A Digital Forensics Profiling Methodology for the Cyberstalker

Alice Silde
University of Derby
School of Computing and Mathematics
Derby, United Kingdom

Dr Olga Angelopoulou
University of Derby
School of Computing and Mathematics
Derby, United Kingdom

Abstract—Cyberstalking has many new manifestations with the opportunities offered by modern technology. It is a growing issue that presents significant distress to its victims. This type of crime presents difficulty to computer forensic investigators in large amounts of digital evidence accumulated over time and in offender apprehension. There is currently little research in the profiling the cyberstalker through a digital investigation. However, techniques employed for dead-end offline investigations, including criminal profiling are also be applied to cybercrime. This research focuses on developing a profiling methodology for profiling the cyberstalker. We adopted the Behavioural Evidence Analysis methodology to the digital investigation process and attempted to construct our profiling methodology. The methodology was evaluated by using a simulation of stalking behaviours to produce a technical, as well as a criminal profile. The aim of this work is to provide a starting point on the profiling of the cyberstalker. The investigators could then proceed with informed assumptions about the offender, according to his digital traces and also achieve reduced time for the digital investigation.

Keywords—profiling; cybercrime; cyberstalking; digital forensic investigation

I. INTRODUCTION

Stalking has been a known issue for a long time. However, with the development of technology that has introduced the Internet, such as social networking, Cloud computing and various mobile computing devices; there are many new manifestations to this crime [24]. In addition, there are various technologies that the offenders can use to avoid detection and pursuit [28]. Cyberstalking is now more common than offline stalking, with 20% victims being stalked through social networks [21]. Statistics show that cyberstalking is on the growth [31; 30]. Just like offline stalking, cyberstalking causes the victim serious psychological harm through terror and helplessness, with 30% of women and 20% of men seeking professional psychological help after being cyberstalked [21].

Some of the main difficulties to the investigation are posed by the large amounts of the collected digital evidence. The crime occurs over time, and there is difficulty to apprehend the offender, as they may not be connected to the victim in the physical world. Fortunately, due to the repetitive, persistent and personal nature of the crime, it lends itself to criminal profiling. Criminal profiles are widely used in investigations to link together different crimes, to narrow down lists of suspects and to aid the process of investigation [25]. Profiles help build a bigger picture and reconstruct the crime, when there are too many unknowns [33]. An estimated 77% investigations, that make use of profiling, have a successful outcome [25].

There have been attempts to profile different types of cyber criminals, such as hackers [9], identity thieves [1] proving that the idea is viable. However, the study area is relatively new [22] and there is no formal methodology of profiling cybercriminals [2].

The next section of this paper will present the current research in profiling offenders and particularly cyberstalkers, as well as the existing methodologies in digital investigations as the base of this research. Following that, our proposed methodology will be presented and discussed. The focus for the purpose of this paper has been given to our experiment. The results of the technical profiling stage will be then summarised and discussed in the context of their significance to the field.

II. STATE OF THE ART

There is a lack of consensus of whether cyberstalking should be seen as a new type of crime or simply as an extension of offline stalking. Burmester et al. [5, p. 2] define cyberstalking as

“repeated and persistent attempt by one individual, the stalker, to harass another individual, the victim, using the Internet or other open networks”.

Other authors provide similar descriptions, focusing on the unwanted, repetitive behaviour that involves the Internet or related technology. Different types of stalking can be identified, based on the technology or approach used by the stalker. Email is used in 79% of stalking cases [10], and hence is the most popular stalking technology.

...
A. Profiling Cyberstalkers

Other than the complexity of the involved behaviours and motivations, lack of information and research on cyberstalking [22; 25; 26; 4] has also been a factor in preventing successful profiling. However, the current research in the field provides a strong starting point to producing a more detailed profile, as well as an overview of available methods.

[4], [2] and [15] all study the victimology of cyberstalkers, providing insight into victim demographics, based on statistical research, as well as setting out approaches for examining victim lifestyle and exposure. [22] also examine victim-offender relationships, which is a significant element in understanding the offender’s motivations [33].

