or the probability that the system is secure (Ramos, Lazar, Holanda Filho, & Rodrigues, 2017). (Jha, Sheyner, & Wing, 2002) proposed the reliability metric, which represents the probability of an adversary not succeeding in an attack. This metric was obtained from a continuous time Markov chain generated from assigning transition probabilities to the edges of an attack graph. Formally, the reliability of the network is the probability that, in a sufficiently long execution time, the Markov chain will not be in a security failure state. In case not all transition probabilities are available, due to, for example, lack of data about attacks, the authors proposed a Decision Markov Process approach to compute the reliability

http://doi.org/10.18080/jtde.v7n1.181

metric. (Li, Parker, & Xu, 2011) used a renewal stochastic process to estimate the likelihood that an adversary exploits a randomly selected system vulnerability.

Markov theory in security metrics

For a Markov process, the conditional probability distribution of future states of the process (conditional on both past and present states) depends only on the present state, not on the sequence of events that preceded it. Based on this property, several studies have deployed Markov models for security metrics. (Bar, Shapira, Rokach, & Unger, 2016) used a Discrete Markov Chain Model to predict next honeypot attacks. In (Patcha & Park, 2007), to detect anomaly attacks in an intrusion detection system (IDS), the authors used a Hidden Markov Chain to model this system. (Madan, Goševa-Popstojanova, Vaidyanathan, & Trivedi, 2004) used a Semi Markov Model (SMM) to quantify the security state for an intrusion tolerant system. In this work, Discrete Time Markov Chain (DTMC) steady-state probability was applied to compute the mean time to security failure (MTTSF). Anderson *et al.* (2011) proposed a malware detection algorithm based on the analysis of graphs that represent Markov chains from dynamically collected instruction traces of the target executable. (Almasizadeh & Azgomi, 2013) used an attack path concept and time was used to calculate transition probabilities. In terms of security metrics, most research used Markov models in predicting security attacks or malware propagations. To our best knowledge, few studies consider applying Markov chains and for computing the probability distribution of security threats.

The relationship between cloud security threats and vulnerabilities

In this section, we explore the relationship between security threats and vulnerabilities to identify potential attacks.

A security threat is considered as a potential attack leading to a misuse of information or resources, and vulnerability is defined as some flaws in a cyber space (system) that can be exploited by hackers. As a result, a security threat is a potential attack that may or may not eventuate, but with a potential to cause damage. First, we clarify the cloud security threats based on the Cloud Security Alliance (CSA) report (ALLIANCE, 2016; D. B. Hoang & Farahmandian, 2017). The report released twelve critical security threats specifically related to the shared, on-demand nature of cloud computing with the highest impact on enterprise business.

1. Data Breaches (DB). These are security incidents in which confidential or protected information is released, stolen or used without permission by an attacker.

2. Weak Identity, Credential and Access Management (IAM). Attacks may occur because of inadequate identity access management systems, failure to use multifactor authentication, weak password use, and a lack of continuous automated rotation of cryptographic keys, passwords, and certificates.
3. Insecure APIs (Application Programming Interfaces). The security of fundamental APIs is a vital key role in availability of cloud services. From authentication and access control to encryption and activity monitoring, these interfaces must be designed to protect against both accidental and malicious attempts to circumvent policy.
4. System Vulnerabilities (SV). These are exploitable bugs in programs that attackers can use to infiltrate a computer system for stealing data, taking control of the system or disrupting service operations. Vulnerabilities within the components of the operating system – kernel, system libraries and application tools – put the security of all services and data at significant risk.
5. Account Hijacking (AH). It is a traditional threat with attack methods such as phishing, fraud, and exploitation of software vulnerabilities.
6. Malicious Insiders (MI). It is defined as a malicious insider threat created by people in organizations who have privileged access to the system and intentionally misuse that access in a manner that negatively affects the confidentiality, integrity, or availability of the organization's information system.
7. Advanced Persistent Threats (APTs). These are parasitical-form cyber-attacks that infiltrate systems to establish a foothold in the computing infrastructure of target companies from which they smuggle data and intellectual property.
8. Data Loss (DL): for reasons like the deletion by the cloud service provider or a physical catastrophe (including earthquake or a fire) leading to the permanent loss of customer data. Providers or cloud consumers have to take adequate measures to back up data, following best practice in business continuity and disaster recovery – as well as daily data backup and possibly off-site storage.
9. Insufficient Due Diligence (IDD). An organization that rushes to adopt cloud technologies and chooses cloud service providers (CSPs) without performing due diligence exposes itself to a myriad of commercial, financial, technical, legal and compliance risks.
10. Abuse and Nefarious Use of Cloud Services (ANU). Poorly secured cloud service deployments, free cloud service trials, and fraudulent account sign-ups via payment

instrument fraud expose cloud computing models such as IaaS, PaaS, and SaaS to malicious attacks.

11. Denial of Service (DOS). DOS attacks are meant to prevent users of a service from being able to access their data or their applications by forcing the targeted cloud service to consume inordinate amounts of finite system resources so that the service cannot respond to legitimate users.

