Chapter 5

APPLYING MACHINE TRUST MODELS TO FORENSIC INVESTIGATIONS

M. Wojcik, H. Venter, J. Eloff and M. Olivier

Abstract Digital forensics involves the identification, preservation, analysis and presentation of electronic evidence for use in legal proceedings. In the presence of contradictory evidence, forensic investigators need a means to determine which evidence can be trusted. This is particularly true in a trust model environment where computerised agents may make trust-based decisions that influence interactions within the system. This paper focuses on the analysis of evidence in trust-based environments and the determination of the degree to which evidence can be trusted. The trust model proposed in this work may be implemented in a tool for conducting trust-based forensic investigations. The model takes into account the trust environment and parameters that influence interactions in a computer network being investigated. Also, it allows for crimes to be reenacted to create more substantial evidentiary proof.

Keywords: Trust models, forensic investigations, digital evidence

1. Introduction

Digital forensics involves the identification, preservation, analysis and presentation of electronic evidence for use in legal proceedings [1, 10, 14]. Clearly, digital evidence must be trustworthy for it to have any probative value in a courtroom. However, a dilemma arises when an investigator encounters evidence with varying interpretations, some of which contradict each other. In such an instance, a means is needed for determining which evidence can be trusted.

The problem is especially critical when the network containing the evidence in question is running some form of trust model architecture. Such a network allows computerised agents to participate in transactions on behalf of a user to find the most efficient way to conduct these interactions. Thus, it is possible that some of the files (especially sys-

Please use the following format when citing this chapter:

Wojcik, M., Venter, H., Eloff, J., Olivier, M., 2006 in International Federation for Information Processing, Volume 222, Advances in Digital Forensics II, eds. Olivier, M., Shenoi, S., (Boston: Springer), pp. 55–65.

tem files) that may look suspect were in actuality created by the agents executing the trust model and not by human users.

This paper proposes a trust-based model consisting of three phases to address the dilemma. A tool based on this trust model can be run by an investigator to determine the trustworthiness of network nodes and the influence of the trust environment on the files that are created. The term "nodes" in this context denotes devices running a trust model. The trust-based forensic model helps evaluate evidence to determine which evidence can be trusted and which evidence has been tampered with. It also allows for a crime to be reenacted and the evidence to be recreated to produce evidentiary proof that is complete, reliable and believable when admitted in court [14].

This paper is organised as follows. The next section, Section 2, provides an overview of trust models. Section 3 describes the proposed model that integrates key concepts from trust models and digital forensics to enhance digital forensic investigations. Section 4 presents the advantages and shortcomings of the model, and Section 5 provides concluding remarks.

2. Trust Models

New technologies have changed the business world to such an extent that even methods for establishing trust during business transactions have had to be revised to keep up with how transactions are performed. This has led to the formulation of trust models.

Trust is an abstract concept, the exact definition of which is unique for every individual. Trust relies on the formulation of templates for similar situational experiences. This allows an individual to group various experiences and their associated trust representations.

Nooteboom [11] defines trust as a four-place predicate: "Someone has trust in something, in some respect and under some conditions." The individuals participating in a trust relationship in the context of trust models are called agents. Agents, in our work, refer to non-human, coded entities. These coded entities are defined by a programmer and embody logical rules [7] and restrictions against which interactions are analysed and processed to obtain a trust value. A trust value, which is calculated by a trust model, indicates the level of trust one agent has in another. The exact values that indicate trust, distrust and partial trust depend on the specific trust model. The "someone" and "something" in Nooteboom's predicate refer to two agents participating in an interaction. Each agent has some form of trust in the other. The respect under which the trust is given refers to the situational factors that instigated

the need for the transaction and the conditions refer to the limitations under which the transaction occurs.

Trust models [3, 4, 9, 12] are used to analyse the trustworthiness of other agents. This includes the trustworthiness of information shared by agents, since this information is often used to make important decisions.

Trust values are obtained and assigned in various ways. Dynamic means of evaluating an agent and calculating a trust value include observation, experience and negotiation. Observation allows an agent to examine the interactions of other agents before attempting an interaction. Direct experience allows an agent to participate in an interaction and analyse the outcome [5]. Negotiation, on the other hand, requires that two agents share trust-related information contained in their security policies before commencing an interaction [8].