Many authors [24; 11; 29; 20] examine the cyberstalker’s motivation statistically by using typologies for offline stalking or developing new typologies. There is very little consideration for the technical aspects of the crime, such as the evidence left by the offender, in the studied work. [28] apply crime scene assessment to explore some of the technical aspects of the crime. Rogers [25] also touches on the importance of assessing the offender’s technical skill and identifying potential locations of digital evidence.

B. Profiling Methodologies

Profilers make extensive use of knowledge of law, sociology, criminology and psychology [30] to make assumptions about the unknown offender. While many argue that profiling is rarely scientific [14], there are several methodological approaches to assessing the crime scene, the victim and the evidence.

Most known methods are inductive and use statistical information about known offenders [25]. Deductive methods, however, are deemed more suitable for cybercrime investigations [25; 30]. The most popular deductive profiling methodology is the Behavioural Evidence Analysis (BEA) framework that has four main stages: equivocal analysis, crime scene assessment, victimology and offender characteristics [33]. [33] considers the application of cybercrime as well as physical crime, demonstrating the flexibility and extensibility of involved methods. Famous inductive approaches, such as the FBI Model [13] or Investigative Psychology [6] extend the range of available tools by outlining processes of classifying the crime, examining victim-offender relationship and providing more ways of evaluating offender motivation. For example, the organised/disorganised [29] and commuter/marauder [25] dichotomies stem from inductive approaches.

C. Digital Investigation Methodologies

There are different investigation specifics to consider when dealing with different types of crime. The ultimate goal remains the same - identify the responsible parties and reconstruct the event in as much detail as possible [25]. The main difference of online crime is in the extreme volatility of digital evidence [8], indirect and tamperable time-frames [17] and technically complex anti-forensics [29]. These specifics need to be considered when profiling a cybercrime, as they shift the focus towards the handling and interpretation of digital evidence. This can be ensured by integrating profiling into an established digital investigation framework, which provides a methodology for working with digital evidence and using it to reconstruct criminal events.

Casey’s [8] Digital Forensic Investigation framework (DFIF) is one of the industry standards, provides the necessary tools for examining evidence, and ensures adaptability of set methods to different types of crime. Other methodologies, such as the Objectives-Based Hierarchical framework [3] or the Computer History Process Model [7] can also be instrumental for guiding the type of thinking the investigator or the profiler should adopt for investigating cyberstalking.

Most of the existing work does not focus on the technical aspects of the crime, provides typologies rather than profiles and bears mostly academic significance. This paper addresses the gap in research into the common digital evidence of cyberstalking behaviours on the victim’s and the offender’s computer, proposing a method for examination of digital traces of cyberstalking. It also introduces a deductive non-statistical approach to criminal profiling with consideration of the digital investigation process. In both cases, a generic scenario of email/IM cyberstalking behaviours, obtained from case studies and current research is used to build the profile.

III. METHODOLOGY

We propose a combination of BEA [33] and DFIF [8] as the methodology for constructing the criminal profile of the cyberstalker. A high level representation of the proposed methodology is demonstrated in Figure 1.

The Discovery/Accusation, Examination and Analysis phases of DFIF were used, as profiling will only occur in the practical phases of the framework. Each phase contains investigative processes. This methodology inserts profiling stages into each phase. These stages are one or several of the four main profiling stages of BEA: equivocal analysis, crime scene characteristics, victimology and offender characteristics. Equivocal analysis occurs only in Discovery/Accusation, as it is the equivalent of initial forensic analysis of crime scene evidence in physical crimes [33]. The other three stages occur throughout the investigation, continuously enhanced by the investigative processes. Profiling stages receive inputs from investigative processes and produce outputs that can be used in another profiling stage or the next investigative process. The majority of inputs are digital evidence found, extracted and processed within the investigative processes. However, deductions based on this evidence can also serve as inputs, ensuring the profile builds up as the investigation progresses. It is obvious that technical aspects of working with digital evidence are crucial. In this methodology, the findings of technical investigation also contribute to the composite profile.
The novelty of this method lies in the awareness of the digital evidence; its significance and availability at each stage of profiling process. In addition, the offender’s technical skills and forensic awareness are evaluated, using motivational typologies and other approaches. This stage occurs after the reconstruction process in the model, as it draws on conclusions of criminal profiling and knowledge of criminal events.