12. Shared Technology Vulnerabilities (STV). Cloud service providers deliver their services by sharing infrastructure, platforms or applications. The infrastructure supporting cloud services deployment may not have been designed to offer strong isolation properties for a multi-tenant architecture (IaaS), re-deployable platforms (PaaS) or multi-customer applications (SaaS). This can lead to shared technology vulnerabilities that can potentially be exploited in all delivery models.

A security threat usually exploits one or more vulnerabilities in components of a system to compromise it. The relationship between security vulnerabilities and these recognised threats is thus essential for threat modelling. Hashizume *et al.* (Hashizume, Rosado, Fernández-Medina, & Fernandez, 2013) identified seven major security vulnerabilities in cloud computing:

1. Insecure interfaces and APIs (V1). Cloud providers offer services that can be accessed through APIs (SOAP, REST, or HTTP with XML/JSON). The security of the cloud depends upon the security of these interfaces. Vulnerabilities are weak credentials, insufficient authorization checks, and insufficient input-data validation. Furthermore, cloud APIs are still immature, which means that they are frequently changed and updated. A fixed bug can introduce another security hole in the application.

2. Unlimited allocation of resources (V2). Inaccurate modelling of resource usage can lead to overbooking or over-provisioning.

3. Data-related vulnerabilities (V3). This is one of the biggest cloud challenges involving data issues. Data can be co-located with the data of unknown owners (competitors, or intruders) with a weak separation. Data may be located in different jurisdictions which have different laws. Incomplete data deletion – data cannot be completely removed. Data backup is done by untrusted third-party providers. Information about the location of the data usually is unavailable or not disclosed to users. Data is often stored, processed, and transferred in clear plain text.

4. Vulnerabilities in Virtual Machines (V4). Beside data-related issues, vulnerability in Virtual Machines is a big challenge in cloud security. It includes several aspects: possible covert channels in the colocation of VMs; unrestricted allocation and de-

allocation of resources with VMs; uncontrolled migration – VMs can be migrated from one server to another server due to fault tolerance, load balance, or hardware maintenance; uncontrolled snapshots – VMs can be copied in order to provide flexibility, which may lead to data leakage. Uncontrolled rollback could lead to reset vulnerabilities – VMs can be backed up to a previous state for restoration, but patches applied after the previous state disappear. VMs have IP addresses that are visible to anyone within the cloud – attackers can map where the target VM is located within the cloud.

5. Vulnerabilities in Virtual Machine Images (V5). Uncontrolled placement of VM images in public repositories. VM images are not able to be patched since they are dormant artefacts.
6. Vulnerabilities in Hypervisors (V6). These vulnerabilities stem from the complexity of the hypervisor code.
7. Vulnerabilities in Virtual Networks (V7). The vulnerabilities are associated with the sharing of virtual bridges by several virtual machines.

Table 1: Relationship between security threats and vulnerabilities

	Threat	Description	Vulnerabilities	Incidents
1	DB	Data Breaches	V1, V3, V4, V5, V7	An attacker can use several attack techniques involved, like SQL, command injection, and cross-site scripting. Virtualization vulnerabilities can be exploited to extract data.
2	IAM	Weak Identity, Credential and Access Management	V1, V3	An attacker can leverage the failure to use multifactor authentication, or weak password uses.
3	API	Insecure interfaces and APIs	V1	An attacker can take advantage of weaknesses in using APIs like SOAP, HTTP protocol. Bugs in APIs can be also exploited.
4	SV	System Vulnerabilities	V4, V5, V6, V7	An attacker can attack via vulnerabilities in Virtual Machine images, in Hypervisors, and in Virtual Networks.
5	AH	Account Hijacking	V1	To get system access, attackers can use the victim's account
6	MI	Malicious Insiders	V5, V7	An attacker can generate a VM image embracing malware, then propagate it.
7	APT	Advanced Persistent Threats	V1, V4, V5, V6, V7	An attacker can use several kinds of vulnerabilities from specific virtual cloud or APIs to infect bugs permanently in the target system for mainly scavenging data.
8	DL	Data Loss	V3, V4, V7	An attacker can use data-driven attack techniques to gain confidential information from other VMs co-located in the same server; or use the risk of data backup, storing process to scavenge data.
9	IDD	Insufficient Due Diligence	V4, V6	An attacker can leverage weaknesses in complying with rules in using cloud system like configuration of VMs, data and technology shares.
10	ANU	Abuse and Nefarious Use of Cloud Services	V4	An attacker can attack, through use and share of servers, data of customers by using an anonymous account.
11	DOS	Denial of Service	V1, V2	An attacker can request more IT resources, so authorised users cannot get access to the cloud services.
12	STV	Shared Technology Vulnerabilities	V4, V6	An attacker can sniff and spoof virtual networks or exploit the flexible configuration of Virtual Machines or hypervisors.

