# KF7004 – MComp Computing Research Project Reflection

16018262

Sept 22

Word count:

## Contents



# 1 Background

## 1.1 Scope

This research attempted to look at whether there was enough data in website log files for a formula to be applied, in order to detect attacks. It was part of the scope not to generate my own data and collect a representative sample of real world data to measure the effectiveness. I think the scope 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. But I was still 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 country's different internet patterns, it would be difficult to understand different traffic patterns.

### 1.2 Objectives

The objectives in my proposal were to; identify how attacks are evading current techniques, understand attack characteristics and refine formula that can detect and determine risk. 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. However, due to time constraints I encountered due to a lack of support from the University, Another reason for its removal was that I was unable to reach the lecturer who suggested this idea, because they now work at another university. Therefore, I kept the original methodology used in the previous research. This was just looking at the number of attacks from within a country. I think the fact that the formula has worked shows that I have understood current attacks and their evasion techniques proving a successful detection methodology.

### 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 talking they were more than happy to cooperate. Largely due to the fact they wanted to know if their website was under attack; if it was how the attack could be stopped. The biggest risk associated with this project was the lack of previous work. Because, there was a question in my mind that this methodology seemed so simple and obvious, why had nobody tried this before? Was it because it did not work? 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 to take and one that I would gladly take again.

## 2 Group vs Individual

Although I was primarily working alone, I did receive a lot of assistance from Nick. Thurow et al. 1999 argues that multidisciplinary settings allows individuals from different disciplines to contribute their disciplinary perspectives in an attempt to solve complex problems. Nick does not have any knowledge of cybersecurity, but does have extensive knowledge of algorithms; unlike myself who does have knowledge of cyber security. Hence, I was free to experiment and he was able to advise on all the methods used without getting caught up in the smaller details. For example, I had identified that user agents could be an identifier of risk, I had originally tried a Levenshtein distance to try and find the nearest user agent. However, the academic 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 projects I will now begin by looking at simpler approaches first. Furthermore, Nick had prior knowledge of the project, due to supervision at undergraduate level and therefore understood the project and its complexities. This occurred organically in the present research as ideas could be effectively implemented without disagreement, regarding the direction of the project.

Throughout the year myself and Nick tried to get into a regular habit of meeting; this enabled us to see what did or did not work. Also preventing delays in the research if I got stuck on something. Despite this, one downside of this was if there had been a slow progress week it felt like I hadn't done anything, which negatively impacted my motivation. The project was never planned out at the start, this led to a more agile development. Momentum easily kept going and I was able to establish if a method was working or not. Furthermore, due to the novel nature of the work there was a danger that I would set unrealistic deadlines to get the work done, which is deemed by **GanttPRO** to be one of ten major pitfalls of projects success(**GanttPRO**). By not having afixed plan I was able to adapt to the level of support I received. On reflection I think this was the correct way. Sometimes I would get called at 6pm, saying their was someone

available the next day. But it meant some of our meetings involved brainstorming on whiteboards. This was sometimes difficult to read if the meeting was on Teams, due to bad internet connection. On reflection the whiteboard feature on Teams would have been better. This work required a different approach to manage the project. This is inline with AJ Shenhar's theory on project management (Shenhar 1998), who suggests that different projects require different styles of management. So due to external factors an agile project management style was needed.

I worked alone largely due to the fact that the work was heavily linked to the work carried out for my undergrad dissertation. Most of the underlying software was written as a part of my dissertation, therefore I worked to identify small improvements that could be made. Then developing a methodology to prove that accuracy of the program. It could have been useful to get new ideas in, which may have led to a better outcome. If I were to have worked in a group the collaboration would have been beneficial as Robert E. Levasseur (2010) states, "to initiate and... the high level of two-way communication" (Levasseur 2010). Retrospectively, the initiation process occurred very naturally between myself and Nick. While on placement, we discussed a problem that I was having with an IP address attacking my website and I suggested that I programme 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. 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. For example, 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.

# 3 Background, Motivation, Sources of Knowledge

Throughout this project, there have been many stages where reflection has been needed in order to understand the progress that has been made. At times, progress has felt limited and non-existent. However, these periods have allowed for reflection. 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). However, at other times throughout the year, there have been ongoing issues with a lack of available support in order to complete work. 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'd change 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 was a continuation of my undergraduate project, looking at a new way to detect attacks on websites. The motivation of this 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'. (CITE) This shows that I was be 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

# 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 validated 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.

# 5 Ethics, Legal, Social, Security and Professional Issues

## 5.1 Ethics

Due to the ambiguity of user consent in the collection of website log data, and whether users had given informed consent about the collection. Websites say that data can be analysed however due to the university policy of expressed consent, there were questions about whether the data could be used. However, due to the study wanting real-world data, this might have changed user behaviour, knowing that they would've been tracked. This was done to mitigate the effects of possible confounding variables, such as social desirability. Studies have suggested that individuals online behaviour changed when they are being monitored.

Another ethical implication of the work may be the fact that entire country is given a risk. This is mainly done to give the software an idea of context, however a user from a high risk country can still get a low overall risk score.

# 6 Legal Issues

There was a potential legal issue as people could not opt out of the data analysis. This is due to the fact that as soon as they went on to the website their IP address was logged. Most websites privacy policy state that IP addresses will be logged and used for analysis, therefore most users should be aware of how their data will be used.

# 7 bibliography

# References

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.

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.

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.