The result of the trust analysis process is a trust value that is used to restrict an interaction. In particular, it limits the information that is shared and it defines the behaviour of the interaction. Higher trust values result in freer interactions and higher trust in the information shared during the interactions.

Since the trust model influences how interactions are conducted, it also influences how information about the interactions is stored. This has a direct influence on forensic investigations because it influences potential evidence of criminal activity. Trust models are also able to determine which nodes are suspect in the trust environment. Such nodes are given distrust values based on their behaviour.

3. Defining Trust in Forensics

An investigator needs to know which evidence can be trusted in order to make sound judgments. This is an issue because criminals may attempt to tamper with evidence to affect its trustworthiness. Tampering, which includes planting false evidence, modifying data or deleting files, may impede evidence gathering as well as evidence analysis.

An investigator looks for anomalies, failures and specific results when running tests on a system. If the information has been tampered with to the extent that anomalies, failures and specific results are not discernible, an investigation can be led away from the source of criminal activity [13]. It is easiest to tamper with evidence contained in user-created files, which are easier to locate, understand and modify. System files often contain a wealth of information. However, system files are typically in obscure locations and protected by the operating system; tampering with these files requires specialised technical knowledge.

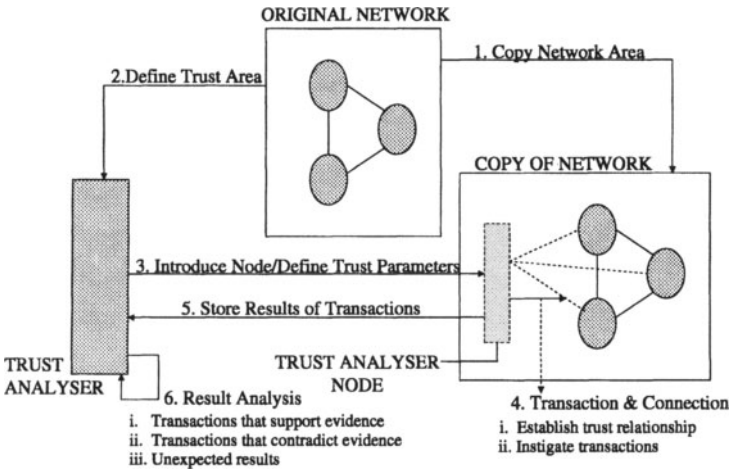

Figure 1. Using trust models to gather forensic evidence.

In a network that relies on a trust architecture, system files are created according to trust rules that govern the processing of interactions. Trust models define the level of trust given to agents participating in interactions. A system file that is created or modified during an interaction between two agents depends on the trust level assigned to the agents and on the nature of the interaction.

The results of these interactions update the state of trust in the system, influencing the processing of future interactions [6, 16]. Keeping this in mind, it is possible to test for criminal activities that have occurred over a network by testing the state of the trust relationships within the network and the reactions of various nodes to similar activities.

Figure 1 presents a scheme for using trust models to determine the presence and trustworthiness of forensic evidence. The numbered and labelled arrows in Figure 1 indicate processes that occur on four logical components. The first component is the original network in which devices containing evidence of criminal activity exist.

The second logical component is the copy of the original network. This is made by copying data from suspect devices and the network. To preserve the trust environment, the copy includes copies of devices that surround the suspect devices. It is important to note that only certain portions of the original network are copied for simulation and reenactment purposes. Therefore, it is important that these portions be carefully selected to include all the devices that may have influenced the

criminal activity. Each device will have files, usually some form of text files created by a particular trust model, that relate to the trust model in place. These files determine how each device reacts to trust-based interactions. The copy also contains the current state of these files.

A copy of the original state of system files (before the trust model was run) can assist in predicting the behaviour of the system and its new state. Although these files could contain trust-related data that may have been tampered with, the investigator should realise that any interaction that has occurred influences all the agents that participated in the interaction. For instance, tampering is to be suspected if one agent that participated in an interaction is found to trust another agent unconditionally while the other agents involved in the interaction show levels of distrust. To leave no traces, a criminal entity would have to tamper with all the agents that may have directly or indirectly participated in the interaction.

The third and fourth logical components, the trust analyser and a logical trust analyser node, make up one physical component, the trust analysis unit. The trust analyser is an investigation tool to be used in trust environments when the influence of the prevailing trust environment is to be determined. The logical trust analyser node is created by the trust analyser and is introduced into the copy of the network to act as an additional node in the network.

The proposed trust model for digital forensics has three phases: the establishment phase, the evidence gathering phase and the analysis phase. The first three processes in the Figure 1 are part of the establishment phase. Processes 4 and 5 are part of the evidence gathering phase, and Process 6 is a phase on its own (analysis phase). The establishment phase sets up the necessary criteria and environmental variables to conduct an investigation. The evidence gathering phase actively gathers evidence for analysis, while the analysis phase produces a conclusion based on the evidence.

3.1 Establishment Phase

The establishment phase begins with the identification of evidence. This phase is paramount as it influences the progress and results of all the following phases [14].

Once an investigator has identified what evidence is present and how it is stored, the evidence must be isolated and collected. This process must not damage or cause any loss of evidence. Also, it should allow for the evidence to be analysed to acquire the relevant information [15].

The first step required by the model is to copy the data and network settings from suspect devices. This is done to preserve the original state of the evidence. The model makes use of distributed computing to ease the computational load. This is implemented by duplicating the network on several machines to simulate the "live" state of the network on a so-called "dead" copy. However, it is not necessary to have a physical machine for every physical machine in the target network. Several subsections of the network can be duplicated on a single machine. Suspect nodes are placed on their own computer; nodes that support suspect nodes are grouped with the nodes they support to create a sub-domain. The use of multiple machines allows for the duplication of some of the more vital physical links. Also, it provides for a more accurate representation of the live state of the original network for subsequent analysis.

After the network area of interest has been duplicated, an analysis is conducted to determine the appropriate trust attributes. One way of determining the trust attributes is to query the people involved with network setup. Should this not be possible, the information can be gleaned from the network itself by searching for global policies that have been defined. The investigator should also be able to directly access the list of rules governing trust from any of the physical nodes in the network. Whether trust is built by reputation, observation or direct interaction with the new nodes depends on the rules that influence the prevailing trust environment. This process must be done with as much, if not more, care as making the copy of the system: it is important that no changes are made to the system state while extracting information. The rules are input into the trust analyser as text documents.

Various activities linked to the suspected crime must be defined as a set of attributes. A crime involving an information leak could have attributes corresponding to the manner in which the information was leaked and the confidentiality level of the leaked information. These could be represented as values and logical rules. For instance, the confidentiality level of information could be a set of values and the means by which the information was leaked can be indicated as parameters. For example, if email is the medium for leakage, the parameters would be the sender's and recipient's addresses. The trust analyser uses these attributes to attempt to recreate the crime.

The trust analyser uses the rules and attributes to define a virtual node for the network that runs according to the same rules and attributes defined by the network and trust environment. This virtual node is then introduced into the copy of the network where it is required to run and gather trust-related information.

3.2 Evidence Gathering Phase

The goal of the evidence gathering phase is to obtain information of value to a forensic investigation. The evidence must be gathered under the restrictions placed by the establishment phase. The virtual node introduced into the copy of the original network controls this process.

The driving force of the evidence gathering phase is the transaction and connection process, which is made up of two key sub-processes: trust establishment and transactions. Trust establishment takes into consideration the trust area and trust parameters received from the establishment phase. It uses this information to establish communication links with the other nodes. This establishes the trust levels between nodes and ensures that the new node is governed by the same context as the original nodes in the suspect network. The new node, therefore, instigates transactions in the same manner as nodes in the original network.

To successfully recreate the evidence, it is important to have a clear definition of the suspected crime and the context in which the crime occurred. Both of these factors are derived from the establishment phase. Once the trust context has been established, the virtual node conducts a detailed analysis of the attributes that are related to the suspected crime. These attributes are used to deduce interactions that should have occurred for the suspected criminal activity to take place.

The transactions sub-process makes use of the already-established trust connections to recreate the forensic evidence that is being questioned. It involves the recreation of events that created the suspect evidence to test whether the results correlate with the suspected crime. For example, the trust analyser may attempt to send confidential information outside the network and examine how this behaviour changes the trust environment.

The responses of the system to the various transactions are recorded and passed back to the trust analyser node. After all the transactions have been finalised, data created by the various nodes, including that created by the virtual node, is collected and returned to the trust analyser for detailed analysis, which occurs in the final (analysis) phase. This data is representative of the system's final state after the transactions have occurred and is used for comparisons with the original evidence.

3.3 Analysis Phase

During the analysis phase, the results are gathered and investigated to reach a conclusion. The trust analyser is supplied with machine-generated data created by the nodes in the network as a result of the

transactions instigated by the virtual trust analyser node. This machine-generated data is compared with other machine-created data that constitutes evidence of a specific crime.

The results are analysed along with the trust rules of the system to determine how the prevailing trust environment influences the representation of the collected evidence. The influence is taken into consideration during the more detailed analysis phase. The analysis may produce one or a combination of three different sets of results: results supporting the evidence, results contradicting the evidence and unexpected results. Various conclusions may be drawn from these results.

The results that support the evidence and contradict the evidence are dependent on the fact that the investigator is expecting certain evidence to correlate and other evidence to contradict. If the results are what the investigator expects, he/she has a means of proving that the suspicions are true. Results that correlate are indicative of a successful recreation of a crime and can be used with the original evidence to prove that the suspected crime did indeed occur. Results that are expected to be contrary, perhaps due to a suspicion that data was tampered with, also support a given theory.

Unexpected results can be scrutinised in two ways depending on an investigator's initial outlook. These results include those that were expected to correlate and do not, and those that were expected to be contradictory but in fact correlate. If the investigator believes his/her theory to be sound and is certain about what results would support the theory, unexpected results could mean that the investigator's entire theory and suspicions are incorrect. The investigator would then have to re-evaluate the evidence and consider alternative possibilities.

If the initial outlook was uncertain as to which evidence is to be trusted and which is to be disregarded, the model is only run until the trust relationships have been established according to the trust parameters in place. For instance, if a recommendation-based trust model is employed, the establishment of trust relationships relies on recommendations from trusted nodes. To recreate the environment as faithfully as possible, nodes that trust the suspect node are modified to trust the new virtual node to a similar degree. The investigator needs to be aware that sometimes the trust value may have to be rolled back to a different value that has since changed due to the effect of the criminal-related transactions on the trust environment itself. The degree to which the state of an environment can be rolled back depends on the prevailing trust model and requires further investigation.

Next, the trust relationship values between the virtual trust analyser node and the other nodes in the network are analysed. This is a fairly

simple concept as the trust relationships between nodes are often represented as single values. Nodes given a high trust value by the virtual trust analyser node are considered to be more trustworthy than those with lower values. Thus, the evidence contained in these nodes has a higher probability of being trustworthy. Trust models only allow a transaction to take place if the nodes participating in the transaction are trusted; otherwise, the transaction would not have taken place.

4. Discussion

The model proposed in this paper can be used by investigators to determine which evidence can be trusted and which evidence is suspect. Also, it can help recreate the crime and provide supporting evidence for the suspected crime. Note, however, that the model is preliminary in nature, and substantial research is required before it can be used in digital forensic investigations.

This model assumes that the network being examined for evidentiary purposes has certain trust mechanisms in place that control the interactions occurring in the context of the network. However, the model should also be applicable to networks that do not have explicit trust architectures in place. In such instances, the process of defining the trust parameters and trust environment will change. Instead of defining a trust model as in a network with trust mechanisms, a default trust context will have to be employed. Further research is necessary to define appropriate default contexts.

This model also assumes that the trust mechanisms work and have not been subverted by a criminal. It is necessary to examine how trust mechanisms might be subverted. The fact that trust models and their workings vary must be taken into account while researching this issue.

The reenacted transactions must be similar to those involved in the suspected crime. However, these transactions must be conducted carefully so that they do not modify data left by the original crime, but only add to it. Should the investigator find that the original data was altered during a reenactment, the transactions used to recreate the crime must be analysed and controlled more carefully. This is an interesting area for future research.

