A systematic mapping study on identifying attack traffic in IP networks

Leandro Cavalcanti de Almeida
Instituto Federal de Educação, Ciência e Tecnologia da Paraíba - IFPB
CEP – 580700-00 - +55 83 3423-9676
Patos - PB - Brazil
leandro.almeida@ifpb.edu.br

Felipe G. dos Santos Universidade Federal da Paraíba - UFPB Centro de Informática – CEP 58051-900
João Pessoa - PB – Brazil
fgs4ntos@gmail.com

Victor Guimarães Pinheiro
Universidade Federal da Paraíba - UFPB
Centro de Informática – CEP 58051-900
João Pessoa - PB – Brazil
victor.tecnologo@gmail.com

Iguatemi Eduardo da Fonseca
Universidade Federal da Paraíba - UFPB
Centro de Informática – CEP 58051-900
João Pessoa - PB – Brazil
iguatemi@ci.ufpb.br

***Abstract*— Attacks on IP networks are increasingly become a problem for the sysadmins, who must try to protect their services, so that they can detect attacks such as DDoS, IP Spoofing or DNS Cache Poisoning. Understanding how are scientific publications in this area is extremely important to achieve this goal. A mapping study is a secondary study type of allow the researcher to have a consistent result in publications in a particular area. This paper aims to present a mapping study in the area of identification and/or detection of attack traffic on IP networks.**

***Index Terms*— Mapping Study, Traffic Attack, IP Networks Attacks, DDoS.**

I. Introduction

Attacks on IP networks are increasingly become a problem for the network administrators, who must try to protect their services, so that they can detect attacks such as DDoS, IP Spoofing or DNS Cache Poisoning. Only in 2012 launched DDoS attacks against U.S. banks generated about 70 Gbps of useless traffic [1].

Understanding how are scientific publications in this area is extremely important to achieve this goal. Usually when trying to perform a scientific research, a researcher starts his work conducting primary studies, which is an empirical study that investigates a specific question. In the next step, the researcher leaves for secondary study, which reviews all primary studies related to a specific research question, in order to integrate and synthesize the evidence related to a specific research question [2].

A mapping study is a secondary study type of allow the researcher to have a consistent result in publications in a particular area [3]. Being a widespread methodology in areas such as medicine, the first works in systematic mapping connected with information technology emerged with the area to software engineering [4].

II. Mapping Study

A systematic mapping study aims at to provide an overview of a research area, identifying the amount, types of search and results available [5]. In order to start a study of systematic mapping, the first step is to set a topic to be mapped. The next step is identification the research questions, that are the questions to be answered by the study. Based on these questions we must create search strings and submit to scientific search engines. The works listed should be analyzed based on inclusion and exclusion criteria defined in the study protocol. The last step is the classification of the works following a predefined taxonomy displaying results in graphs and tables [6].

The aim of the protocol is to define the steps to be followed by the participants of mapping systematic study. The protocol is organized as follows:

A. Setting Theme

In this case the theme was set to “identify attack in IP networks”.

B. Research Questions

- How are distributed publications a over the years about identifying attack traffic in IP networks?
- We can distribute the publications following taxonomy? If yes, how is the distribution?
- How is the distribution of works included in relation to publishing events (Journals, Workshops, Symposium)?
- How is the distribution of works included in the relation to countries?

C. Search Strings

For this mapping we use the automatic search process. This search should follow the same semantics for all search engines used in the research. The semantics used for the search is as follows:

("IDENTIFICATION" OR "DETECTION") AND (("NETWORK IP ATTACK") OR ("ATTACK TRAFFIC") OR ("NETWORK HACKING") OR ("TRAFFIC HACKING"))

The search engines used for this research were:

-IEEEXplore – http://www.ieeexplore.ieee.org

-ACM Digital Library – http://portal.acm.org

-ScienceDirect-http://www.sciencedirect.com/

The unrestricted access to the work of search engines above, through the network of the Federal University of Paraíba is what motivated the choice for them.

D. Exclusion Criteria

Be excluded from the research works that are in agreement with some of the criteria below:

-Papers that are duplicates;

-Papers that are related to the identification or detection of attacks, but not using the network traffic for this identification or detection;

-Papers that are related to the identification or detection of traffic but not using this relationship to identify or detection attacks;

-Papers that present incomplete results;

-Papers that are not related to the identification or detection of attacks in IP networks.

