Digital Investigation 10 (2013) 89–98

Contents lists available at SciVerse ScienceDirect

Digital Investigation

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

Hunting in the enterprise: Forensic triage and incident response

Andreas Moser, Michael I. Cohen*

Google Inc., Brandschenkestrassa 110, Zurich 8002, Switzerland

ARTICLE INFO

Article history:
Received 22 November 2012
Received in revised form 11 March 2013
Accepted 14 March 2013

Keywords:
Triage
Digital forensics
Incident response
Information security
Malware detection
Scalable investigation
Enterprise fleet management

ABSTRACT

In enterprise environments, digital forensic analysis generates data volumes that traditional forensic methods are no longer prepared to handle. Triaging has been proposed as a solution to systematically prioritize the acquisition and analysis of digital evidence. We explore the application of automated triaging processes in such settings, where reliability and customizability are crucial for a successful deployment.

We specifically examine the use of GRR Rapid Response (GRR) – an advanced open source distributed enterprise forensics system – in the triaging stage of common incident response investigations. We show how this system can be leveraged for automated prioritization of evidence across the whole enterprise fleet and describe the implementation details required to obtain sufficient robustness for large scale enterprise deployment. We analyze the performance of the system by simulating several realistic incidents and discuss some of the limitations of distributed agent based systems for enterprise triaging.

Triage is a process commonly applied in the medical field in order to ration limited resources and to maximize their overall effectiveness (Iserson and Moskop, 2007). In the medical arena, responders follow a systemic approach to assess the severity of injuries and the likelihood of successful treatments for medical casualties. Triage is essentially a prioritization process optimizing the usage of limited resources toward achieving the best overall outcome.

Digital forensics is increasingly used in more diverse contexts, such as criminal and civil cases as well as in incident response. Increase in the utilization of digital forensic capabilities has lead to larger case load and longer backlogs of evidence which must be analyzed by a limited number of highly trained investigators (Richard and Roussev, 2006). The increased data volume places challenges on traditional forensic methods and has led to several proposals for updating best practice techniques in order to cope with the work load (Jones et al., 2012).

1. Triage in digital forensics

Triaging has been proposed for managing long case backlogs (Rogers et al., 2006). Drawing inspiration from medical triage techniques (Hogan and Burstein, 2007), the goal of digital forensic triaging is to prioritize evidence for acquisition and analysis in order to maximize case throughput.

Particularly, many digital forensic procedures are designed to discover if evidence exists at all on a computer which was potentially relevant to the case. Drawing an analogy from medical triaging (Iserson and Moskop, 2007), one can define triaging as a process which classifies digital evidence into three groups:

- The system is likely to contain crucial evidence but this evidence is unlikely to be destroyed in the near future.
- The system is likely to contain crucial evidence which may be imminently destroyed.
- The system is not likely to contain relevant evidence and therefore should not be acquired or processed.

* Corresponding author. Tel.: +41 788399834.
E-mail address: scudette@gmail.com (M.I. Cohen).

1742-2876/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.diin.2013.03.003

Classifying potential systems into these categories allows to prioritize evidence acquisition and analysis. Digital evidence collection must follow the Order Of Volatility (OOV) (Farmer and Venema, 2005; Brezinski and Killalea, 2002), in that some sources of evidence are more volatile and likely to change, hence should be collected sooner. For example, memory images are more volatile than disk images, since system state is likely to change quickly (Schuster, 2008), or be completely lost if the system is powered down. Yet the contents of memory contain extremely valuable evidence in many investigations (Walters and Petroni, 2007). In this context, it is critical to obtain the memory image as fast as possible.

It is important to contrast the aims of triaging analysis with traditional digital forensic analysis. While traditional digital forensic analysis aims to establish irrefutable findings upon which a solid case may be built, triage analysis has a much lower evidentiary burden of proof. The triaging step is merely trying to establish whether the system is likely to be involved with the case. This lower burden of proof opens the possibility for wider automation in the triaging process, with a higher acceptable false positive rate. However, the danger with automated triaging is that subtle evidence might be missed.

For example, consider a case where the investigation requires analysis of the URLs the suspect accessed using a browser. A full forensic analysis might require the cache material to be examined, browsing history reconstructed and time lines created. On the other hand, the triaging step merely discovers whether the browser cache contains any references to the user name, or web site in question. Thus, the triaging step can be implemented as a simple keyword search, where a significant number of hits result in classifying the system as potentially containing evidence, leading to acquisition and further analysis of the system.

While triaging analysis is less rigorous than a full digital forensic examination, it must necessarily be applied to many more systems. This can be achieved by recruiting less trained investigators to perform the analysis using standard tools (e.g. recruiting police officers, provided with minimal digital forensic training and using commercial tools). Alternatively, specialized tools may be developed to ensure that triaging analysis is as automated as possible, for example the FBI's image scan tool – a law enforcement only tool used to triage for contraband images (Cantrell et al., 2012). Ideally, the triaging strategy is tailored to the specifics of each case.

1.1. Privacy considerations

Another complication with applying the triaging process is the effect it has on the privacy of the system's owner. In traditional digital forensic analysis, the systems acquired and inspected have a very high probability of being relevant to the case. In criminal matters, these systems must fall under the terms of the relevant warrant before they can be examined at all. Usually, the warrant lists all systems that are to be examined in advance, in order to minimize the chances of examining unrelated systems.

In contrast, triaging affects a wider selection of systems, some of which may turn out to be irrelevant to the case, or unrelated to the suspect. A triage step using too general search criteria may therefore reveal private information irrelevant to the case, as well as select unrelated systems for further inspection with traditional forensic analysis procedures, thus inadvertently violating the owners' privacy. Carefully tailored search criteria lower the false positive rate, resulting in a more focused and effective triage, while simultaneously protecting the privacy of system owners. For example, consider again a search for the presence of specific keywords in the user's web history. A specific and unique term such as an email address or domain name is likely to produce a lower false positive rate than a general term which is likely to occur in unrelated web history.

1.2. Exculpatory evidence

In legal proceedings, exculpatory evidence is any evidence which is favorable to the defendant and might prove the defendant innocence. Legal due process requires that exculpatory evidence be collected and disclosed to the defendant before trial (Supreme Court of United States, 1963).

By its nature, triaging might not collect all evidence, and might systematically omit to collect exculpatory evidence. The digital investigator must keep this requirement in mind, and design triaging processes which include exculpatory evidence collection.

2. Triaging and incident response

Aside from traditional forensic investigations accompanying criminal cases, digital forensics is increasingly employed as part of enterprise incident response procedures. Forensic readiness is defined as the procedures that an organization can take in advance of an intrusion in order to expedite the incident response process (Endicott-Popovsky et al., 2007; Mitropoulos et al., 2006; Tan, 2001).

Enterprise incident response is typically time constrained, requiring rapid collection and analysis of digital evidence (Casey, 2006). For example, when responding to a potential security compromise, the need for acquisition of forensically sound evidence must be balanced with rapid disruption and neutralization of the attacker threat and minimizing the resulting economic loss (Endicott-Popovsky et al., 2007).

In this enterprise context, applying a systemic triaging process is crucial (Lim et al., 2009). Not only does triaging reduce the number of systems which must be manually examined to a manageable level, but triaging also ensures that investigators have the opportunity to acquire forensically sound evidence of systems of value, while maintaining the Order Of Volatility – thus ensuring the possibility for post-incident legal proceedings.

For example, consider a typical network forensic investigation (Casey, 2006). Often, there are several compromised systems under the control of the attacker, all using the same kind of Remote Access Tool (RAT). A triaging procedure might require searching for unique evidence of the RAT in memory, on disk or in the Windows Registry. This might include specific registry keys used by the tool or