Fig. 1. Integration of BEA into DFIF

However, the technical phase of the methodology is where we focus this paper. It is the most under-researched aspect of cyberstalker profiling to date. It was originally intended to confirm the evidence priorities and search locations proposed in the criminal profile, but can also guide future cyberstalking investigations by focusing the search and recovery of evidence. Figure 2 outlines the basic inputs and outputs of the technical profiling stage.

1. **Offender skill** – an evaluation of the offender’s technical aptitude, based on their known criminal behaviour can help establish the likely conduct of the unknown behaviours, such as the likelihood of use of anti-forensics, as well as the extent and complexity of such measures.

2. **MO details** – knowing the details of criminal actions, such as the IM or email client used, is useful in establishing likely locations of evidence, as well as types of evidence involved in these actions.

3. **Offender motivation** – can provide insight into potential unknown actions or behaviours or help explain digital traces that do not describe known activities.

4. **Communication files** – communication is central to most types of cyberstalking [16], hence any communication-related files are necessary for the technical profile, and they may point to additional communication-related locations.

5. **Evidence priorities** – the technical profile tests the evidence priorities suggested in the criminal profile, as well as provides additional, more detailed insight into evidence that should be prioritised in case of specific stalking behaviours.

6. **Crime classification** – the type of crime can point towards typical tools of committing it. For example, in case of email/IM stalking, the investigator should expect to find traces of email communication or an IM client installed on the victim’s and the offender’s machines. The type of crime also suggests typical methods of concealment/evasion.

7. **Application logs** – if use of specific applications is part of the MO, application logs can confirm or deny any related assumptions.

8. **Evidential files** – this input refers to any other files (documents, images, logs) that are related to the MO or suspected unknown behaviours.

9. **System logs** – investigating system logs can reveal the use of malware or other suspicious activity.

10. **Registry** – registry entries are also indicative of spyware/malware installed on the victim’s machine. It is also useful to examine the registry on the offender’s machine to establish configuration and settings of any crime-related software.

![Technical Profiling Diagram](image)

Fig. 2. Inputs and outputs of the technical profiling stage.

The outputs describe the goals of the technical profiling stage - to establish most likely evidence locations, examine the types of evidence inherent to the crime and suggest how and where evidence of specific behaviours would most likely appear on the victim’s and the offender’s machines. Anti-forensic techniques and related evidence on both sides of the crime are also examined.

**IV. TESTING**

The method chosen for the technical profiling is inspired by the honeypot approach [19; 9], and involves the use of two virtual machines (VMs) to simulate the cyberstalkers’ and the victim’s behaviours accordingly. Virtualisation ensures a controlled environment and ethical conduct of the experiment. Following the simulation, a formal process and a standard forensic toolkit for imaging and examination of the VMs are used to ensure soundness and validity of obtained results. The
methodology of the experiment is outlined in Figure 3. While the method of profiling involves a thorough investigation of VM images, detailed knowledge of simulated behaviours is used to eliminate the unknown factors.

The setup and configuration of the virtual machines involves selecting the necessary software and configuring user profiles both on the system and online. For the purpose of the experiment Windows 7 was selected as one of the most commonly used and current operating systems [34], and VM hardware specifications were selected so that the two machines could run on the same host with minimal performance restrictions. The rest of these choices depend on the selected behaviours and anti-forensic techniques. In this case, email/IM stalking was simulated, hence IM, email and social network accounts were necessary for both the victim and the offender, as well as client software or Web interfaces to access these services. Basic anti-forensic behaviours were simulated, using Web-based proxies, Web-based remailers and a client-based Proxy. The anti-forensic behaviours simulated were designed to be accessible to anyone who has slightly above average computer literacy and can perform an online search for appropriate tools and techniques.

![Fig. 3. The technical profiling methodology.](image)

Some of the victim behaviours were simulated using a Web-based email and then repeated, using a local email client. In order for the simulation to produce results close to those likely to appear in a real investigation, some normal behaviours were also simulated, for example, downloading and setting up a desktop background or visiting Websites related to victim’s lifestyle or interests.

The focus of the investigation is on the victim’s machine, as that is commonly the primary source of evidence in a cyberstalking investigation [33]. However, the offender’s machine was also examined to provide more insight into the evidence of simulated behaviours.

V. RESULTS

During the course of the investigation, most of the evidential priorities and their locations on the victim’s device were confirmed. IM conversation logs, email communications, victim’s personal files, victim’s social network profiles, frequented websites/forums, recent searches and other online activity, cached images and system logs were indeed the most useful evidence in the case of selected email/IM stalking behaviours.

A. Victim’s Machine

The majority of user files were contained in C:/Users/<user name>, including personal files, documents, images and online browsing preferences. These revealed information about the victim’s lifestyle, preferences, as well as contained the evidence that the victim gathered on the offensive behaviours as part of the simulation.

In this case, files related to their online activity were mostly located in C:/Users/<user name>/AppData Local and Roaming folders. These files contained evidence of sites visited by the victim, including the online chatroom, where the stalking began, and the Web-based email client used as part of the simulation. Additionally, the victim’s online chatroom nickname, cache of images received from the stalker or saved by the victim, as part of the evidence gathering behaviour, as well as confirmation of normal behaviours were located among the Internet files.

Examination of C:/Program Files and C:/ProgramData provided information on the installed software, such as the Yahoo IM client and Thunderbird email client, used for the simulation. However, it was established that this information has little to offer to the technical profile, other than a confirmation of the use of these tools.

The majority of messenger and email client information can be found in the C:/Users/<user name>/AppData Local and Roaming folders. Only emails received via an email client were recovered. Email conversations from the Thunderbird email client were found in the Roaming folder for the specific user. Some additional Yahoo! Messenger files and configuration information can be found in the Local/VirtualStore subfolders. However, conversation log recovery for Yahoo! Messenger is more complex in the latest versions of the tool, as it stores the logs remotely [32]. To obtain access to conversation logs, the investigators will likely need to obtain a search warrant [32] prior to requesting the logs from Yahoo.

System logs are most commonly located in the C:/Windows/System32/winevt folder. In this case, the evidence in system logs helps establish user account activity and software installation, such as Yahoo!, Avast! and Thunderbird. This evidence is more easily available from other sources, so system log files are not of high significance in this investigation.
Obtaining various pieces of information about the offender, such as their online nickname or email address, from other evidence, it was also possible to perform indexed searches to uncover more details, as well as do Web searches to access the offender’s online profiles. For example, an indexed search for the victim’s or offender’s nickname could help recover fragments of conversations or other online activity. Unfortunately, this was not achieved in this experiment, because no conversations were stored on the local machines by the chosen IM software.

B. Offender’s Machine

The findings of investigating the image of the offender’s machine were similar. Majority of user files were contained in the standard user folders, however, in this case they also included the images sent to the victim as part of the stalking behaviours.

Despite the same major location of Internet files, the type of files was different due to a different browser being used on the offender’s machine. For example, bookmarks and visited sites were contained in SQLite databases which are not natively supported by EnCase. However, more temporary files and cached images were found than in the victim’s case. These included evidence of image searches, relevant chatroom and Website access, performed Web searches and other behavioural evidence. In this case, most of the communications evidence was also found in cached temporary Internet files and thumbnails.

In addition to the usual program folders, in this case, examination of other software-related locations and the registry was more useful, as software configurations, such as the last used Proxy IP address, were discovered.

C. Anti-forensics

Unfortunately, the simulated anti-forensic behaviours were only evident on the offender’s machine through their browsing history and installed software and its configurations. Only evidence of the use of a Web-based remailer was found on the victim’s computer. This was obvious in one of the emails received by the victim, which had vendor disclaimers stating that the email was not sent by a legitimate source.

VI. DISCUSSION

The results are limited to the environment of the simulation and the tools and specific applications chosen to carry out the outlined criminal activities. For example, in this case no encrypted or hidden volumes were used, while in other investigations the majority of evidence could potentially reside on such volumes. Additionally, in a real investigation, system, security and application logs would have a greater significance in establishing the event timeline, reconstruction of the crime as a whole, and placing user activity within the relevant timeline. This is especially relevant, if the offender has installed malware on the victim’s computer. Furthermore, in a real case, some of the evidence might be deleted, and will require additional search and recovery efforts.

It can also be seen that some of the locations, such as system event logs and the registry, have not been as useful as expected in this case. However, this is largely due to the limited number of simulated behaviours and the available detail. With different case details and conditions, these locations might be more fruitful. For example, investigating the registry on the victim’s machine can help reveal malware and spyware [13], which was not used in this experiment. In addition, a realistic time-frame was not simulated in this experiment, making the majority of timestamps useless. In a real case, timestamp analysis will be a more significant part of the investigation; hence there will be more benefit in using Windows event logs in crime reconstruction.

The experiment failed to establish signs of anti-forensic measures on the victim’s machine. However, this might be due to the specifics of the anti-forensic measures chosen for simulation. Different types of methods of concealment might have a different effect on the victim’s machine. A repetition of the same experiment with different data collection settings might also produce different results.

It was also not possible to obtain conversation records from the IM client, as expected. Selecting an IM client, that stores conversations locally, would produce different results, and make conversation analysis easier. However, communication storage is rapidly moving to the Cloud, meaning that it will be increasingly difficult to find logs of IM conversations locally. Hence illustration of this aspect can be seen as a strength of the experiment. In addition, in this case, victim behaviours allowed us to recover some aspects of the conversation. This outlines the importance of the victims keeping evidence of cyberstalking.

Finally, the simulated social profiles and the details produced for them were limited, hence allowing for fewer results to be obtained. This is due to the difficulty to simulate long-term interactions with the cyberspace in a short-term time-frame. For example, the victim did not have any friends on Facebook, and had no online presence in terms of online shopping, games and other activities that a typical person is likely to engage in over a long period of time [2]. As a result, online behaviours of the victim and the stalker could not be examined in depth.

All in all, the experiment demonstrates the use of the outlined technique. However, more work is necessary to fully test it and obtain more meaningful results. The current technical profile is minimalistic, but it can be taken further and enhanced in various ways. The simulation should be attempted on different computer systems (in terms of OS, hardware, configuration), including cross-system offences, and also include mobile devices. There is variation to be introduced in terms of choice of software (IM and email clients, Web browsers), social networks and Web applications. This would allow building a more comprehensive collection of common search locations for specific types of software and technology. Furthermore, in future experiments, extra detail should be added to victim and offender profiles, including their connections on social networks. This will allow us to study the evidence from a more realistic, less isolated environment. A larger number of virtual machines could be used to offer better
Validation techniques and simulate group-stalking behaviours. To take it further, physical machines could be used to note any potential differences in produced evidence. Most of all, more complex anti-forensic behaviours should be investigated to account for a higher level of offenders’ technical skill.

VII. CONCLUSIONS

The results of the technical profile enhance the criminal profile by providing it with the technical details unobtainable from existing research or profiling a generic scenario. Additionally, the technical profile also illustrates that technical suggestions for evidence priorities and search locations in the criminal profile are realistic and likely to appear in a real case. The resulting profile can be applied to investigations where the exact behaviours are unknown to locate evidence of actions inherent to email or IM cyberstalking.

Overall, the results show a degree of accuracy to the extent of possible verification. However, there is scope for more work to be done in refining the profile, evaluating it more thoroughly and performing more varied technical experiments.

References