http://doi.org/10.18080/jtde.v7n1.181

We identify and tabulate the connection between security threats and vulnerabilities in Table 1. It is seen that a security threat may have several security vulnerabilities and one vulnerability may be exploited by several security threats. For example, in terms of threat Data Breaches (DB), five vulnerabilities are involved in this security threat: Insecure interfaces and APIs (V1), Data-related vulnerabilities (V3), Vulnerability in Virtual Machines (V4), Vulnerabilities in Virtual Machine Image (V5), and Vulnerabilities in Virtual Networks (V7). Ristenpart *et al.* (Ristenpart, Tromer, Shacham, & Savage, 2009) indicated that confidential information can be extracted from VMs co-located in the same server. An attacker may use several attacks to collect data by exploiting vulnerabilities in brute-forcing, measuring cache usage, and load-based co-residence detection data processing techniques in cloud systems. Therefore, data leakage depends not only on data-related vulnerabilities but also on virtualization vulnerabilities.

Table 1 indicates that the data-related vulnerability (V3) is involved in three security threats. First, it may cause the threat Data Breaches (DB), when an attacker uses several techniques like SQL injection or cross-site scripting to attack the cloud system. Second, it may lead to the threat Weak Identity, Credential and Access Management (IAM), where an attacker may leverage the data that is often stored, processed, and transferred in clear plain text to gain access to the cloud system. Third, it may cause the threat Data Loss (DL), when an attacker exploits several related vulnerabilities like different located data, incomplete data deletion, and data backup.

Markov model for successful attacks

We introduce a Markov process to describe a cloud attack model and use the CVSS to determine the transition matrix of the proposed Markov model.

A security threat is a stochastic process. We model it as a Markov chain. The probability of transition from one state to others is based on the vulnerabilities present in the current state. An attacker exploits various vulnerabilities to arrive at a security threat state and eventually reaches the final failure state. At this stage, we mainly focus on a first level of abstraction with visible and quantifiable states and construct 3 states, namely the secure state (S), the threat state (T), and the failure state (F). Figure 1 depicts the proposed Markov model for modelling security threats and attacks with state transition probabilities, where α denotes the transient probability from state S to state T, β denotes the transient probability from T back to S, γ denotes the probability to change the state from T to F, δ denotes the transient probability from F state back to T state, ε denotes the possibility from F state back to S state. The model takes all elements of an attack mode into account, including attack, defense and recovery factors of the system. We do not present the direct transition probability from state

S to state F for several reasons. First, we are investigating the impact of security threats on system failure and how an attacker takes advantage of security threats. An attacker tries to exploit vulnerabilities to change from secure state to threat state. Second, the system collapses (goes directly from S to F) mainly in the case of natural disasters or similar catastrophes. This model is simple and practical for our consideration. Even with this 3-state model, it is difficult to derive a set of data for its complete description. We refine the model in several steps of our investigation.

Figure 2 shows the attack model with the defense elements absorbed into the failure state. It means there is no transient probability from F to T or from F to S. When the process reaches F, it stays there with probability 1. This means the recovery process is not taken into account.

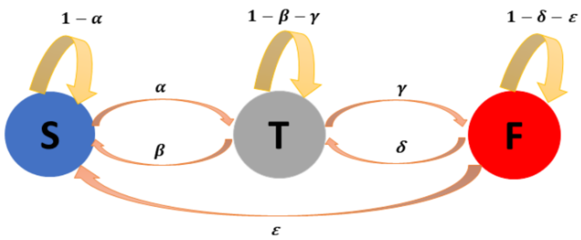

Figure 1. Diagram of attack model with defence and recovery

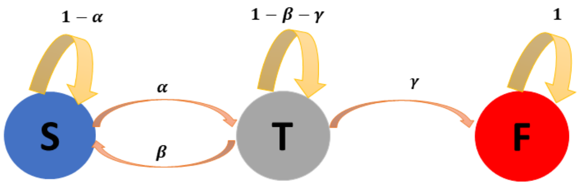

Figure 2. Diagram of attack model with defence and without recovery

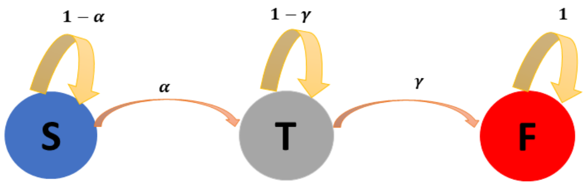

Figure 3. Diagram of attack model without defence and recovery

Figure 3 shows the attack model with the defense efforts absorbed both at the threat state and the failure state. We focus on this kind of abstraction of this model. The aim is to compute the successful chance of attacks by an attacker deploying vulnerabilities of a threat. We do not take into account the recovery element of the system at this stage of investigation, as it can be incorporated at a later stage. Furthermore, recovery efforts largely depend on the manager of the system and relevant data is not often disclosed. The probability from S to T also means the overall probability that includes the defense element that the system tries to change state from T back to S.