E. Taxonomy

The research protocol classifies works according to the following taxonomy:

- Algorithm: sequence in instructions or well-defined steps used in the study for a specific task;
- Analysis: studies that allow separation in parts in order to explore an efficient way each part independently of the others;
- Model: reference proposed to be followed in the study to accomplish a particular task;
- Case study: simulation of a real environment in order to draw conclusions about this environment;
- Tool: application used in study to perform a particular task;
- Evaluation: method for determining as a technique is implemented in practice.

III. RUNNING THE MAPPING STUDY

In this work the implementation of the mapping study followed the steps defined in the research protocol.

Following, all articles were for inclusion process in the mapping, followed the criteria defined. The next step was to sort the jobs included in a taxonomy defined in the protocol.

Together, the researchers analyzed the results of the individual classification. The differences in the classification process were put to a vote, where a third investigator performed the tiebreaker.

The final result of the classification process was stored in a spreadsheet, where it was possible to extract information such as date of publication, country of authors, search engines and taxonomy. From this spreadsheet was possible to cross collected information to create graphs and to understand better the area.

IV. RESULTS

According to the research questions defined in section II, we extracted responses achieved by conducting the systematic mapping, which will be answered from the results presented in this section.

The first research issue relates to the distribution of publications over the years. In accordance with figure 1 shows that the year 2002 has a small number of articles. In 2009, we found a large number of publications. This growth in the number of publications follows a rate of 5 new papers every year, over the previous year, demonstrating an active concern related to the identification of attacks in IP networks.

The figure 2 answer the research question related to the distribution of work in relation to countries.

This chart highlights the United States, where we concluded that the highest rate of publications on identification and / or detection of attack traffic in IP networks.

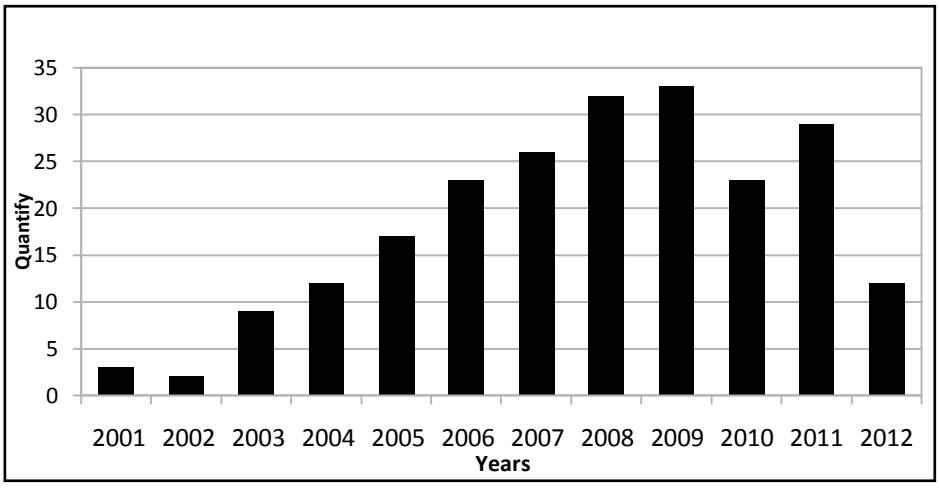

Figure 1: Distribution papers - Years of publication

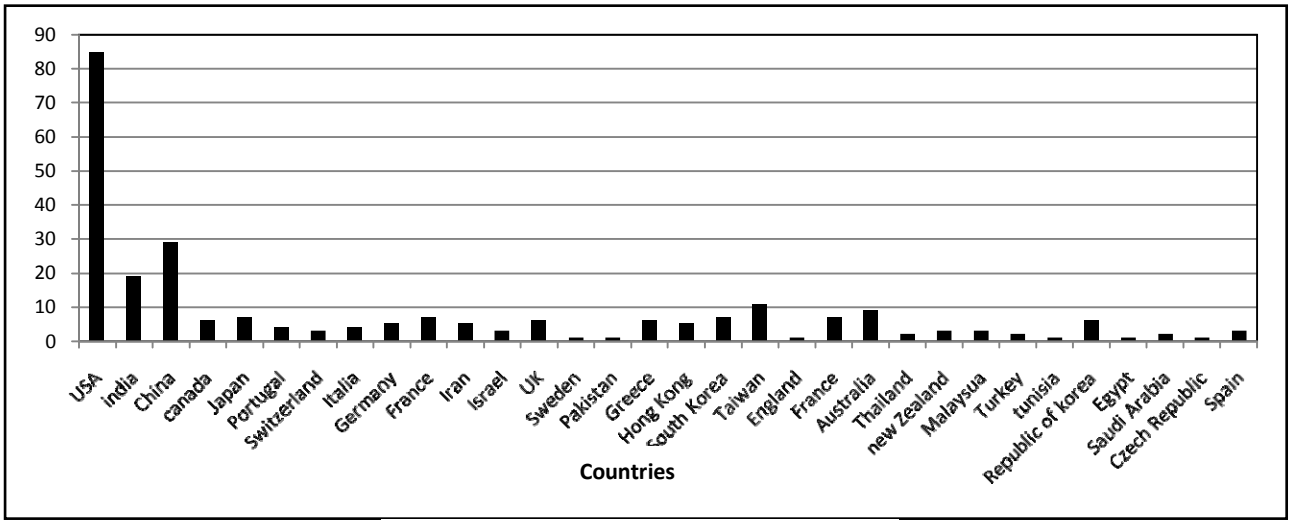

Figure 2: Countries of publications

One possible explanation for this may be the culture that exists in American citizens regarding security, especially after the incidents of September 11, 2001. One concern is noted that Brazil was not present in the statistics presented, demonstrating the deficiency of research related to this issue.

Answering another question, Figure 3 shows the distribution of articles related to search engines. It is noteworthy in this figure that the search engine has indexed more articles was IEEEXplore. The Science Direct had a median income in relation to Articles included. The search engine that got the fewest number of publications was the ACM Digital Library.

One conclusion that can be drawn from these information related to search engines, is that the possibility of an article on the area of identification or detection of attack traffic in IP networks to be found in IEEEXplorer is much higher than in other presented in this article.

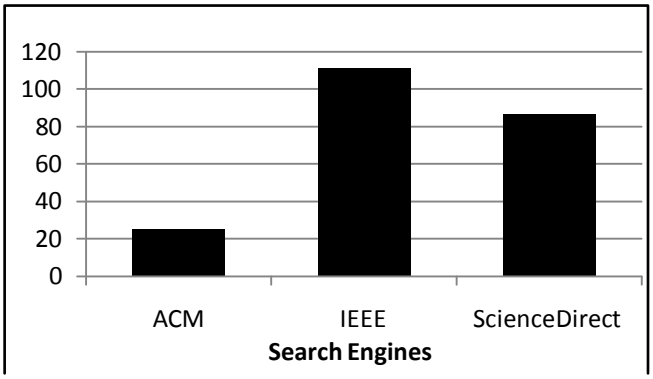

Figure 3: Distribution papers for search engines

Another research question is answered in this study if the publications follow a profile classification based on taxonomy. Figure 4 shows the response from the graph of bubble, which has three variables: taxonomy, facet and intensity.

This type of chart is interesting because from it we can have an holistic view on a particular area, finding the sub-areas surveyed, as well as gaps, i.e., the subtopics where there is no or few publications.

This information shows a general behavior about the research on the area of identification or detection of attack traffic in IP networks and can be used by researchers to define the direction of their studies.

Going contrary to the trend in our mapping only we find an article [7] related to cloud computing. The recent publications year and uniqueness of the article demonstrates that researches related to the identification of traffic in cloud computing environments are not exploiting this sub-area.

Other information that we can draw is that there are high number of articles that propose a model to identify denial of service attacks (DoS) through network traffic.

Within this area there are articles that suggest an analysis of detecting DoS attacks that use the HTTP [8] protocol as well as models for detecting DoS attacks like TCP Syn Flood [9].

We also note the existence of gaps, or areas that are poorly researched, as IP Prefix Hijacking, which had only a article [10] indexed. The lack of studies dealing with algorithms, models, case studies and tools related to the subject demonstrates a range of possibilities for new research.

Other gaps that were highlighted are articles related to algorithms, case studies, tools and reviews on topics such as: DNS Cache Poisoining, RoQ Attacks, Billing Attacks and Low-rate TCP Attacks.

These gaps can expose a potential area for future researches.

V. CONCLUSIONS

This paper presented a study of systematic mapping, in the area of identification and/or detection of attack traffic in IP networks.

One of the most important results achieved in mapping is the bubble chart (Bubble Plot), which allows us to have a holistic view of the searched area.

