Proc. of the Second Intl. Conf. on Advances In Computing, Electronics and Communication - ACEC 2014.
Copyright © Institute of Research Engineers and Doctors, USA .All rights reserved.
ISBN: 978-1-63248-029-3 doi: 10.15224/ 978-1-63248-029-3-87

Towards Quantitative Metrics for Evaluation of VoIP Security Systems

Sheeba Armoogum & Nawaz Mohamudally

***Abstract*—VoIP systems is the new voice communication technology and is playing a key role in various sectors (Government, business, education etc) for the development of a country. Unfortunately, industry and academic researchers find difficulties to evaluate existing systems with focus on security to countermeasure the various attacks for either implementation in a real operator or future enhancement. This paper aims at enhancing the set of evaluation criteria presented in our last work by introducing metrics in a similar way done in software engineering. The V-Model is used to extract data. At the end, a metrics mechanism is developed together with the computation of each metric.**

***Keywords*—Evaluation Criteria, IDS, Metrics, SIP, VoIP Systems**

I. Introduction

Voice over IP (VoIP) technology has gained its maturity worldwide. The rapid change is now very fast for developing countries. During a workshop conducted in 2011, ITU stated that the percentage of penetration of this technology in African countries is already 33% [1]. However, two issues that are always queried by researchers and industry (SIP providers) are the architecture to be deployed in a real operator network and the counermeasures due to VoIP attacks. Since its existence, aroung forty various software or hardware solutions, also known as SIP Defenders or VoIP Protectors (VPs), were proposed [2]. Unfortunately, many systems have limitations upon deployment at clients' premises. For instance, some solutions are appropriate for a particular condition (single attack countermeasure) only [3,4]. Many systems are not scalable (bandwidth) [5,6] while some are not dynamic as they do not have the facilities to automatically upgrade or add new IDS/IDPS at run mode [6,7,8]. Moreover, attackers recognise that securitiy measures are weakly being implemented in many developing countries.

The quality of a VoIP systems with security is largely attributed by their architectures. Thus, stakeholders such as SIP providers, VoP systems designers and researchers need to evaluate systems before implementation or future enhancement.

In our last paper [2], we developed a new technique of evaluation known as the DADMV set of evaluation criteria. From the best of our knowledge, we were not aware of any such technique used. This method will help both researchers and industry to choose the most suitable VoIP system for implementation in a particular company. We will give an outline of this technique in the next section.

In this paper, the quantitative metrics for evaluating the systems using the DADMV method which were ommited in ur previous study, will be addressed. Infact, researchers have difficulties in identifying the exact place of a document or report related to system analysis and design (SAD) to extract information. This paper addresses this issue too. The benefits of this study are as follows: firstly, the VoIP designers can validate their security system using the proposed metrics mechanism before presenting it in front of the clients; secondly, researchers can do an indepth comparative study of VoIP architectures before conducting further enhancement of the most suitable one. The metrics model will also determine the level of transnsparency of research articles published by VoIP Security researchers and SAD report for industry/clients purpose.

The rest of the paper is structured as follows: the methodology is discussed in Section II; we present the enhanced evaluation criteria with metrics in Section III; Section IV introduces the Metrics model and computation;we conclude in section V with some future work.

II. Methodology adopted and Input Data

In industry, the various phases towards the completion of a project is always documented in a report. In this work, we have adopted the software enginering principles, concepts and metrics to indentify the input data for the evaluation of a VoIP system. In software engineering, the Software Development Life Cycle (SDLC) describes development processes for the planning, analysis, design, implementation, testing, documenting, deployment and maintenance of an information system [9]. There are various models used such as the Waterfall model [10], the Spiral model [11], the Agile development model [12] etc. In this study, we adopt the V-Model [13] Hardware Development Life Cycle (HDLC) to describe the various processes. These processes/activities will be used to identify the input data. The data obtained will then be related to the DADMV set of criteria.