We are interested in finding the transition probability from state S to state F in the attack sequence. The Chapman–Kolmogorov equation (Ross, 2014) is available to find the transient

http://doi.org/10.18080/jtde.v7n1.181

probability between two states after a number of jump-steps. The transition probability can be calculated by matrix multiplication. Therefore, to derive the transition probability between two states in a number of steps, the Chapman–Kolmogorov equation can be used as follows:

$$P_{ij}^{m+n} = \sum P_{ik}^{m} P_{kj}^{n} \quad (1)$$

where P is the probability matrix of transitions in the state space. P_{ij}^{m+n} is the transition probability from state i to state j after (m + n) steps via any state k.

Distribution of security threat probabilities

To compute the distribution of security threat probabilities based on a Markov chain, 3 phases can be presented as follows: modelling security threats as a Markov chain; building a transition probability matrix; computing the transition probability from state S to state F via each threat T.

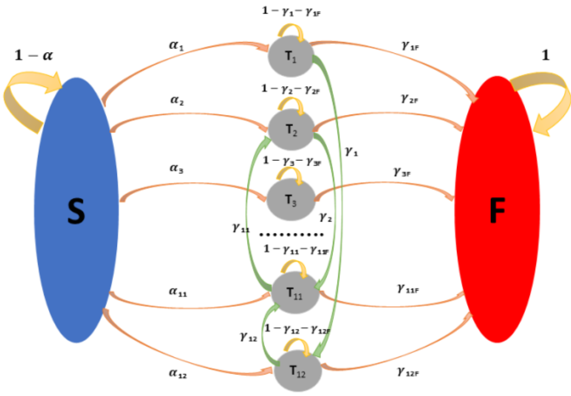

Figure 4. Security threat model with attack process

Phase 1: modelling security threats as a Markov chain. Figure 4 shows an attack model that expands the general model in Figure 3 with twelve attack paths. This is modelled as a Markov chain with fourteen states, including a security state, a failure state, and twelve threat states. The security state is defined as a state of the system that has no failure or security threats. The failure state is a state when the system fails to meet its minimum requirements. The threat state is considered as a middle state that an attacker could exploit a specific set of vulnerabilities. Attack path can be defined as a possible way that an attacker starts from security threat to reach failure state through threat states. In this model, we assume that the probability of an attack path is the overall probability that includes the defense element. This is a simplification, as it is possible that the system can move from one threat state to other determined threat states to reach the failure state.

Phase 2: building transition probability matrix. The probability of each attack path is considered as the probability of changing state security to failure caused by each security threat. An attacker leverages security vulnerability of each security threat (the attack path) to attack to reach the failure state of the cloud system. From the attack model (see Figure 4) we arrive at a transition probability P_{ij} matrix with fourteen states including security, failure, and twelve threat states.

$$P = \begin{bmatrix} 1-\alpha & \alpha_1 & \cdots & \alpha_{12} & 0 \\ 0 & 1-\gamma_1-\gamma_{1F} & \cdots & \gamma_1 & \gamma_{1F} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-\gamma_{12}-\gamma_{12F} & \gamma_{12F} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

In this matrix, α is the sum of probability of all attack paths from S state to T states; and γ_F is the sum of the probability of all threat states to the failure state. Once the system is in the security state, it will remain in this state with probability (1-α) and, once the system is in the failure state, the probability of remaining in this state is 1 (the absorbing state). The probabilities of attack paths representing from S to T states are α_1, α_2, α_3 etc. The probabilities of attack paths representing from threat states to the failure state are γ_{1F}, γ_{2F}, γ_{3F} etc. There are also transition probabilities from one state to other states. However, for demonstration purposes, it is assumed that there is one path from one threat state to another threat state. These probabilities are presented as $\gamma_1, \gamma_2, \gamma_3$ etcetera.

Phase 3: computing the transition probability from state S to state F via threats T_i. According to attack paths theory, each attack-path represents the path that the attacker will take advantage of to reach the failure state (F) from a threat state (T) by exploiting the set of vulnerabilities (v_{ij}) of each security threat. For example, we assume that attack path 1 represents the path where the attacker exploits vulnerability of threat 1 (Data Breaches-DB). Thus, there is a distribution of probability of attack paths when attackers may choose one path to attack in the space of attack paths. To quantify this distribution, we use the concept of weight of each path. CVSS (NVD, 2018) can be used to weigh each path from S to T, from T to F, or between threats to calculate transition probabilities. The weight associated with the transition from S to T_i is determined by computing the ratio between vulnerability scores from S to T_i and all vulnerability scores from S to all threats. By using (2) below, the transition probabilities (α_i) from S to T_i can be calculated. Similarly, the transition probabilities (γ_{iF}) from T_i to F can be computed using (3). To compute the transient probability S to F via T_i, $(P(SF)_i)$, (1) can be used to compute the value in any number of jump-steps. However, at this stage, for the purpose of demonstrating the threat model based on the Markov chain, we compute $P(SF)_i$ in two jump-steps using (4). In this case, the probability between threats may not be considered.

$$\alpha_i = \frac{\sum_j v_{ij}}{\sum_{k,l} v_{kl}} * \alpha \quad (2)$$

$$\gamma_{iF} = \frac{\sum_j v_{ij}}{\sum_{k,l} v_{kl}} * \gamma_F \quad (3)$$

$$P^2(SF)_i = P^2 = \alpha_i * \gamma_{iF} \quad (4)$$

In these equations, i is the index of an attack path, v_{ij} is the vulnerability score of vulnerability j associated path i, $k \epsilon P$ is the set of attack paths.