From this study, many masters and doctoral students may start their research having an overall picture of area, really knowing which sub-area is more developed in his line of research. The next steps of our research include the investigation of the sub-areas not yet discussed, seeking to understand why the lack of published work in this sub-area.

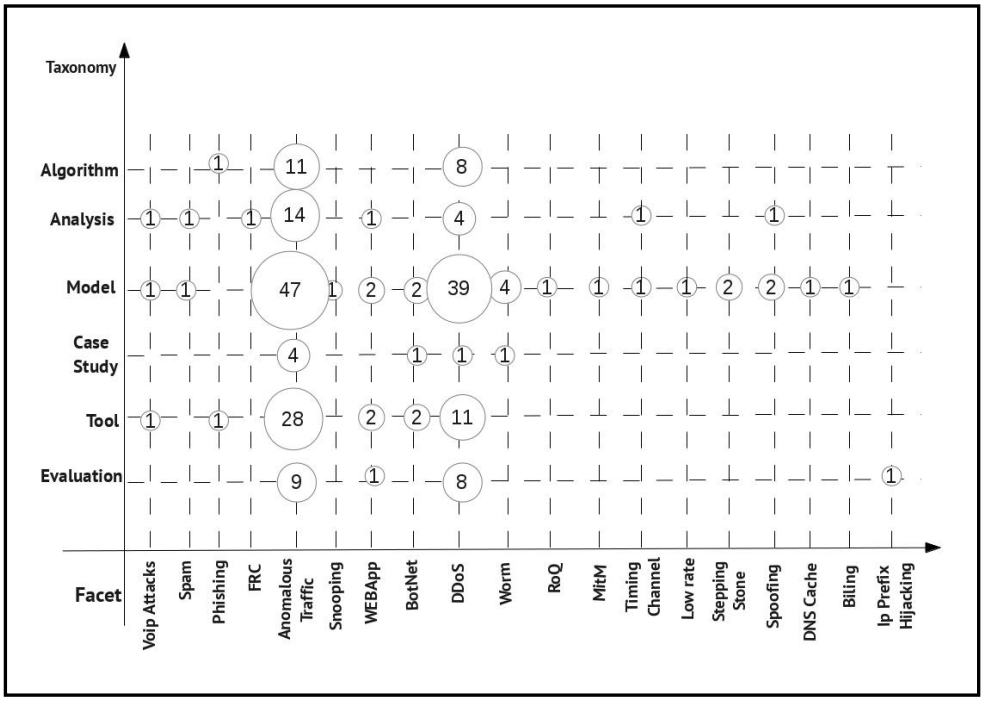

Figure 4: Bubble Plot

REFERENCES

[1] A. Sawas, Computerworld – United Kingdom – 2013 *http://computerworld.uol.com.br/seguranca/2013/02/25/gartner-recomenda-revisao-dos-planos-de-combate-a-ddos-nas-companhias/*. Access in 25/02/2013.

[2] F. Kenji, Introdução à Revisão Sistemática da Literatura. Centro de Informática – UFPE, 2011.

[3] S. Jacinto; F. Silva, Um mapeamento Sistemático da Pesquisa sobre a Influência da Personalidade na Engenharia de Software, 2010.

[4] J. Bailey; R. Feldt; M. Turner; B. Kitchenham; P. Brereton; S. Linkman. Evidence relating to object-oriented software design: A survey, in Proc. of the 1st (ESEM 2007), pp. 482-484.

[5] K. Petersen; R. Feldt; S. Mujtaba; M. Mattsson, Systematic Mapping Studies in Software Engineering, 2008.

[6] B, Kitchenhan; S. Charters, Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007, Keele University.

[7] R. Vanathi; S. Gunasekarans, Comparison of Network Intrusion Detection Systems in Cloud Computing Enviroment. International Conference on Computer Communication and Informatics, 2012.

[8] D. Das; U. Sharma; D. Bhattacharyya Detection of HTTP Flooding Attacks in Multiple Scenarios. Proceedings of the 2011 International Conference on Communication, Computing & Security.

[9] H. Wang; D. Zhang; K. Shin. Detecting SYN Flooding Attacks. Twenty-First Annual Conference of the IEEE Computer and Communications Societies. INFOCOM 2002

[10] Y. Zhang; C. Hu; M. Mao; R. Bush. iSPY: Detecting IP Prefix Hijacking on My Own. IEEE/ACM Transactions on Networking. 2010.