A cyber-crime investigation framework

Article in Computer Standards & Interfaces · May 2008
DOI: 10.1016/j.csi.2007.10.003 · Source: DBLP

CITATIONS 16
READS 2,249

2 authors:

Vasilios Katos
Bournemouth University
77 PUBLICATIONS 242 CITATIONS

Peter M. Bednar
Lund University
133 PUBLICATIONS 355 CITATIONS

Some of the authors of this publication are also working on these related projects:

contextual inquiry into complex, ambiguous and uncertain problem spaces View project

International Conference ICTO2017 – ICT for a better life and a better world, Paris March 16-17, 2017.
View project

All content following this page was uploaded by Vasilios Katos on 29 December 2016.
The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.
A cyber-crime investigation framework

Vasilios Katos a,⁎, Peter M. Bednar a,b

a School of Computing, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth PO1 3HE, UK
b Department of Informatics, Lund University, Ole Römers väg 6, 223 63 Lund, Sweden

Available online 14 October 2007

Abstract

Epistemic uncertainty is an unavoidable attribute which is present in criminal investigations and could affect negatively the effectiveness of the process. A cyber-crime investigation involves a potentially large number of individuals and groups who need to communicate, share and make decisions across many levels and boundaries. This paper presents an approach adopting elements of the Strategic Systems Thinking Framework (SST) by which conflicting information due to the unavoidable uncertainty can be captured and processed, in support of the investigation process. A formal description of this approach is proposed as a basis for developing a cyber-crime investigation support system.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Strategic systems thinking; Dempster–Shafer Theory; Cyber-crime scene

1. The expanding crime scene

Although that there are commonalities between a cyber and a physical crime scene, there are also significant differences [3], making the topic of cyber-crime an important area of research. This has been acknowledged by European governmental agencies, see for example the UK’s Parliamentary Office of Science and Technology Report [10]. The main difference is that the boundaries of a digital crime scene are not clearly outlined and the crime scene area may extend beyond a room, a city, a country, or even a continent. For instance, a computer virus outbreak may impact a large proportion of computers connected to the Internet. Formally, the crime scene would then be defined as the area that includes all infected computers. Alternatively, an identity thief or a paedophile may use a remote server to host illegal material and that server may be miles away from the person’s physical location. In this case, the crime scene would be the person’s physical location, the remote server, and the network paths that the relevant network protocols utilise.

It should be highlighted that the term ‘scene’ in this paper is used in the investigator’s sense. More specifically, a ‘crime scene’ for an investigator would be any area where they believe they may be able to identify facts or evidence which they can produce to a court or from which other evidential inferences may be made, whereas for a lawyer, the ‘scene’ is determined by the definition of the crime and is therefore conceptual. Therefore the purpose of the crime scene investigator is the opposite of that of the lawyer, as the investigator starts out with an ill-defined crime scene and one of the objectives is to conclude to a defined crime scene. The challenge for the investigator is not only to deal with the constantly evolving and diversifying technologies, but also to deal with an increasingly complex problem space. In systems research and practice there is a long tradition to deal with issues related to inquiries into complex problem spaces [1,2,8]. This research has concluded that conventional theory bears little relation to the experiences most people have in their real-world practice in organisations. While there has been a continuous technical development in contemporary society, the development of concepts which bring structure and meaning to people lags behind. There is a need to support professional analysts in discussions which allows for sense-making models to develop which can include any particular assumptions about the nature of their experienced ‘reality’.

The contribution of this paper is the development of a formal framework to aid cyber-crime investigations and more specifically a forensic investigation. If we consider T0 to be the point in time where the crime is committed, the crime scene space is infinite and the boundaries start appearing at time Td relating to the time where the crime is detected. Formally, a forensic investigation focuses mainly on the space defined by Td−T0, where a chronology of
events is projected on a universe of discourse in order to produce a (static) Dempster–Shafer frame of discernment, $\Theta$ and this paper describes an approach to perform this.