Table 2. Vulnerability scores

Vulnerability	Acronym	Exploitability score
CVE-2017-14925	V1	8
CVE-2014-4064	V2	2
CVE-2015-5255	V3	3
CVE-2015-4165	V4	5
CVE-2016-0264	V5	7
CVE-2015-1914	V6	5
CVE-2017-6710	V7	7

To calculate the probability distribution of security threats, we need to determine elements of the Markov transition matrix based on the vulnerabilities associated with a threat. From the security state S, the total probability that the system moves to one of the threat states is assumed to be α (α = 0.0318 (Jouini & Rabai, 2015)). We can determine the transition probability that the system moves from S to T_i as the ratio of the sum of vulnerability scores of threats associated with T_i over the total CVSS scores of all threats.

Table 2 shows the CVSS scores (NVD, 2018) associated with relevant vulnerabilities considered in this paper. According to CVSS, this number is a score out of ten. For example, V1 scores eight out of ten because the severity of this vulnerability is very high once it is related to cloud data breach vulnerabilities. In addition, to go to state T_1 from S, an attacker needs to exploit the certain set of vulnerabilities associated with the security threat state T_1. In this case, vulnerabilities one, three, four, five, and seven will be exploited (see Table 1). Therefore, the number of vulnerability scores for the attack path one is W1=V1+V3+V4+V5+V7=30 and the total number of all vulnerability score from S to any T_i is

http://doi.org/10.18080/jtde.v7n1.181

W=177. We can estimate the transition probability from S to T1 ($\alpha_1 = 30/177 * \alpha = 0.00539$). Similarly, other transition probabilities from S to T_i will be computed by using (2). We assume that the transition probability from state T_i to F is highly likely with probability $\gamma_{iF} = 0.95$ for any attack paths (see Figure 4). By computing α_i and γ_{iF}, the transition probability matrix P is obtained. Then by using (1) and (4), we have the probabilistic distribution of twelve security threats expressed in Table 3.

Table 3 Probability distribution of twelve security threats

	Threats	Formula	Probability ($\times 10^{-3}$)
1	DB	$\alpha_1 * \gamma_{1F}$	5.1203
2	IAM	$\alpha_2 * \gamma_{2F}$	1.8774
3	API	$\alpha_3 * \gamma_{3F}$	1.3654
4	SV	$\alpha_4 * \gamma_{4F}$	4.0962
5	AH	$\alpha_5 * \gamma_{5F}$	1.3654
6	MI	$\alpha_6 * \gamma_{6F}$	2.3894
7	APT	$\alpha_7 * \gamma_{7F}$	5.4616
8	DL	$\alpha_8 * \gamma_{8F}$	2.5601
9	IDD	$\alpha_9 * \gamma_{9F}$	1.7067
10	ANU	$\alpha_{10} * \gamma_{10F}$	0.8533
11	DOS	$\alpha_{11} * \gamma_{11F}$	1.7067
12	STV	$\alpha_{12} * \gamma_{12F}$	1.7067

As seen in Table 3, threat Advanced Persistent Threat (APT) has the highest probability (0.55%). The second highest probability is threat Data Breach with 0.51%. Threat Abuse and Nefarious Use of Cloud Services (ANU) has lowest probability with 0.08%. From the distribution of security threat probability, the highest chance for attacking the cyber system relates to threat Data Breaches (DB). In terms of security management, security experts needs to give a decision to protect data or to protect against advanced persistent attacks

Estimation of security attack probability

In this section, to compute the security attack probability, the relationship between attack types and security threats will be investigated. Then, we introduce the probabilistic method to determine the security attack probability distribution.

Relationship between attack types and security threats

A security attack is an information security threat that involves an attempt to obtain, alter, destroy, remove, implant or reveal information without authorised access or permission. In other words, a security attack is an attempt to gain unauthorised access to information resources or services, or to cause harm or damage to cyber systems. It is clear that an attack type relates to security threats. An attack type can use one or several security threats and one threat can involve several attack types. We investigate the relationship between attack types and security threats (Table 4). In (Singh & Shrivastava, 2012), there are five major types of security attack in cloud computing. It is impossible that an attacker can exploit all vulnerabilities in the vulnerability space. Apparently, an attacker or a group of attackers just can exploit several determined security vulnerabilities. These vulnerabilities often are grouped into categories. These categories can be identified by different security threats. Each of these groups of attacks will have specific features that can be recognised and differentiated from other groups. Each group of attacks will fit several security threats. Five different groups of attack and their connection with security threats will be investigated as follows.