We use the V-Model activities and classify the input metrics data into seven classes and sub-classes as follows:

Sheeba Armoogum
University of Mauritius (UOM)
Mauritius

Nawaz Mohamudally
University of Technology, Mauritius (UTM)
Mauritius

a) Class A: System Analysis - This includes Project Defintion (A1) and Requirement specifications (A2) such as Business Requirement Specs and the System Requirement Specs and Change Scenario.

b) Class B: System Design - This includes Functional Design (B1), Design aspects like Extensibilty, Fault tolerance, Maintenability, Modularity, etc (B2), High Level Design such Process view (PV), Deployment view (DV), Physical view (PhV) and Logical view (LV) (B3) and finally Low level Design like PV, DV, PhV and LV (B4).

c) Class C: System Implementation - This includes Unit/Hardware implementation (C1) and System Implementation (C2).

d) Class D: Deployment and Testing - This includes Component tesing (D1) and System Integration testing (D2).

e) Class E: Client testing, that is, Business requirement specification tesing (E1).

f) Class F: Maintenance (F1).

III. The DADMV Set of Criteria and its Metrics

As explained in our previous paper [2], the DADMV set of evaluation criteria which consists of five main groups: the **D**epiction group, the **A**rchitecture group, the **D**etection Group, the **M**itigation group and the **V**alidation group. The Depiction group defines and examines the types of attacks which the architecture can address. The Architecture group analyses the components of the systems and determines if ever there is some similarities with previous studies. The Detection group states the mathematical algorithm(s) or software & hardware technique(s) used in the system proposed. The Mitigation Group states whether or not the attacks are reduced by providing some kind of results through simulation. Finally, the Validation group checks the reliability and effectiveness of the architecture using various means. The DADMV set of evaluation criteria is explained below.

a) The Depiction group:
- Attack/Threat: The attacks that are addressed are stated (M1).
- Target victim: This can be the SIP proxy, User agent Servers (UAS), User agent (UA) or any other nodes (M2).
- Other protocol-based attacks: Despite that we constraint our study on SIP protocol, however, we believe that other protocol-based attacks will interest academic researchers and industry. RTP as discussed by many authors [14][15] is one of them (M3).

b) The Architecture group:
- Components' Description: A clear description of each component used followed by the architectural diagram is mandatory. We assume that this will help SIP providers when deploying the system at the SIP clients' premises as regards to many set up requirements (M4).
- Node visibility: The node protector (e.g router, computer with firewall etc) is to protect the SIP proxy and other nodes of the VoIP system. The node protector must be invisible to legitimate and illegitimate nodes. If the node protector is a router, its existence can be easily guessed by an attacker [16]. Also, an attacker can act as a proxy and steals IP addresses, intercepts SIP traffic and finally bypasses the node protector to access freely the various VoIP services. Here, we want to ensure that there is no IP routing and no SIP proxying [16] for the selected proposals (M5).
- Set-Up Simplicity: During installation/deployment some specifications are required. Here, we will state how far this process is complexed (M6).
- Similarity: This criterion is important in order to prevent redundancy in the comparative analysis process as many studies/defenders are based on or are the extension of the some previous work. This criterion hence simplifies the process by decreasing the number of studies to be compared (M7).
- Flexibility: The system should be flexible in case of new algorithms are added whenever new attacks are discovered (M8).

c) The Detection group:
- Detection scheme: The researchers propose various methods (Mathematical models(s) or programming tools (e.g Shield [17][18] language) & hardware technique(s) used in the system proposed (M9).
- Detection approach: The approaches or principles can be either signature based or anomaly based (M10).

d) The Mitigation group:
- Protection scheme: IDSs are applied algorithms that monitor system activities for malicious activities or policy violations and report them to a console for further study and action. However, some IDSs may attempt to stop attacks as well. In that case, they are known as the Intrusion Detection and Prevention Systems (IDPSs). This issue will be checked (M11).
- Theoretical results: Some results may prove that there is a reduction of attacks (M12).

e) The Validation group:
- Bandwidth Scalability: For the case of very larger amount of DDoS attacks (billions of fake requests per second), the system should be automatically catered for higher bandwidths to avoid network failure (M13).
- Memory usage: Again for the case of very larger amount of DDoS attacks, we want to know the duration the victim SIP proxy can survive before it stops working and crashes for a certain memory used. We want to investigate as well what appropriate

memory could be used to avoid that the SIP proxy being attacked (M14).

- CPU power: we want to investigate the maximum number of attacks before the processor gets overloaded (M15).
- Other practical Results: Other parameters can be analysed and be useful (M16).
- Financial feasibility: Despite that no financial details are described in the studies, based on the architectures, an attempt to inform SIP providers is made whether the system is expensive or not (M17).
- Applications: based on the above criteria (finance and architecture), an attempt to inform SIP providers of the applicability of the defence system. Three situations are considered: small, medium and large organisations (M18).

IV. The Metric Model and Discussion

This section defines the relation between the classes/subclasses and the metrics. The metrics are then computed.

A. *The link*

The link will help the readers/researchers to identify the exact location of an information from a document for the computation of the metrics. The list below illustrates the types of data required for each metrics.

a) M1 depends on A1 and E1.

b) M2 depends on A1 and E1.

c) M3 depends on A1.

d) M4 depends on A1, A2, B1, B3 and B4.

e) M5 depends on B2.

f) M6 depends on F1.

g) M7 depends on A1 .

h) M8 depends on B2 and A3.

i) M9 depends on A1.

j) M10 depends on A1.

k) M11 depends on C2 and B2.

l) M12 depends on C2.

m) M13 depends on B2 and D2.

n) M14 depends on B2 and D2.

o) M15 depends on D2.

p) M16 depends on D2.

q) M17 depends on A2.

r) M18 depends on A2.

B. *Computation of the Metrics*

Various factors are considered when defining each metric's value as follows:

a) A minimum value of -1 and a maximum value of 1 are considered for each metric.

b) The Negative Effect (N-Effect): The metric will take the minimum value of -1.

c) The Positive Effect (P-Effect): The metric will take the maximum value of +1.

d) The Negative Fractional Effect (NF-Effect): If a metric has a negative impact with respect to various cases/conditions, the metric will take a negative number between -1 and 0.

e) The Positive Fractional Effect (PF-Effect): If a metric has a positive impact with respect to various cases/conditions, the metric will take a positive number between 0 and +1.

The metrics are computed based on the above defined values:

a) M1: It will have a N-Effect in the assessment. For each attack, a minimum of -1 mark is assigned. The total marks assigned is –N1, where N1 is the number of attacks (e.g Floods, DNS, etc).

b) M2: It will have a N-Effect in the assessment. For each victim, a minimum of -1 mark is assigned. The total marks assigned is –N2, where N2 is the number of victims (SIP proxy, UA, etc).

c) M3: It will have a N-Effect in the assessment. In addition to metrics M1 and M2, protocol-based attacks (e.g. RTP) are considered. For each such attack, a minimum of -1 is assigned. The total marks assigned is –N3,where N3 is the number of attacks.

d) M4: The set-up requirements and components's desciptions will help SIP providers during deployment. Hence, it will have a P-Effect in the assessment (+1 mark) for transparency and 0 mark if no or ambiguous information is provided.

e) M5: The node protector (e.g. router) should be invisible to atackers and legitimate nodes. Hence, it will have a N-Effect in the assessment, that is, -1 mark is assigned for the two situations: attackers steal IP addresses after IP routing and after SIP proxying.

f) M6: The researchers / industry specialists should be able to judge on the complexity of the deployment process (e.g. whether too many controllers are used in a subnet). A P-Effect and a PF-Effect are considered, that is, for simple process and complex processes, +1 mark and +0.6 mark are assigned respectively.

g) M7: Many new work are based on existing ones. For a certain number of existing work (say N4) where VoIP designers have referenced to a certain number of articles (say X), a PF-Effect value is computed as follows: +X/N4.

h) M8: The sytem will have a P-Effect (+1) if it is dynamic (Extensible when a new algorithm is addded).

i) M9: It will have a P-Effect in the assessment. For each detection scheme, a maximum of +1 mark is assigned. The total marks assigned is +N5, where N5 is the number of IDS used.

j) M10: It will have a P-Effect in the assessment. A maximum of +1 mark is assigned if a signature based or an anomaly based approach is used.

k) M11: Some IDSs can can also defend and mitigate attacks. A P-Effect is considered (+1 mark).

l) M12: Results are important in research and development. A P-Effect is applied for testing in laboratory test-beds.

m) M13: This metric is very critical in security. We believe that severe penalty should be applied if an architecture does not allow to change the bandwidth when there is a large number of DDoS attacks. In this case, the N-Effect (non-flexibility of the system) and the P-Effect are applied (flexibility of the system).

n) M14: Memory is an expensive component. If the proxy does not show any sign of failure for a certain number of Requests per Second (RPS), then a P-Effect is applied. Otherwise, in case of failure, the following NF-Effects are applied:a) for a minimum failure time of 5 mins, -0.2 mark is assigned, b) for a minimum failure time of 4 mins, -0.4 mark is assigned, c) for a minimum failure time of 3 mins, -0.6 mark is assigned, c) for a minimum failure time of 2 mins, -0.8 mark is assigned, and d) for a minimum failure time of 1 mins, -1 mark is assigned. To note that a 5 mins failure time is considered for a maximum length of a phone call of 5 mins.

o) M15: Here the performance of the CPU is tested. If the processor is not overloaded for a certain number of Calls per Second (CPS), then a P-Effect is applied. Otherwise, in case of failure, the NF-Effects are applied similar to metric M14.

p) M16: Other types of performance can be considered here. However, we believe that the most important one will be the resistance of the system due to single and composite attacks. The marks are evaluated based on 3 composite attacks. The PF-Effects are applied as follows: +0.33 mark for resisting a single attack, +0.67 mark for resisting two simultanuous attacks and +1 mark for resisting three or more simultanuous attacks.

q) M17: Cost can be used to calculate the cost/mark ratio at the end of the evaluation exercise before recommendation.

r) M18: This metric will not be used in the calculation of the overall marks. However, it will indicate us the types of organisation where the VoIP system is more applicable.

The Overall Marks = $\sum_{i=1}^{17} M_i$, where M_i is the marks scored for a metric i.

The Cost/Mark Ratio = $\frac{Cost(M17)}{OverallMarks}$.

The recommended system with security for the organisation as per the metric M18 will be the architecture having the lowest cost/mark ratio.

V. Conclusion

In this paper, we introduce a novel metrics model for evaluating VoIP systems with focus on security. The issue of lack of evaluation criteria is addressed. The DADMV method which was introduced in our previous work is further enhanced by introducing its metrics. To obtain the input data, we adopt the same principles that is used in software engineering. The V-Model which is normally being used in SDLC is here used in HDLC. This adopted V-Model help researchers and industry to extract information from SAD report and articles. To obtain the input data in an easier way, seven classes are created. The output results are detailed in section IV. To compute the metrics, the N-Effect, the P-Effect, the NF-Effect and the PF-Effect are used.

The novel metrics model is tested for one VoIP system, but it is beyond the scope of our study, hence not included in this paper since we are not focusing on comparative analysis of various VoIP security systems.

However, there are still rooms for improvements. In our future work, we will firstly make an attempt to provide more metrics. We will also implement testing of this model on various systems. The computation of some important metrics are too straightforward. For example, an enhanced study needs to be conducted for the bandwidth & memory usage and CPU load (M13, M14 and M15). In this work, the exact number of fake calls per second and the number of calls that are passed through the system are not considered.

References

[1] International Telecommunication Union, ITU News Magazine, http://www.itu.int/itunews/manager/ display.asp, visited in July 2014.

[2] S. Armoogum & N.Mohamudally, "Survey of Practical Security Frameworks for Defending SIP Based VoIP Systems against DoS/DDoS Attacks". Proceedings of IST-Africa 2014 International Conference, May 2014, Publication in EEE Xplore Digital Library, in press.

[3] F. Huici, S. Niccolini and N. d'Heureuse,"Protecting SIP against very large Flooding DoS Attacks", In Proc. IEEE GLOBECOM 2009, pp. 1369–1374, ACM 2009.

[4] A. Lahmadi & O. Festor," A Framework for Automated Exploit Prevention from Known Vulnerabilities in VoIP Services", IEEE Trans. On Network and Service Management, pp. 114–127, 2012.

[5] J. Fielder, T. Kupta, S. Ehlert, T. Magedanz, and D. Sisalem, "VoIP Defender: Highly scalable sip-based security architecture," in

International Conference on Principles, Systems and Applications of IP Telecommunications (IPTComm), ACM, Ed., New York, USA, 19-20 July 2007, pp. 11–17, ISBN: 978-1-60558-006-7.

[6] G. Ormazabal, S. Nagpal, E. Yardeni, and H. Schulzrinne, “Secure sip: A scalable prevention mechanism for dos attacks on sip based voip systems,” in Principles, Systems and Applications of IP Telecommunications. IPTComm 2008, Heidelberg, Germany, pp. 107–132, ACM 2008.

[7] S. Ehlert, G. Zhang, D. Geneiatakis, G. Kambourakis and T. Dagiuklas, “ Two Layer Denial of Service prevention on SIP VoIP infrastructure”, Computers Communications, vol. 31, pp. 2443–2456, ELSEVIER 2008.

[8] M.Z. Rafique, M. Ali Akbar, M. Farooq, “Evaluating DoS Attacks Against SIP-Based VoIP Systems”, In Proc. IEEE GLOBECOM, 2009.

[9] J. Patel, “Architectural View in Software Development Life-Cycle Practices”, In Proc. 6th IEEE/ACIS International Conference on Computer and Information Science, IEEE Xplore, pp. 194-199, ISBN: 0-7695-2841-4, 2007.

[10] Zhaohao Sun, “A waterfall model for knowledge management and experience management”, in Proc. Fourth International Conference on Hybrid Intelligent Systems, IEEE Xplore, pp. 472 – 475, ISBN: 0-7695-2291-2, 2004.

[11] B.W. Boehm, “A spiral model of software development and enhancement”, IEEE Computer Society, vol. 21, Issue 5, pp. 61–72, May 1988.

[12] A. Ahmed et. al., “Agile software development: Impact on productivity and quality”, In Proc. IEEE International Conference on Management of Innovation and Technology, IEEE Xplore, pp. 287-291, ISBN: 978-1-4244-6565-1, 2010.

[13] M. McHugh et. al. , “An agile V-model for medical device software development to overcome the challenges with plan-driven software development life-cycles”, In Proc. IEEE International Workshop on Software Engineering in Health Care , IEEE Xplore, pp. 12-19, 2013.

[14] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, “SIP: Session Initiation Protocol,” RFC 3261, Jun. 2002.

[15] S. Ehlert, D. Geneiatakis, and T. Magedanz, “Survey of network security systems to counter sip-based denial-of-service attacks”, Computers & Security, vol. 29, no. 2, pp. 225–243, 2010.

[16] G. Ormazabal, S. Nagpal, E. Yardeni, and H. Schulzrinne, “Secure sip: A scalable prevention mechanism for dos attacks on sip based voip systems,” in Principles, Systems and Applications of IP Telecommunications. IPTComm 2008, Heidelberg, Germany, pp. 107–132, ACM 2008

[17] A. Lahmadi and O. Festor, “Veto: An exploit prevention language from known vulnerabilities in sip services,” in IEEE/IFIP Network Operations and Management Symposium, NOMS, 19-23 April 2010, Osaka, Japan. IEEE, 2010, pp. 216–223.

[18] A. Lahmadi and O. Festor, “Secsip: A stateful firewall for sip-based networks,” in 11th IFIP/IEEE International Symposium on Integrated Network Management, IM’09, Long Island, New York, USA, June 2009.

About Authors :

Sheeba Amoogum received her BSc in Maths, Physics and Electronics in 1997 and Master in Computer Applications in 2000. She is currently doing her PhD in the field of VoIP security. She is a lecturer at the University of Mauritius. Her fields of study are the software engineering, networks and security.

Dr Nawaz Mohamudally is graduated in telecommunications from the University of Science and Technology of Lille I in France and is currently an Associate Professor at the University of Technology, Mauritius where he had been Head of the School of Software Engineering and Business Informatics and the School of Innovative Technologies and Engineering. He is presently the Head of Department of Industrial Systems Engineering and the chairman of the Internet Management Committee, an advisory unit to the Mauritian government.