From the above it is clear that once there is a suspicion that a cyber-crime is committed and the relevant enforcement agency is aware of the event, there is a likelihood that expert investigators with different roles (such as police enforcement agents, forensic investigators, system administrators, IT specialists, and so on) would need to form groups and contribute with their expertise and information they gathered in order to understand the modus operandi of the cyber criminal. In addition, legal constraints and policies, both national and international will influence the decision making process and the information flows on different levels, including individuals within groups as well as between groups.

The purpose of this paper is to present an information system capable of capturing the information provided by the different members during a cyber-crime investigation in such a way that:

- conflicting ideas will not be mutually cancelled or demoted;
- communication can take place on different levels (between individuals, groups, and super-groups);
- policy and legal constraints could be trivially incorporated;
- the progress of the information exchange is monitored. This suggests an existence of metrics, which are also presented.

The paper is organised as follows. Section 2 presents a brief overview of the relevant elements of the Framework for Strategic Systems Thinking (SST). In Section 3 a formal description of a selection of components of the SST is presented by using Dempster Shafer Theory of Evidence and showing how the latter is an ideal tool for meeting the requirements stated above. Finally Section 4 presents the conclusions and areas for further research.

2. The SST framework

Human beings have no difficulty in keeping contradictory understandings in mind whilst considering resolutions in everyday life - whether complementary, alternative or incompatible. However, traditional logic upon which e.g. decision support system and smart software are built does not reflect this human capacity. Such logic has difficulty in dealing with the maintenance of underlying contradictions as valid parts of resolutions [8].

The SST framework [1,8] represents a systematic attempt to support systemic inquiry into uncertain and complex problem spaces. It involves exploring a problem space from each participating expert investigators’ unique perspectives, both separately and in group contexts. One outcome of such an inquiry is a collection of expressions of inherently contradictory resolutions. The complex process and the amount of data involved in such inquiries can make the whole analytical task overwhelming.

As shown in Fig. 1, the framework for Strategic Systemic Thinking (SST) consists of three main aspects: intra-analysis, inter-analysis, value-analysis. In the intra-analysis each investigator makes an effort to consolidate and develop descriptions of the problem space from their own unique perspective. This is done in the form of exchange and development of ‘narratives’ using methods such as brainstorming, mind maps and rich pictures. In the inter-analysis these narratives are then explained to and exchanged with other investigators. Finally in the value analysis the focus by the participating experts is put upon dialogues. Here, participants attempt to develop understandings of the conditions under which any one narrative may be acknowledged as valid or acceptable.

An analysis is grounded in each participating expert investigators’ efforts to inquire into, create and share knowledge, by creating and exchanging messages — or equivalently, hypotheses. These hypotheses are derived from different perspectives and therefore, if truthful, could contain contradictions if combined. When applying an analysis based upon a methodology such as the SST framework, a human expert may, in practice, take these contradictory matters into account, and can thus follow through the whole complex analytical process of investigation. This is, nevertheless, a challenging task.

In practise complex inquiries built on a model such as the SST framework, there will be many different intra-analyses from different perspectives. Not only is each analysis done by different experts, but also each individual expert may have several, and sometimes incompatible, perspectives. Even though perspectives may be incompatible they could still all be justifiable as reasonable alternatives. As each individual expert makes efforts to develop her/his own understandings about relevant problem spaces, many ‘hypotheses’ will be created. These hypotheses are later used as a basis for further elaboration, as part of self-reflection and sharing. This ‘sharing’ takes the form of storytelling (e.g. exchange of narratives). In the inter-analysis there is a conscious exchange of hypotheses for the purpose of knowledge sharing, knowledge creation and rationalization.

The rationalization aspect comes about through a purposeful classification of hypotheses as messages (narratives). Such a classification exercise is based upon negotiation regarding what characterizes each narrative. Examples of four types of narratives are: compatible, incompatible, complementary or unidentified. This classification exercise is not intended to bring about exclusion of alternative (for example ‘incompatible’) perspectives or narratives. The purpose of inter-analysis is rather to widen understandings of different perspectives - no single alternative is excluded no matter how ‘different’ or ‘crazy’ (e.g. estranged to the majority) they may seem. The result is not only rationalization, (similar narratives are grouped to limit the number of alternative stories but not their ‘scope’) but also further complexification and acceptance of contradictions. In the value-analysis, the participating expert investigators are elaborating and reflecting upon hierarchy and
priority. Here again, it is not intended to create a consensus or compromise, but to understand the complex diversity of the problem spaces and their scope.

