KF7004 – MComp Computing Research Project Reflection

16018262

Sept 22

Word count:

Contents

1 Background

1.1 Scope

This research undertaken as part of this module, attempted to assess if there is sufficient data within website log files for a formula to be applied, in order to detect attacks. Part of the scope was to collect a representative sample of real-world data, to measure the effectiveness of the proposed methodology. By using real-world data, I could prove it is more accurate when validating security models; therefore, challenging conventional thinking that studies have to generate their own data.

The scope of this work was correct due to the novel nature of the work and the difficulty I had getting ethical approval, which will be explored in a later section. Even given these challenges, I was able to gather enough data to test my hypothesis. The methodology was tested on a local website and could be applied to any across the world. The risks associated with this region were known and acted as a control. However, due to countries' differing internet patterns, the determination of risk attached to an IP in relation to the country the website is based in may have been harder to assess.

1.2 Objectives

The objectives in my proposal were to; identify how attacks are evading current techniques, understand attack characteristics and refine a formula that can detect and determine risk. The high accuracy of the formula indicates I have understood current attacks and their evasion techniques proving a successful detection methodology. Although, I had to remove an objective, which would have been to develop a better way to assign risk to a country. I intended to look at the number of attacks coming from IP addresses within a country compared to the size of the population. But, I was unable to achieve this. I had issues with my support throughout most of the year, which forced time-constraints upon me; also, this data is difficult to determine. Another reason for its removal was that I was unable to contact the lecturer who suggested this idea, because they now work at another university. Therefore, I kept the original methodology used in the previous research, which still was able to prove that looking at risk for a country is a valid factor to consider.

1.3 Risk

In my proposal I identified the risk that website owners may be unwilling to give me their data for analysis. However, after engaging with them, they were happy for their data to be used. The research could potentially inform them of attacks on their website; if any happened to be found, I could advise them on how they could be mitigated. The biggest risk associated with this project was the lack of previous work. There was a question in my mind that as this methodology seemed relatively straight-forward, why had nobody tried this before? Was it because it was unsuccessful? This risk was mitigated due to the fact that, even if the conclusion did not support my hypothesis I would still be adding knowledge to this field. Therefore, this was a worthwhile risk and one that I would gladly take again.

2 Project Management

Primarily I worked alone, although I did receive some guidance from Nick (a Computer Science lecturer). Thurow et al. 1999 argues multidisciplinary settings allow individuals from different disciplines to contribute their perspectives attempting to solve complex problems. Myself and Nick have complimentary skill-sets that enabled a multidisciplinary approach; focussing on different areas of the research. Nick used his knowledge of programming algorithms to help improve the formula. Whilst, I focussed on aspects of the data pertaining to cyber-security. Hence, I was able to identify risk factors and Nick would then advise on how to efficiently calculate any risk attached to the data. For example, I determined user agents could be an identifier of risk, originally trying a Levenshtein distance to find the nearest user agent. This proved unsuccessful. This approach, may have resulted in misidentification of user agents due to their similarities. However, Nick suggested doing a bag of words approach, which was ultimately a stronger and more accurate methodology. This stopped me overcomplicating the detection of user agents. In future I will begin by looking at simpler approaches.

Looking back on this project, working with Nick was enjoyable, due to our different approaches. Nick always had a positive outlook on the project and its challenges; therefore keeping my motivation high and I came away from our meetings believing the project could be completed with ease. He also helped by, not only generating ideas for the project but helping to implement them. This further helped reduce any stress on me and he enabled me to have faith in our methodology. However, a slight problem was, Nick was difficult to reach at times. The primary way to ask a question to him was looking for a space in his calendar that was mutually convenient. I often felt bad for booking in meetings for issues that could have been solved via email but, this was unavoidable. Which added to my stress, particularly at key times; for example, when submitting my ethics for the third time. I was going to wait for Nick to double-check, but I was getting

more stressed about the ethics not coming back on time; rather than the potential of it being rejected for the fourth time for problems Nick may have spotted. Overall, working with Nick proved beneficial and if I were to do a similar project I would ensure we had better communication outside of meetings if possible.

AJ Shenhar suggests that different projects require different styles of management (Shenhar 1998). The research undertaken required very little project management in the traditional sense, due to the small team involved. By using an unplanned, agile methodology I was still able to deliver a well-rounded research project. This also led to a high degree of being able to question each stage of the project, because we were not rushing to hit deadlines. This is further backed up by **GanttPRO** who suggest that setting unrealistic deadlines is one of the major pitfalls of projects success. Given varying levels of support offered to me throughout the year, it was important that I kept setting realistic weekly deadlines. By engaging as regularly as possible with Nick he was able to critically evaluate my proposed methodology; leading to a robust exchange of views. This also helped me write the paper, as it forced me to concisely explain my thought processes in a clear way and see where the pitfalls of the methodology may be, so I could address these in the paper. Some meetings with Nick only lasted an hour, which when trying to talk over complex areas of code may have been too short. This was compounded by poor internet connection that sometimes made communication difficult.

The work was a follow-on from my undergraduate dissertation. Most of the underlying software was already written, therefore I worked to identify small improvements that could be made. Then developing a methodology to assess the accuracy of the program. However, new ideas from different researchers may have led to a better outcome, because of the opportunity for collaboration. It can be shown that the more collaboration that exists in research projects, the higher the quality of the research (Figg et al. 2006). Therefore the collaboration within this project led to a better output but other collaborators might have increased this further. Myself and Nick would often challenge one another's ideas in order to ensure our methodology was robust. It may have been useful to include another cyber security researcher, in order to further challenge our thinking. Additionally, when classifying the accuracy of the program; having someone to classify the data as well, helping to remove any bias that I might have had. However, I did get a high degree of collaboration with Nick; this is in line with Robert E. Levasseur (2010) who states "to initiate and... the high level of two-way communication" (Levasseur 2010). The initiation process occurred very naturally between myself and Nick. While on placement, we discussed a problem caused by an IP address attacking my website and I proposed that I program something to detect the attack. This sparked a reciprocal flow of ideas that continued throughout the research, even if some elements were absent from the final project.

Throughout this project, there have been many stages where reflection has been needed in order to understand the progress that has been made. By regularly putting my research on hold and taking a step back, I was able to develop a regular reflective habit (Dyment and O'connell 2010). For example, after making changes to the formula, I would run the data again and reflect on the output. However, at other times throughout the year, there have been ongoing issues with a lack of available support in order to complete work. It added more stress and I often felt like I was rushing work that I should have taken more time over. There have been periods where reflection in action have been required, often when doing this we are unaware that we have derived meaning from an experience. Through this process we find ourselves doing the same action again, as it has a positive outcome (Schon 1983). When I reflected on where I was in the project, and what I needed to do to progress it resulted in more effective decision making and problem solving. This was mainly evident by last minutes changes made to my proposal. This meant having to do an online zoom session, so that someone was able to assist me in completing my final draft. The positive outcome here was that I got the work done and I was able to learn that doing work on zoom, although not ideal, still contributed to the success of the project. One thing I've learnt is, not asking other people to proofread my work near the deadline at risk of being too last minute to get any meaningful changes in. This is an example of how I am putting Bourner's theory of reflection, being the experience of turning reflection into learning in action (Bourner 2003).

3 Background, Motivation, Sources of Knowledge

The motivation behind this research was to solve a challenge that I faced, running my own website. The lack of literature in this area to do with cyber security was surprising, however this led to freedom in creating a new methodology. I had to combine some existing techniques and test some new ones. Due to the lack of relevant literature, the literature review was challenging to construct. Therefore it became more about identifying gaps in previous work. Hertz (1997) points out that 'the reflective researcher does not merely report the findings of the research but at the same time questions and explains how those findings are constructed'. (Hertz 1996) This shows that I was able to reflect on the available literature and ask questions that led to my hypothesis. A lot of these questions were not complicated in nature, but came from a working knowledge of websites and servers, the key lesson that I learned in this was that just by asking simple questions about the research, it was easy to identify flaws. Moving forward I will read research with a more critical point of view.

Throughout the project, my motivation fluctuated at differing points. As a student with cerebral palsy, I have an intrinsic motivation to succeed. My disability gives me an obligation to prove myself in order to give myself a sense of purpose. This is parallel to Bye, Pushkar, and Conway 2007 who suggests that disabled students exhibit greater intrinsic motivation in comparison to non-disabled counterparts. At the start of the year, my motivation was high, due to the fact that I wanted to prove my methodology worked due to some people telling me that a mathematical model would be impossible. However, this quickly diminished due to a discourse with the university regarding my level of care. Holloway asserts that with the right level of support disabled students are able to complete their studies at the same rate as their non disabled peers (Holloway 2001) Even though my care needs were not met, I was adamant that I would not be taking a year out, as some had suggested and instead used this adversity as a source of motivation to produce a high-quality research paper. After Christmas, I secured an appropriate level of support from the university, as they agreed that I could reduce my timetable as well as resit some modules next year. This further increased my motivation as I now found that the project was achievable. Even when my motivation was low, I found that, by setting small achievable targets, and completing them would keep me motivated.