Substantial resources may be needed to conduct investigations on large networks. To reduce the complexity and investigative overhead, an investigator has the option of copying only critical portions of a large network and running the tool on those portions.

5. Conclusions

The trust-based forensic model proposed in this paper is intended to help evaluate forensic evidence to determine which evidence can be trusted and which evidence has been tampered with. This model also allows for crimes to be reenacted to create more substantial evidentiary proof.

The three phases of the model have been investigated from a conceptual point of view. More research is necessary to explicitly define what happens in each phase and how it should be accomplished. Areas that warrant attention are how the network may be copied to preserve the prevailing trust environment, how protocols will work on the network copy and how to explicitly define the crime activities being reenacted.

An interesting dilemma arises when a computerised agent is able to actively instigate transactions on behalf of a user. In such an environment, an agent is given rights to participate in transactions without the user's direct knowledge. Investigations must take into account the fact that criminal activity could have been caused by a code flaw or by a malicious act by the programmer of the agent code and not directly by the user. Methods for testing and proving code must be evaluated and incorporated in the proposed model.

Acknowledgements

This material is based on work supported by the National Research Foundation (NRF) under Grant No. 2054024. Any opinions, findings, conclusions and recommendations expressed in this material are those of the author(s) and, therefore, the NRF does not accept any liability thereto.

References

[1] V. Baryamureeba and F. Tushabe, The enhanced digital investigation process model, presented at the *Digital Forensics Research Workshop*, 2004.

[2] S. Bui, M. Enyeart and J. Luong, Issues in computer forensics (www.cse.scu.edu/~jholliday/COEN150sp03/projects/Forensic%20Investigation.pdf), 2003.

[3] M. Carbone, M. Nielsen and V. Sassone, Gigascope: A formal model for trust in dynamic networks, *Proceedings of the First International Conference on Software Engineering and Formal Methods*, pp. 54-61, 2003.

[4] M. Coetzee and J. Eloff, Towards web services access control, *Computers & Security*, vol. 23(7), pp. 559-570, 2004.

[5] B. Esfandiari and S. Chandrasekharan, On how agents make friends: Mechanisms for trust acquisition, *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, pp. 27-34, 2001.

[6] C. Jonker and J. Treur, Formal analysis of models for the dynamics of trust based on experiences, *Proceedings of the Ninth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (Lecture Notes in Computer Science, Volume 1647)*, pp. 221-232, Springer, Berlin, Germany, 1999.

[7] A. Josang, Prospectives for modeling trust in information security, *Proceedings of the Australasian Conference on Information Security and Privacy (Lecture Notes in Computer Science, Volume 1270)*, pp. 2-13, Springer, Berlin, Germany, 1997.

[8] L. Kagal, T. Finin and A. Joshi, Trust-based security in pervasive computing environments, *IEEE Computer*, vol. 34(12), pp. 154-157, 2001.

[9] M. Marx and J. Treur, Trust dynamics formalized in temporal logic, *Proceedings of the Third International Conference on Cognitive Science*, pp. 359-363, 2001.

[10] R. McKemmish, What is forensic computing? *Trends and Issues in Crime and Criminal Justice*, no. 118, Australian Institute of Criminology, 1999.

[11] B. Nooteboom, *Trust: Forms, Foundations, Functions, Failures and Figures*, Edward Elgar Publishing, Cheltenham, United Kingdom, 2002.

[12] M. Patton and A. Josang, Technologies for trust in electronic commerce, *Electronic Commerce Research*, vol. 4, pp. 9-21, 2004.

[13] S. Peron and M. Legary, Digital anti-forensics: Emerging trends in data transformation techniques, presented at the *E-Crime and Computer Evidence Conference*, 2005.

[14] K. Ryder, Computer forensics: We've had an incident, who do we get to investigate? (www.sans.org/rr/incident/investigate.php), 2002.

[15] A. Svensson, Computer Forensics Applied to Windows NTFS Computers, Master's Thesis, Stockholm University/Royal Institute of Technology, Stockholm, Sweden, 2005.

[16] L. Xiong and L. Liu, A reputation-based trust model for peer-to-peer e-commerce communities, *Proceedings of the Fourth ACM Conference on E-Commerce*, pp. 228-229, 2003.