The whole purpose of complex inquiry, in our view, is not ‘to make’ decisions but ‘to be able to make’ informed decisions. The purpose is to enrich the foundation upon which decisions could be made while still keeping an overview of available ‘knowledge’.

3. Representation of the SST methodology

Traditional probability theory is handicapped in the sense that it cannot capture and represent events in an uncertain domain. That is, probabilistic analysis requires that the probability distributions are known for all events. This limitation was initially addressed by Dempster [5] and further refined by Shafer [11]. According to the Dempster Shafer mathematical theory of evidence (DST), classical probability is extended in such a way that events can be described at a higher level of abstraction, without requiring one to resort to assumptions within the evidential set. Furthermore, Dempster and Shafer developed an algebra to combine events and produce measures for events that can be contradictory. Classic probability could be viewed as a special case of DST.

In DST hypotheses are represented as subsets of a given set. A hypothesis is a statement which holds with some probability. An interesting feature in DST is that the probability assigned to a hypothesis need not be calculated or proven in the classic probability sense. Therefore, a probability can be a person’s view on the validity of the respective hypothesis.

In the SST a participant is at some stage presented with a series of hypotheses created by themselves or others. Let \( \Theta \) denote the set of mutually exclusive alternatives. This set is the frame of discernment. The powerset \( 2^\Theta \) would then contain all subsets of \( \Theta \). By considering the powerset, we are able to form complex hypotheses based on the building blocks given by the frame of discernment. Furthermore, the DST includes three measures to represent opinions on hypotheses and express their uncertainty. This is done by considering the mass assignment function \( m : 2^\Theta \rightarrow [0,1] \), which assigns probabilities to any subset of \( 2^\Theta \).

There are two restrictions for the mass assignment function:

\[
\sum_{A \subseteq \Theta} m(A) = 1
\]

\[
m(\emptyset) = 0
\]

In addition to the mass assignment function, there is also the belief \( \text{Bel} : 2^\Theta \rightarrow [0,1] \) and the plausibility \( \text{Pl} : 2^\Theta \rightarrow [0,1] \) function:

\[
\text{Bel}(A) = \sum_{B \subseteq A} m(B)
\]

\[
\text{Pl}(A) = \sum_{B' \supseteq A \cap \emptyset} m(B)
\]

It can be seen that these two measures are related as follows:

\[
\text{Bel}(A) = 1 - \text{Pl}(\complement A)
\]

\[
\text{Pl}(A) = 1 - \text{Bel}(\complement A).
\]

It should be noted that \( \text{Bel} \) and \( \text{Pl} \) are non-additive; intuitively this is correct, since there is no reason to require that the sum of all the Belief and Plausibility measures to be 1.

3.1. Combination of evidence

During the intra-analysis stage, each investigator contributes with his beliefs about the validity of the hypotheses. From the perspective of the SST framework, the intra-analysis stage is an exercise in combining evidence contributed by different parties. In DST the combination of evidence is performed by applying Dempster’s rule:

\[
m(A) = \sum_{X \cap Y = A} m_1(X) \cdot m_2(Y)
\]

where \( m \) is the resulting mass assignment function and \( m_1 \) and \( m_2 \) are the mass assignments of the original evidence. An interesting aspect about Dempster’s rule is that hypotheses which are originally considered to be unlikely may result to having high levels of belief. This has been criticised and considered undesirable, so further research in different combination rules was undertaken [6, 7, 12, 14, 15]. However, in the case of the SST methodology this property is highly desirable, thus making the ”original” DST a suitable candidate. A low probability does not necessarily vouch for the quality or appropriateness of a hypothesis. Many effective ideas were originally considered to be radical, controversial, or outside the norm. In traditional systems thinking, such ideas are ignored, as they are considered infeasible, inappropriate or glitches of the underlying analysis process. On the contrary, in the SST philosophy there is a systematic effort to highlight and explore these ideas and DST seems to provide the mathematical tools to enable this. All these characteristics strengthen the adoption of the SST with DST for a crime investigation, as criminals often think ‘outside the box’, contributing to high uncertainty. Therefore, the candidate framework should be capable of highlighting ‘extreme’ hypotheses.

3.2. The open world assumption

One of the key characteristics of the SST framework is the ability to operate in an open problem space. In DST terms, this means that there can be conflicting evidence which may suggest information which was not considered. This is detected by using the Dempster’s combination rule, where the resulting mass assignment for \( \emptyset \) is nonzero:

\[
m(\emptyset) > 0.
\]

There are two approaches to rectify this, depending on the view of the world. If a closed world is considered, then the nonzero mass allocated to \( \emptyset \) must be removed and redistributed to the remaining elements. This is typically done by normalising Dempster’s calculation:

\[
m(A) = K \sum_{X \cap Y = A} m_1(X) \cdot m_2(Y)
\]

\[
K^{-1} = 1 - \sum_{X \cap Y = \emptyset} m_1(X) \cdot m_2(Y)
\]

with \( A \neq \emptyset \). If an open world is considered, we then accept the violation, as it is interpreted that the solution exists in a
hypothesis which we haven’t considered. It is obvious that in the case of a criminal investigation framework, an open world view should be adopted. This can also be seen if we consider the three aspects of the SST: intra-analysis, inter-analysis, and value-analysis. In intra-analysis an expert investigator creates and explores his personal perspectives. Although this alone could be viewed as a closed world, at some stage these perspectives will be challenged during the inter-analysis stage. From an individual’s perspective, the inter-analysis experience is an attempt of hypotheses external to his world to enter and influence the frame of discernment.

The normalisation coefficient $K$ in the above equation can give discriminative information. The quantity $\log K$ is called weight of conflict between $\text{Bel}_1$ and $\text{Bel}_2$ [9]. Therefore, although in an open world assumption the combination of rules should not be normalised, $K$ must be calculated in order to assist to understanding the conflict. Following this convention, a nonzero mass assignment of the empty set would suggest that a solution exists outside a person’s frame of discernment.

3.3. Formal description of the system

The proposed methodology is based on the following assumptions:

- There exists an agreed coding and representation of the hypotheses. This is required in order to avoid both ambiguities and redundancy, and to enable a more successful application of the combination rule(s).
- There can be any order of execution of the three main stages (inter, intra, or value analysis), as well as any number of iterations between them. This would allow the framework to integrate with a variety of protocols.
- It is possible and anticipated for each party’s frame of discernment to change over time. This will have a direct impact on the cardinality of the powerset. Although it is expected that the frame of discernment will inflate due to interaction with the world outside the frame and accumulated experience, there is no reason to disallow deflation.
- All assumptions surrounding the combination of evidence based on DST must also be inherited by the proposed methodology. More specifically, the combination rules are applied on evidence that is supplied by independent sources and on the same frames of discernment. This can be particularly troublesome especially when it is considered that the frames of discernment may change over time. An approach to address this is presented below.

Fig. 2 presents an outline of the investigation process in DST terms. For simplicity, the temporal indexes have been omitted. The three stages can repeat at any sequence, i.e., each of the stages can precede any of the other two. The reason is that all investigations are initiated by some prior event. This event could be relating to historical data, other ongoing investigations and newly discovered phenomena.

In order to formally describe the SST processes, the following notation is considered. Let $n$ be the number of participants, each of them identified by $U_i$, $1 \leq i \leq n$. Let $\Theta_{Tk}^{U_i}$ denote the frame of discernment of investigator $U_i$ at time $T_k$, $k=0,1,\ldots$. The respective mass distribution function of $U_i$ at time $T_k$ is represented by $m_{Tk}^{U_i}(\cdot)$. Let $\Theta_{Tk}^g$ denote the collective frame of discernment of the group at time $T_k$ with $m_{Tk}^g(\cdot)$ being the mass distribution function. Furthermore, we consider the following metric [8]:

**Definition 1.** For any time $T_k$ where $k > 0$, $\delta_X^{T_k} = \#\{ (\Theta_{Xk}^{T_k} \cup \Theta_{Xk}^{T_k-1}) \cap (\Theta_{Xk}^{T_k-1} \cap \Theta_{Xk}^{T_k-1}) \}$ is the $\Theta$-osmosis of $X$’s frame of discernment.

The $\Theta$-osmosis metric is used in the group or investigator maturity metric to indicate how close the analysis process is to completion or the contribution required by an investigator [8]:

**Definition 2.** For any time $T_k$ where $k > 0$, $X$’s maturity is defined as:

$$\omega_X^{T_k} = e^{-c\delta_X^{T_k}}$$

where $c \in (0,1)$ is a constant.

Although the analysis using the SST approach can start at any of the three stages, the initial problem must be captured and

![Fig. 2. DST within the SST framework.](image-url)
expressed as the collective frame of discernment. Therefore \( \Theta_0 \) needs to be agreed upon. This also favours the first assumption stated above - an agreed coding and representation of the hypotheses. In addition, the initial mass distribution is set to:

\[
m^0_g(A) = \frac{1}{\# \{\Theta^g \} - 1} \quad \forall A \subseteq \Theta^g \setminus \{\emptyset\}
\]

\[
m^0_g(\emptyset) = 0
\]

to start at a maximum entropy. In each of the three main steps, the following actions can be performed:

**Intra-analysis.** An investigator \( U_i \) can join and leave at any stage. For the sake of simplicity, we assume that the investigator joins at the very beginning of the analysis process, i.e. at time \( T_0 \). The investigator would then have the opportunity to assign beliefs and contribute with evidence on \( \Theta^T_{U_i} = \Theta^T_0 \) and its power set, in order to produce \( \Theta^T_{U_i} \). This may result in an inflation of the collective frame of discernment.

If \( k > 0 \), then the new frame \( \Theta^T_{U_k} \) as well as the mass function are calculated as follows:

- If \( \Theta^T_{U_{k-1}} \subseteq \Theta^T_{U_i} \) then \( \Theta^T_{U_k} \subseteq \Theta^T_{U_i} \) and the new evidence \( m(U_k) \) they are combined with \( m^T_{U_{k-1}} \) to produce \( m^T_{U_k} \).

- If \( \Theta^T_{U_{k-1}} \nsubseteq \Theta^T_{U_i} \) then there are events that have been introduced by other investigators into the group’s frame of discernment. In this case \( \Theta^T_{U_k} = \Theta^T_{U_{k-1}} \cup \Theta^T_{U_i} \) and the resulting mass assignment would be:

\[
m^T_{U_k}(A) = \begin{cases} m^T_{U_{k-1}} & \text{if } A \subseteq \Theta^T_{U_{k-1}} \\ 0 & \text{otherwise} \end{cases}
\]

Finally the investigator may introduce new events in order to enrich the list of hypotheses. The calculations for the expansion of the frame of discernment are similar to those in the intra-analysis stage which are explained in detail in the next paragraph.

**Inter-analysis.** During the inter-analysis stage, the pieces of evidence of the different investigators are combined into the collective mass assignment. At this stage it is likely that each individual during their intra-analysis process may have discovered new events and have enriched their personal frame of discernment. Therefore \( \Theta^T_{g_{k-1}} \subseteq \Theta^T_{U_i} \). By considering an open world assumption, we can allow the mass assignment of the empty set \( m_g(\emptyset) \) to include all focal points of \( \Theta^T_{U_i} \) that fall outside \( \Theta^T_{g_{k-1}} \), i.e.

\[
m_g(\emptyset) = \sum_{A \subseteq \Theta^T_{U_i} \setminus \Theta^T_{g_{k-1}}} m^T_{U_i}(A)
\]

whereas for the remaining values:

\[
m_g(A) = m^T_{U_i}(A), \forall A \subseteq \Theta^T_{g_{k-1}}.
\]

This modified mass assignment \( m_g() \) will be combined with \( m^T_{g_{k-1}}() \) instead of \( m^T_{U_i}() \), to produce \( m^T_g() \). The process is repeated for all investigators, \( U_i, 1 \leq i \leq n \).

Finally, the collective frame of discernment would be enriched with the additional elements of the participants, defining it at time \( T_k \):

\[
\Theta^T_g = \bigcup_{1 \leq i \leq n} \Theta^T_{U_i}
\]

and the mass assignment of the empty set could be distributed to the new members of the power set. The simplest way of doing this would be to arrange a uniform distribution, following the maximum entropy consideration. More complex arrangements would require taking into account the historic values at time \( T_{k-1} \) of the investigators’ mass assignment functions, but this is not investigated in this paper and is recommended for future research.

At the end of this stage the collective frame of discernment may have increased in size and therefore a \( U_i \) may exist where \( \Theta^T_{U_i} \nsubseteq \Theta^T_{g_{k-1}} \).

**Value-analysis.** During the value analysis stage, the relevant constraints are incorporated into the mass assignments. This is done by creating a list of hypotheses from the current powerset which cannot be accepted due to political, procedural, or other constraints and setting their masses to zero. This stage is completed once the following actions are performed:

- the masses which are removed from the unacceptable hypotheses are distributed uniformly to the remaining focal points;
- If the group maturity \( g_{T_{k-1}} \) is above a certain threshold (i.e. close to 1), then all the events from the current \( g \)'s that are not found in any of the focal points of \( m^T_{g_{k-1}} \) are removed and the resulting \( g_{T_{k-1}} \)s will have a smaller cardinality. Although that removal would not affect the mass assignment, it should be done once the group has reached an accepted maturity; as if it is done at an earlier stage, it may have an impact on the initial ideas and hypotheses. This is in line for example with problem solving approaches that use brainstorming; during the initial stages of brainstorming, the ideas are not criticised - they are captured to be evaluated at a later stage.

The maturity metric could also be considered for measuring the efficiency of the investigation. For example, if the maturity is high for a certain time period, the involved parties should expect to end, hold or reduce the resources to the case under investigation, for cost efficiency purposes. On the contrary, low maturity indicates that there is new evidence brought in, or new findings are attained and the use of resources is at its optimum. By doing this, we are enforcing a higher level protocol within a group decision support system.

For each stage, there is a set of data that needs to be created, captured and processed. A big challenge is to have a user interface that allows not only encoding and representing the various pieces of evidence and hypotheses effectively, but also to seamlessly integrate with the “DST engine”. That is, the component of the group decision support system will translate the encoded hypotheses and evidence to DST terms in order to allow further processing and application of the DST primitives and algebra such as the DST evidence combination rules.

### 3.4. Interface and standards requirements

Perhaps the biggest challenge from the standards perspective is the need to establish a suitable language in order to facilitate the coding and representation of the hypotheses used within the investigation. The language is needed in order to encourage standardisation on a semantic and ontology level. Forensic
investigations are highly formalised and dependent on strict and regulated procedures. Therefore a continuous need to incorporate, change and develop a suitable language and codes is part of the everyday professional practice.

The need of a suitable language which incorporates new practices of inquiry will challenge interface requirements, as there will be a need to integrate with data mining components and processes. For example, in [4] a crime text mining approach is proposed including a classification schema describing types of criminal activity in the light of terminology used by criminals. The interface should therefore be capable of handling the translation of the mined evidence to DST terminology and standards.

4. Conclusions and areas for future research

An approach by which information sharing and evidence consolidation can be performed by expert investigators for investigating cyber-crime is introduced. The approach adopted elements of the Strategic Systems Thinking framework which were formally developed under DST.

The purpose of this paper is to introduce the aforementioned concepts in the cyber-crime investigations domain and the suitability of the underlying tools is studied. Therefore a generic cyber-crime investigation framework is presented. An ongoing area of research is the application of the framework to specific cyber-crime classes, which could be based either on the technical nature of the malicious activity such as a DDoS, or on the type and purpose of the offence, such as identity theft. Furthermore, a variety of digital forensics tools include digital evidence bags [13], automated logging – network tools in particular – which could be interfaced with the proposed system. By doing this, a more specialised system equipped with the appropriate semantics would allow further exploration of the efficiency and effectiveness of the tool.

References