UNIVERSITY OF TWENTE
50 YEARS OF HIGH TECH, HUMAN TOUCH

Masarykova univerzita
Fakulta informatiky
Datum: 29 -04- 2019
Č.j./E.č.: MU-FI/181552/2019/824608/FI-11
Počet listů dokumentu: .....
Počet příloh a listů/sv.: ...../.....
Počet a druh nelist. příloh: ...../.....

Doctoral thesis report

Study Programme: Informatics
Field of Study: Computer systems and Technologies
Name of Student: Tomáš Jirsik
Thesis Title: Cyber Situation Awareness via IP Flow Monitoring

Reviewer: Prof. Dr. Ir. Aiko Pras
Institution: University of Twente, the Netherlands

Scientific impact

In his thesis the author has listed 2 journal publications (as coauthor), and 16 conference publications, of which 7 as first author. According to Google Scholar he has 131 citations. These are good numbers for a PhD student. Also the venues are good to very good, such as IEEE Communications Magazine, IFIP/IEEE IM, IEEE/IFIP NOMS and TMA.

In Section 1.3 the contributions per chapter are given. This list is extensive and convincing.

Research problem

The thesis focuses on cyber situational awareness, and the use of IP Flow monitoring techniques to improve such awareness. There are three main problems being addressed: 1) how to improve the understanding of application layer attacks, without sacrificing performance, 2) how to improve the understanding of attacks on hosts, even if their IP addresses may not be stable nor unique, and 3) how to modify the monitoring of flows such that information can be obtained at earlier stages than before?
The goals of this thesis are to solve these three main problems; the thesis is also structured according to these three problems.

Allow me to give some minor comments. Although the research goals are clear and relevant, they are formulated in terms of "propose" (RG1), "develop" (RG2) and "provide" (RG3). These are rather "things one could do", instead of "things that others will learn". Personally I would have liked to see research questions', plus clear answers to these questions in the conclusions section. The current conclusion section therefore, unfortunately, reads rather like a summary.

In general the text is well written, and figures are clear. There is no need to make changes.

Thesis Structure

The structure of the thesis is clear and convincing. Chapter 2 gives a nice overview of cyber situational awareness, including its history (from aviation). Three

existing definitions are presented and discussed, and an extensive overview of literature is given.

Chapter 3 introduces IP Flow Monitoring. Although already a number of surveys and tutorials in this area exist (which are being cited by this thesis), the way of presenting this subject is novel and includes several nice additions, such as the Timeline(s) presented (amongst others) in section 3.3.5, the open issues and Challenges (Section 3.5.2).

Chapter 4 focuses on the first research goal: how to identify and analyse encrypted applications. This chapter focuses on the performance impact of including application specific flow information (section 4.1) as well as the identification of SSL/TLS versions and the cipher suits that are being used (section 4.2).

Chapter 5 focuses on the second research goal. It discusses the Top N statistics, IPv6 transition mechanisms as well as identification of individual hosts (in particular OS detection). The results are quite nice, and I enjoyed reading this chapter (especially Section 5.3)

The last research goal is discussed in chapter 6: how to reduce the delays in flow monitoring to facilitate real-time cyber situation awareness. Section 6.1 introduces the concept of stream processing, section 6.2 combines this novel concept with traditional batch processing and section 6.3 provides a framework for real-time cyber situation awareness.

Conclusions are provided in chapter 7. In section 7.1 for each of the research goals the main steps to reach that goal are summarised. Section 7.2 presents the topics for further work.

Defence Questions

Possible questions for the defence are:

1. Although the thesis reflects an impressive piece of work, I would be interested what, in your opinion, are the most interesting "discoveries" you made. In other words: are there any findings that you did not expect, or what, in your opinion, are the main "lessons learned"?
2. The thesis is structured around three research goals. For each of these goals there is a dedicated chapter. But how can a reader verify or validate if the goals are met, or that the "solution" provided in this thesis is indeed the best? In some chapters I see subsections on "experiments" or "evaluations" but, for example in chapter 6, how will you convince the reader that your approach (architectures, frameworks") are indeed the best?
3. In chapter 4 you analysed SSL/TLS (up to 1.2) to identify individual hosts. Will your approach work equally well after the introduction of new protocols that include notions like session resumption and limit the number of cyber suite exchanges, such as Quick or TLS 1.3. What would be the impact of these new protocols on your approach?

4. In the conclusions of chapter 5 you say: *the host was identified with high precision – 77.23 % and 36.49 % of the hosts were identified unambiguously*. But why do you say "high" precision? For me 77 and 36% seem relative low numbers.

Conclusions

The thesis demonstrates that the author is able to work independently and is an expert in his field of research. The thesis is well organised and shows the author's capabilities to work independently in this field. The thesis meets the standard requirements imposed on dissertation thesis in this field and therefore I would like to recommend the thesis for acceptance (grade B).

Prof. Dr. Ir. Aiko Pras

Design and Analysis of Communication Systems (DACS)
University of Twente
The Netherlands