1. DOS attacks (A1)

Attackers will take advantage of the availability feature of a cloud system; they aim to overload a target server with service requests in such a way that it is unable to respond to any new request and hence resources are made unavailable to its users. This can be illustrated in several scenarios: (1) Overloading a target with a large amount of junk data, like UDP floods, ICMP floods etc.; (2) Using blank spaces in various protocols to overload target resources, like SYN floods, fragment packet attack, ping of death; (3) Initiating numerous HTTP requests so that they cannot be handled by the server in an HTTP DDOS attack or XML DDOS attack. It is clear that this attack type is related to the threat DOS (T11) and threat MI (T6), when attackers take advantage of a malicious insider to build the botnet for DDOS attacks.

2. Cloud malware injection attack (A2)

Attackers may try to inject a malicious service or even a virtual machine into a cloud system in order to hijack a user's service for their own purposes. These may include data modification, full functionality changes/reversals or blockings. Cloud malware injection attack groups tend to exploit security vulnerabilities that relate to security threats such as data breach, insecure interfaces and APIs, system vulnerabilities, malicious insider, and advanced persistent attack. This type of attack corresponds to 5 threats: DB (T1), API (T3), MI (T6), APT (T7) and DL (T8), when attackers use malicious insiders or advanced persistent threats to inject malware to take control of a cloud system, especially in database management.

3. Side-channel attacks (A3)

An attacker could attempt to compromise a cloud by placing a malicious virtual machine in close proximity to a target cloud server and then launching a side-channel attack. Side-channel attacks have emerged as an active type of security attack targeting system implementation of cryptographic algorithms. This type of attack has a close relationship with several threats such as: (1) AUN (Abuse and Nefarious Use of Cloud Services – T10) when an attacker attacks through using and sharing the servers so that the attacker can implement its malicious virtual machine to perform a side-channel attack; and (2) STV (Shared Technology Vulnerabilities – T12).

4. Authentication attacks (A4)

Authentication is a weak point in cloud computing services and is frequently targeted by an attacker. Today, most of the services still use simple username and password type of knowledge-based authentication. Some authentication attacks are: (1) Brute Force Attacks, where exhaustive combinations of a password are applied to break the password security. This brute force attack is generally applied to crack encrypted passwords when they are saved in a form of encrypted text. (2) Dictionary Attack: unlike the brute force attack, rather than searching all possibilities, the dictionary attack tries to match a password with most occurring words or words of daily life usage and hence it is more effective in terms of speed. (3) Shoulder Surfing: it is an alternative name for "spying" in which an attacker spies on a user's movements to gain his/her password. Here, the attacker observes the way a user enters the password, i.e. what keys of the keyboard the user has pressed. (4) Other related attacks such as Replay Attacks, Phishing Attacks, and Key Loggers. The authentication attack group is related to password attacks; hence, it is pertinent to security threats including: (1) IAM (Identity and Access Management – T2), when an attacker can take advantage from the failure to use multifactor authentication or strong passwords; (2) AH (Account Hijacking) by using a victim's account to get access to the target's resources; (3) ANU (Abuse and Nefarious Use of Cloud Services – T10), when an attacker attacks through using and sharing the servers to gain access to customers' data through an anonymous account. Therefore, A4 has a relationship with T2, T5, and T10.

5. Man-In-The-Middle Cryptographic attacks (A5)

A man-in-the-middle attack is one in which an attacker intercepts messages in the public key exchange process and then retransmits them, substituting his/her own public key for the requested one, so that the two original parties still appear to be communicating with each other. Through this process, the two original parties appear to communicate normally without being aware of the intruder. The message sender does not recognise that the receiver is an unknown attacker trying to access or modify the message before retransmitting it to the receiver. Thus, the attacker controls the entire communication. MIM attacks include: (1) Address Resolution Protocol Communication (ARP) – in the normal ARP communication, the host PC will send a packet which has the source and destination IP addresses and will broadcast the packet to all the devices connected to the network; (2) ARP Cache Poisoning, in which the attacker sniffs the network by controlling the network switch to monitor the network traffic and spoofs the ARP packets between the host and the destination PCs and then performs a MIM attack; and (3) others including DNS Spoofing or Session Hijacking. This attack group (A5) is related to several threats: (1) IAM (Weak identity, Credential and Access Management – T2), when attackers leverage the weakness in using multifactor authentication or fake information leading to loss of credentials; (2) AH (Account Hijacking – T5) by sniffing the connection to catch the cookies of victims between their PC and the web server, then using the cookies to bypass the system. So A5 has connection with T2 and T5.