4 Relevance

This work was relevant due to the fact that all previous cyber security work in regards to low rate DDoS attacks could be seen as deeply flawed. For example, Tripathi generated their own data to validate their own hypothesis and methods. One of my goals was to show that cyber security work needs to be validated by real data. The work is also relevant due to an increasing reliance on websites and the internet in general. If there aren't effective ways to check attacks, then they have the potential to disrupt daily life. I saw this actualise itself when, after helping a Newcastle University Researcher I was able to reduce the running cost of their web severs by fifty percent. This shows the real world effect that the software has, along with the harm that attacks can cause if they are able to go undetected. The work was also relevant to me on a personal level; as it allowed me to create some research that has never been done before and go on to develop it as a product to sell to other websites. It will also aid in the provision of a safer internet by educating website owners. Once people understand their website's vulnerabilities they're able to take positive action to prevent attacks from occurring post-assessment.

I'm also pleased that I have shown that my work is relevant in the cyber-security field and provides valid conclusions. It may help other researchers think about the way that cyber-security research is conducted.

5 Ethics, Legal, Social, Security and Professional Issues

5.1 Ethics and Legal Issues

Due to complications surrounding consent, ethical approval provided a significant challenge. This study worked using automatically collected data, as a result of this, informed consent could not have been obtained easily. Conventional practice stipulates the requirement for explicit consent before using personal data, resulting in repeated rejection. Retrospectively, a better approach would have been to reference the lawful and relevant use of GDPR in data collection. Having an impartial reviewer giving input on the form may have been beneficial as they could've identified any issues before submission. Furthermore, after the first rejection, it would've been beneficial to enquire who could give appropriate guidance on the issue of consent, which may have led me to Jamie Mahoney quicker.

Jamie was able to point me towards article 14 paragraph 5 of GDPR(European Commission 2016) which states that informed consent inst necessary under the conditions that the provision of such information proves impossible or would involve a disproportionate effort. By citing this information, I was then able to receive ethical approval. After discussing my ethics with other researchers, there was also a suggestion that I could use data that was greater than 6 months old as this falls out of the purview of data protection. Furthermore, I could've tried to anonymise the data in order to simplify the ethics, however this may have added complications to the project. For example, I would be unable to determines the location of an IP address.

6 bibliography

References

Bourner, Tom (2003). "Assessing reflective learning". In: *Education+ training* 45.5, pp. 267–272.

Bye, Dorothea, Dolores Pushkar, and Michael Conway (2007). "Motivation, interest, and positive affect in traditional and nontraditional undergraduate students". In: *Adult education quarterly* 57.2, pp. 141–158.

Dyment, Janet E and Timothy S O'connell (2010). "The quality of reflection in student journals: A review of limiting and enabling factors". In: *Innovative Higher Education* 35, pp. 233–244.

European Commission (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. URL: `https://eur-lex.europa.eu/eli/reg/2016/679/oj`.

Figg, William D et al. (2006). "Scientific collaboration results in higher citation rates of published articles". In: *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 26.6, pp. 759–767.

Hertz, Rosanna (1996). "Introduction: Ethics, reflexivity and voice". In: *Qualitative sociology* 19.1, pp. 3–9.

Holloway, Sarah (2001). "The experience of higher education from the perspective of disabled students". In: *Disability & Society* 16.4, pp. 597–615.

Levasseur, Robert E (2010). "People skills: Ensuring project success—A change management perspective". In: *Interfaces* 40.2, pp. 159–162.

Schon, D (1983). *The reflective practitioner. How professionals think in action*. London: Temple Smith.

Shenhar, A. J. (1998). "From theory to practice: toward a typology of project-management styles". In: *IEEE Transactions on Engineering Management* 45.1, pp. 33–48. ISSN: 0018-9391. DOI: `10.1109/17.658659`.

Thurow, Amy Purvis et al. (1999). "The dynamics of multidisciplinary research teams in academia". In: *The review of higher education* 22.4, pp. 425–440.