Digital Investigation 20 (2017) S37–S46

Contents lists available at ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

DFRWS 2017 Europe — Proceedings of the Fourth Annual DFRWS Europe

Do digital investigators have to program? A controlled experiment in digital investigation

Felix Freiling [a, *], Christian Zoubek [b, **]

[a] Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
[b] Department of Computer Science, Technische Hochschule Nürnberg Georg Simon Ohm, Nürnberg, Germany

ARTICLE INFO

Article history:
Received 26 January 2017
Accepted 26 January 2017

Keywords:
Education
Training
Investigation
Investigator
Experiment

ABSTRACT

We report on the results of an exploratory study in which graduate students played the role of digital investigators within an advanced digital forensics course. Overall, 39 students were split up into 10 groups. Each group had to solve one out of three arguably realistic cases within a time frame of 11 weeks. Participants had to log their actions and the corresponding time effort. The resulting data was analyzed in order to identify differences in investigative strategies as well as factors that influence the quality of the results. As can be expected, the total effort (in minutes) generally positively influences the results, but rather surprisingly, participants did not (have to) program to solve the cases although they were restricted to using publicly available tools.

Introduction

It is established practice today that law enforcement bodies delegate (some) physical evidence acquisition and (more commonly) physical evidence analysis to experts, usually working in forensic laboratories. There is a long tradition in documenting the experiences in handling and interpreting physical evidence (Kirk, 1974; Groß and Geerds, 1977; Lee and Harris, 2000). An increasingly large portion of evidence in criminal cases today, however, is *digital* evidence, i.e., evidence that is stored on or transmitted over digital media. When it comes to digital evidence, common police investigators often do not possess sufficient knowledge and training to process and interpret the evidence. So with digital evidence, much work is performed by special police departments that support investigators in evidence acquisition and evidence analysis. In fortunate circumstances, there are special police departments that do not depend on technical advice but rather possess sufficient technical knowledge to *directly* investigate cybercrime cases and therefore perform digital investigations. The work of such (digital) investigators lies in the intersection of digital forensics (Casey, 2011) and criminalistics (Inman and Rudin, 2000).

Skilled digital investigators are extremely rare but they are also very valuable to law enforcement. Consequently, there are many efforts to qualify an increasing number of personnel for this task. While there is much experience in how to teach *technical* skills, less is known about how to effectively teach *criminalistics* skills to digital investigators. One of the most common approaches in practice is to have experienced investigators train instructors ("train the trainer"). However, this further increases the workload of the specialists and does not help to understand what and how to teach. It is therefore not surprising that work of such experts is not yet as well documented as it is with physical evidence. However, the increase of cybercrime demands that good practices in digital investigation be objectively studied such that investigative "rookies" can be better educated and trained.

Related work

In the literature of classical criminology, there is ample work on the *modus operandi* and criminal profiling of offenders, e.g., the well-known Crime Classification Manual (Douglas et al., 2006) or standard works on Criminal Profiling (Turvey, 2011). This work has extended to cybercrime, both in terms of profiling of cybercriminals (Colombini and Colella, 2011) as well as statistical measures that allow educated guesses about future crime (so-called *predictive policing* (Friend, 2013)). There is much less work on the actual way how *investigators* deal with cybercrime cases.

* Corresponding author.
** Corresponding author.
E-mail addresses: felix.freiling@cs.fau.de (F. Freiling), christian.zoubek@th-nuernberg.de (C. Zoubek).

http://dx.doi.org/10.1016/j.diin.2017.01.004

A common approach to study the work of investigators is to employ *case-based reasoning* (CBR) to extract knowledge from old cases that can be applied in new ones (Horsman et al., 2014; Hoelz et al., 2011). This, however, makes it necessary to formalize investigative processing steps as well as all digital evidence. Out of this necessity, formalisms like CybOX (Casey et al., 2015) as well as CBR systems like DIALOG (Kahvedzić and Kechadi, 2009) and FITCASE (Casey, 2013) have evolved. CBR, however, assumes that knowledge is extracted from "good" cases, i.e., cases where investigators performed well. The selection of such cases must still be done by an expert and generally cannot be automated. CBR therefore does not contribute to an understanding of what makes a good investigative strategy. To understand this, it is necessary to study the work of digital investigators in controlled experiments and look at all behaviors, from which good and bad behaviors can be extracted.