Table 4. Relationship between security attack types and security threats

	Type	Description	Threats	Incident
1	A1	Denial of Service	T6, T11	Making overloaded requests to the system to stop availability of servers
2	A2	Malware Cloud Injection	T1, T3, T6, T7, T8	Injecting malicious virtual machine or service to get the victim's access to the cloud system
3	A3	Side-Channel attack	T10, T12	Using and sharing the servers
4	A4	Authentication attack	T2, T5, T10	Using weak passwords, sharing technology
5	A5	Man-in-the-middle	T2, T5	Using weakness of multifactor authentication and the cookies of users

Computing the attack type probabilities

Probability computation of an attack type is based on the probability of the set of security threats. It is can be presented mathematically as $Pr(A_i) = Pr(T_1 \; and \; T_2 \; or \; T_3 \dots)$. However, in this paper, we assume that each attack path presents a security threat. There are no relations between these security threats: each threat is independent from other threats. Therefore, the probability of an attack type is the union of the probability of the attack-related security threats. It is formulated as follows:

$$P(A_i) = P\left(\bigcup_{j=1}^{N} T_j\right) = \sum_{j}^{N} P(T_j) - \sum_{1 \le j < k \le N} P(T_j \cap T_k) + \sum_{1 \le j < k < l \le N} P(T_j \cap T_k \cap T_l) - \sum_{1 \le j < k < l < m \le N} P(T_j \cap T_k \cap T_l \cap T_m) + \cdots \quad (6)$$

This probability of the union of any number of sets can be expressed as the following steps: (1) Add the probabilities of the individual threats; (2) Subtract the probabilities of the intersections of every pair of events; (3) Add the probabilities of the intersection of every set of three events; (4) Subtract the probabilities of the intersection of every set of four events; (5) Continue this process until the last probability is the probability of the intersection of the total number of sets that we started with (Taylor, 2019). The probability of an attack type is computed by using (6). For example, to compute the probability of attack DOS (A1), we have $Pr(A_1) = \Pr(T_6 \; or \; T_{11}) = \Pr(T_6) + \Pr(T_{11}) - \Pr(T6 \; and \; T11) = \Pr(T_6) + \Pr(T_{11}) - \Pr(T_6) *$

$\Pr(T_{11}|T_6)$. Because T_6 and T_{11} are independent, $\Pr(T_{11}|T_6) = \Pr(T_{11})$, and therefore $Pr(A_1) = \Pr(T_6) + \Pr(T_{11}) - \Pr(T_6) * \Pr(T_{11}) \approx 0.0041$. Similarly, applying the above algorithm by using (6), we will have the probability distribution of five attack types seen in Table 5.

As seen in Table 5, attack type Malware Cloud Injection (A2) has the highest probability at 1.67%. The second highest probability is attack type Denial of Service (A1) at 0.41%. The lowest probability is attack type Side-Channel Attack (A3) with 0.2%. The distribution of attack probability provides several implications. For an attack countermeasure plan, security practitioners need to care about methods to prevent malware cloud injection attacks, because the chance of this type of attack is highest. For a security manager to make a decision on security investment, it may depend on not only the probability of an attack but also the consequences of this successful attack, because, in several scenarios, the probability of an attack is very small, but the impact is very high in terms of money. As a result, the average security cost, which is the product of the probability of an attack and the consequence of this attack, is quite high. In this case, the manager can prioritise security actions against the kind of attack that makes more damage – for example, if the consequence of denial of service attacks (A1) is ten times higher than that of malware cloud injection (A2), at $1,000,000 and $100,000, respectively. In this case, using the figures from Table 5, the security cost for A1 is $1,000,000 x 0.00409=$4,092, while the security cost for A2 is $100,000 x 0.016=$1,667. Therefore, the security cost for A1 is nearly two-and-a-half times higher than the security cost for A2.

Table 5: Probability distribution of five attack type

	Attack	Description	Probability ($\times 10^{-3}$)
1	A1	Denial of Service	4.092
2	A2	Malware Cloud Injection	16.679
3	A3	Side-Channel attack	2.559
4	A4	Authentication attack	4.091
5	A5	Man-in-the-middle	3.240

Conclusion

This paper has proposed a novel security threat model to compute security threat probability as a metric to measure the security of a cyber-system. For this purpose, we applied a Markov chain model with three states to identify the attack paths through various security threats. Twelve security threats reported by the Cloud Security Alliance and seven security vulnerabilities scored by the Common Vulnerability Scoring System were investigated to quantify the parameters of the proposed security threat model and to compute the probability distribution of security threats. The probability distribution for cloud attack types also was calculated based on the security threat model. Several scenarios for using the probability distribution of security threats and attacks in cloud security management were explained. One of the limitations in the model is that the relationships between security threats have not been taken into account. Several directions are being considered for our future work: one would be to extend the proposed model to include the interrelationship among cloud security threats; another direction is to explore a new model for estimating the distribution of the consequences over the system's constituents or stakeholders once the probability of the materialised threat has been estimated. This will open up research into the area of quantitative cyber security risks.

References

Aissa, A. B., Abercrombie, R. K., Sheldon, F. T., & Mili, A. (2012). Defining and computing a value based cyber-security measure. *Information Systems and e-Business Management*, *10*(4), 433-453.

Almasizadeh, J., & Azgomi, M. A. (2013). A stochastic model of attack process for the evaluation of security metrics. *Computer Networks*, *57*(10), 2159-2180.

Anderson, B., Quist, D., Neil, J., Storlie, C., & Lane, T. (2011). Graph-based malware detection using dynamic analysis. *Journal in Computer Virology*, *7*(4), 247-258.

Aroms, E. (2012). *Risk Management Guide for Information Technology Systems*. NIST Special Publication 800-30.

Bar, A., Shapira, B., Rokach, L., & Unger, M. (2016). Identifying Attack Propagation Patterns in Honeypots Using Markov Chains Modeling and Complex Networks Analysis. Paper presented at the *2016 IEEE International Conference on Software Science, Technology and Engineering (SWSTE).*

CIS [Center for Internet Security]. (2010). The CIS security metrics. Available at http://www.itsecure.hu/library/image/CIS_Security_Metrics-Quick_Start_Guide_v1.0.0.pdf

Cloud Security Alliance. (2016). *The Treacherous Twelve - Cloud Computing Top Threats in 2016*. From https://downloads.cloudsecurityalliance.org/assets/research/top-threats/Treacherous-12_Cloud-Computing_Top-Threats.pdf

Ghayvat, H., Mukhopadhyay, S., Liu, J., Babu, A., Alahi, M. E. E., & Gui, X. (2015). Internet of things for smart homes and buildings. *Australian Journal of Telecommunications and the Digital Economy*, *3*(4).

Hashizume, K., Rosado, D. G., Fernández-Medina, E., & Fernandez, E. B. (2013). An analysis of security issues for cloud computing. *Journal of Internet Services and Applications*, *4*(1), 5.

Hoang, D. (2015). Software Defined Networking? Shaping up for the next disruptive step? *Australian Journal of Telecommunications and the Digital Economy*, *3*(4).

Hoang, D. B., & Farahmandian, S. (2017). Security of Software-Defined Infrastructures with SDN, NFV, and Cloud Computing Technologies. In S. Y. Zhu, S. Scott-Hayward, L. Jacquin, & R. Hill (Eds.), *Guide to Security in SDN and NFV: Challenges, Opportunities, and Applications* (pp. 3-32). Cham: Springer International Publishing.

http://doi.org/10.18080/jtde.v7n1.181

Hu, Q., Asghar, M. R., & Brownlee, N. (2017). Evaluating network intrusion detection systems for high-speed networks. Paper presented at the *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)*.

Huang, K., Zhou, C., Tian, Y.-C., Tu, W., & Peng, Y. (2017). Application of Bayesian network to data-driven cyber-security risk assessment in SCADA networks. Paper presented at the *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)*.

Jha, S., Sheyner, O., & Wing, J. (2002). Two formal analyses of attack graphs. *Proceedings. 15th IEEE Computer Security Foundations Workshop, 2002*.

Jouini, M., & Rabai, L. B. A. (2015). Mean Failure Cost Extension Model towards Security Threats Assessment: A Cloud Computing Case Study. *Journal of Computers*, *10*(3), 184-194.

Le, N. T., & Hoang, D. B. (2017). Cloud Maturity Model and metrics framework for cyber cloud security. *Scalable Computing: Practice and Experience*, *4*, 277-290.

Li, X., Parker, P., & Xu, S. (2011). A stochastic model for quantitative security analyses of networked systems. *IEEE Transactions on Dependable and Secure Computing*, *8*(1), 28-43.

Madan, B. B., Goševa-Popstojanova, K., Vaidyanathan, K., & Trivedi, K. S. (2004). A method for modeling and quantifying the security attributes of intrusion tolerant systems. *Performance Evaluation*, *56*(1), 167-186.

NIST. (2018). National Vulnerability Database. Available at https://nvd.nist.gov/

Patcha, A., & Park, J.-M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, *51*(12), 3448-3470. doi: https://doi.org/10.1016/j.comnet.2007.02.001

Patel, S., & Zaveri, J. (2010). A risk-assessment model for cyber attacks on information systems. *Journal of Computers*, *5*(3), 352-359.

Ramos, A., Lazar, M., Holanda Filho, R., & Rodrigues, J. J. (2017). Model-Based Quantitative Network Security Metrics: A Survey. *IEEE Communications Surveys & Tutorials*, *19*(4), 2704-2734.

Ristenpart, T., Tromer, E., Shacham, H., & Savage, S. (2009). Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. *Proceedings of the 16th ACM conference on Computer and communications security*.

Ross, S. M. (2014). *Introduction to probability models*. Academic press.

Singh, A., & Shrivastava, D. M. (2012). Overview of attacks on cloud computing. *International Journal of Engineering and Innovative Technology (IJEIT)*, *1*(4).

Taylor, C. (2019). Probability of the Union of Three or More Sets. Retrieved March 5, 2019, from https://www.thoughtco.com/probability-union-of-three-sets-more-3126263

Thomson, W. (1889). *Lord Kelvin: Electrical units of measurement. Popular lectures and addresses*. Macmillan, London.

http://doi.org/10.18080/jtde.v7n1.181