Research goal and contributions

We performed a controlled experiment in digital investigation within a graduate level course on digital forensics at Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) in Erlangen, Germany. Overall, 39 students (split up into 10 groups) participated in the experiment. Each group had to solve one out of three arguably realistic cases within a time frame of 11 weeks. Our goal was to observe and analyze the strategies by which they performed their work. In order to do this, participants had to log their actions and the corresponding time effort.

In this paper we report on the results of the analysis of the collected data. Since we were not aware of related work that performed similar experiments before, we could only state rough research questions instead of exact hypotheses to evaluate. Still, the data shows that

- the total effort (in minutes) generally positively influences the quality of results,
- previous grades can (probably) be used to predict the quality of future results while individual motivation is not a good predictor,
- about 50% of the total effort is spent on technical analysis and 30% in documentation,
- it was not necessary to program to solve the cases although participants were restricted to using publicly available tools.

While our findings are limited and can be described as preliminary, we believe that the collected data, which is available online (Freiling and Zoubek, 2017), will be helpful to shape further experiments in this relevant field.

Paper outline

This paper is structured as follows: We first formulate the research questions along which the study was designed in Section Research questions. We then describe the experimental design of our study in Section Experimental design. We report on the first qualitative results in Section Results and conclude in Section Summary and conclusions.

Research questions

Since nothing was known about the factors that influence the process of a digital investigation, we are only able to formulate exploratory research questions instead of exact research hypotheses. Before we develop these questions, we need to define some terminology to make the remainder of this paper more understandable.

Terminology

We use the following terminology: A *case* consists of a description of the case context and the investigative goals, as well as a collection of digital evidence that needs to be analyzed. A *participant* is a human that participates in our experiment. A *group* consists of multiple participants (members) that jointly work on a single case.

The *effort* spent in an investigation is measured in the number of work minutes spent on solving the case. We distinguish four different types of work below. We further separate *individual effort*, i.e., the effort in minutes of a participant, and *group/total effort*, i.e., the sum of all efforts of members of a particular group.

To learn more about investigative strategies, we distinguish four different task types that participants should report on:

- task type T1: conceptual work with pen and paper, including documentation.
- task type T2: group meeting/discussion.
- task type T3: programming new tools and/or interfacing with old ones.
- task type T4: performing the actual investigation by applying tools.

Since we were in control of the cases, we had knowledge of the *ground truth*, i.e., we knew all items of vital evidence that needed to be found to fully solve the case. We define the *result quality* or *grade* as the percentage of vital evidence found and correctly interpreted by the group.

Characterization of different case types

The first line of research questions refers to the ability to characterize different classes of cases. In practice, it is often stated that every case must be treated with a new and fresh view onto the subject matter to prevent overlooking crucial evidence (Casey, 2011). However, it might still be the case that certain types of cases inherently require more effort than others. In this respect, we formulate the following questions:

- Is there a difference between the total effort of the groups to solve different cases?
- Is there a difference between the total efforts of groups that all solved one particular case?

Characterization of investigative strategies

Another area of interest was to observe how each group approached the case and what types of tasks were performed in which order to solve the case. The corresponding research questions are:

- Do groups use different strategies when trying to solve different cases?
- Is the distribution of total effort to individual task types different for different cases and group?
- What types of tools are used to solve different cases?

Factors influencing total effort

In practice, the effort to solve a case is usually the driving factor of cost. Therefore, we are interested in factors that influence total effort, e.g., experience, knowledge, motivation of participants etc. We ask the following research questions: