Ontology-Driven Approach for Evidence Admissibility in Network Forensics

Milica Matijević*, Stevan Gostojić*
* University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia
{matijevicmilica,gostojic}@uns.ac.rs

***Abstract*—Digital forensic investigators are commonly involved in legal proceedings aiming to support decision making. Therefore, the investigator's testimony must rely on admissible digital evidence without grounds for impugning. This paper presents an ontology-driven application for assisting digital investigators to conduct a network investigation that leads to court admissible results. The application also enables collaboration among digital forensic experts, thus maintaining awareness of the latest technology, techniques, and tools.**

Keywords—digital forensics, network forensics, forensic soundness, evidence, admissibility, ontology

I. INTRODUCTION

In forensic science, certainty is a word that is used with great care, cause even if a prosecution is not the goal of digital investigation, digital forensic cases may result in legal action. Therefore, it is advisable to handle digital evidence as if it were going to be presented to the court. To derive digital evidence from traces designated as potential digital evidence, an investigator should perform forensically sound and admissible procedures [1]. Accordingly, an investigator should be aware of the internationally recognized standards, legal regulations, and latest technology to ensure adherence to forensic principles, such as evidence authenticity and integrity, a chain of custody, tool reliability, and conclusion objectivity. [2] states internationally recognized standards such as ISO/IEC 27037, ISO/IEC 27041, ISO/IEC 27042, and ISO/IEC 27043, and gives a comprehensive overview of forensic phases: identification, collection, acquisition, conservation, analysis, evaluation, and presentation.

Besides the knowledge of the generic investigation process, every forensic discipline requires specific knowledge. Since information technology has become network centered, network forensics is often conducted and requires specialized knowledge of tools and techniques.

The question that we address is whether the specific network forensic process produces admissible digital evidence. Regarding the existence of a wide range of standards and regulations, how can one be sure that all of them were properly conducted during the forensic investigation. Besides the legal aspect of an investigation process, there is also a technical aspect that includes the tools, techniques, and adequate expert knowledge. Accordingly, providing admissible digital evidence is grounded on awareness of all legal and technical aspects of forensics investigation.

The increasing importance of digital evidence in legal procedures has imposed a formalization of the forensic processes and procedures that were conducted in an unstructured and ad hoc manner [3]. Therefore, in addressing the described problem, we developed an ontology-driven web application. The ontology model consists of three modules: digital forensics module, network forensics module, and computer network module. The digital forensics module represents the legal aspect and the network forensics module represents the technical aspect of an investigation process, while the computer network module describes all necessary concepts of computer networks.

The structure of this paper is as follows: Section 2 describes the background of ontology development and existing works on applying ontologies in the field of network forensics. Section 3 provides an overview of the ontology model. Section 4 explains the design, implementation and functions of the web application based on the ontology and Section 5 provides an example of using this web application. Finally, section 6 provides the summary of the work, describes advantages and disadvantages in comparing with similar ideas, and specifies directions of the future work.

II. BACKGROUND

To improve the efficiency and effectiveness of a digital investigation and to ensure standard procedures, different artificial intelligence techniques are being used. Many researchers apply ontology engineering to digital forensic data processing. Ontologies are used to support communication, computational inference, and reuse and organization of knowledge. Different types of ontologies exist in accordance with their granularity, formality, generality, and computational capability. In terms of generality, ontologies may be classified as top-level ontologies, mid-level ontologies, task ontologies, domain ontologies, and application ontologies. Domain ontologies specify concepts and inter-concept relations particular to a domain of interest. Task ontologies are ontologies developed for specific tasks, and application ontologies are ontologies used in specific applications where typically utilize both domain and task ontologies.

Web Ontology Language (OWL) is a standard for representing ontologies on the Semantic Web. OWL provides expressing semantics and allows automated reasoners to carry out logical inferences and derive knowledge. As it is not possible to fully achieve both of these objectives, OWL exists in three dialects known as OWL-Lite, OWL-DL, and OWL Full. OWL DL has a decidable inference procedure [4]. [5] states that OWL can express complex knowledge domains such as property cardinality constraints, domain, and range restrictions, and enumerated classes that define concepts. In OWL many constructors are also available: atomic and complex concept negation, concept intersection, universal restrictions, limited existential quantification, transitivity, role hierarchies, inverse roles, functional properties, datatypes, nominals, and cardinality restrictions. In the second version of OWL (OWL 2.0), these are

complemented by the union of a finite set of complex role inclusions and a role hierarchy, and unqualified cardinality restrictions are replaced by qualified cardinality restrictions.

[6] provides a comprehensive overview of ontologies utilized in the realm of digital forensics. [7] propose the software tool for ensuring digital forensics-aware web services design. They utilize an ontology that describes admissibility requirements, which can be modified by forensics experts. In such a manner, forensics experts assess the attribute weights of admissibility property based on their experience and knowledge about jurisdiction laws.

This paper focuses on the computer network forensics. Network forensics handles data related to the network traffic and data logged by end-points, switches, routers, proxies, monitoring systems, and intrusion detection systems. We applied ontology engineering to develop web application for supporting network investigation aiming to achieve admissible digital evidence and reliable investigation conclusions.

III. Ontology Model

For creating the ontology, we used the OWL DL language. Protégé 5.5, an open-source ontology editor based on OWL 2.0, has been used during the development of the ontology [8]. The ontology model [9] consists of three modules that describe digital forensics concepts, network forensics concepts, and essential concepts of computer networks. To develop the ontology modules mentioned, we considered relevant literature and reused some existing ontologies. The digital forensics module is oriented to the general principles of an investigation and prerequisites of digital evidence admissibility and performing forensically sound processes.

The admissibility term in [1] refers to the following: authorized obtaining of digital evidence, authentic digital evidence, assessing the reliability of digital evidence, presenting the best evidence to the court, and digital evidence not designated as hearsay. [1] also provides a comprehensive overview of computer network aspects that are important for investigation based on particular layers – the physical layer and data-link layer, the network layer and transport layer, and the application layer of the Internet.

According to [3] there are many definitions of the term "forensically sound" published by NIST, Association of Chief Police Officers (United Kingdom), Digital Evidence Standards Working Group, U.S. Department of Justice, and Australian Law Reform Commission. Taking into account all of them, authors concluded with four criteria for evidence evaluation – meaning, errors, transparency, and experience, claiming that the existence of a prescriptive set of rules that can be performed in every situation is not possible. The meaning criterion represents the need for digital evidence preservation in unaltered form. The error criterion suggests identification of all software and hardware errors encountered during the digital investigation process and a detailed description and explanation of their impact on digital evidence. The third criterion, transparency, describes an obligation of the investigator to document the investigation process thus it can be verified by other investigators under the same conditions. Finally, the fourth criterion, experience, is related to the appropriate and sufficient experience of the digital investigator who performs an investigation.

Scientific Working Group on Digital Evidence publishes guidelines for digital forensic investigation. Specifically, we considered [10] and [11] to create and populate the digital forensics module of the ontology. We took into account [12] to create the part of the digital forensics ontology module related to evidence admissibility. The concept of a tool is also proposed within the digital forensics module and populated by instances according to the detailed list of tools in [13].

According to the previously described literature, the digital forensics ontology module consists of classes presented in Figure 1.

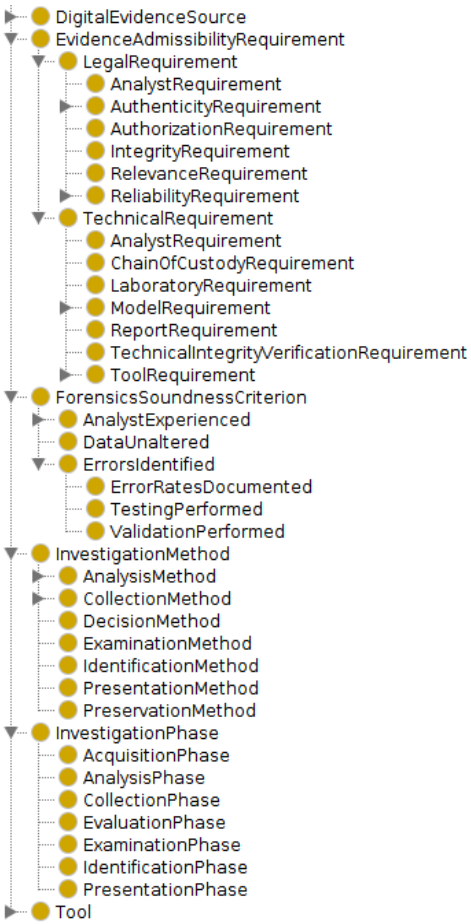

Figure 1. Digital forensics ontology module

The source of data that the network forensics ontology module is based on is [14]. We also considered the admissibility of Internet and cloud forensic evidence from [15] to enrich the network forensics ontology module.

To describe computer network concepts within the computer network ontology model, we reused the computer network ontology from [2], [16] and [17]. A very important thing an investigator should be aware of is obfuscation methods applied to some aspects of computer networks. [18] gives a comprehensive overview of such methods, hence we took them into account and enhanced the knowledge constituted within the network ontology module. Figure 2 presents the computer network ontology module which consists of essential concepts aiming to support network forensics knowledge.

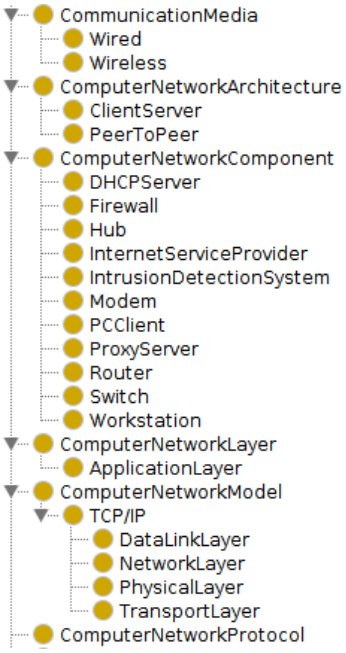

Figure 2. Computer network ontology module

To enrich the ontology knowledge, it is not sufficient to consider only the relevant literature. Therefore, we propose the way of how to include the knowledge of experienced forensic investigators. The ontology root classes are only unchangeable, but all subclasses, properties, and instances can be upgraded by competent digital investigators. This is described in the next section.

IV. IMPLEMENTATION

The web application [19] that helps digital investigators to perform forensically sound investigation process and provides experienced investigators with an ability to share their experience, is implemented using the Flask [20], a Python web framework. Considering the double role of the web application – to help inexperienced investigators and to provide engagement of experienced investigators in enhancing the formally described knowledge, the two working modes are provided – one for inexperienced investigators and the other for experienced and skillful investigators.

For accessing the content of the ontology in Python, Owlready2 Python module [21] has been used. Owlready2 is an Application Programming Interface (API) in Python [22], that provides manipulation of ontology classes, instances, and annotations in the same manner as Python objects are manipulated. It also allows for automatic determination and reclassification of objects using HermiT reasoner [23].

The functioning mechanism of the application is as follows: the user chooses the working mode according to his expertise. If the inexperienced investigator mode is chosen, the user gets insight into computer network concepts and instances. Then he/she is able to select those which contain relevant digital pieces of evidence related to his case. Selected computer network concepts or instances determine the digital evidence source which stores information about the targeted computer network mechanism. The digital evidence source then exposes the digital evidence admissibility requirement related to the forensic soundness criterion satisfied by achieving that admissibility requirement and the investigation method by which that requirement should be satisfied. Finally, the system shows information about tools that can be utilized in performing a particular method and information about the forensic investigation phase in which that method is performed. Therefore, an investigator being aware of admissibility requirements perform adequate testimony and make a report that can not be impugned.

If an experienced investigator approaches the web application, then he/she is able to upgrade the ontology model. Root classes of the ontology model are the base an investigator begins from, hence the root classes are unchangeable. As presented in Figure 3, the basic computer network and digital forensics concepts the ontology model consists of are CommunicationMedia, ComputerNetworkArchitecture, ComputerNetworkComponent, ComputerNetworkLayer, ComputerNetworkModel, ComputerNetworkProtocol, EvidenceAdmissibilityRequirement, ForensicsSoundnessCriterion, InvestigationMethod, InvestigationPhase and Tool.

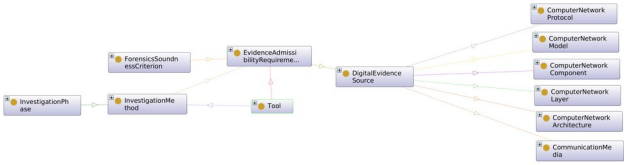

Figure 3. Root classes and their relations within Network forensics ontology model

A skillful investigator then chooses the specific subclass of the chosen root class and the system provides him/her with the ability to modify the instances or the hierarchy of that subclass. Figure 4 shows the flowchart of the described algorithm.

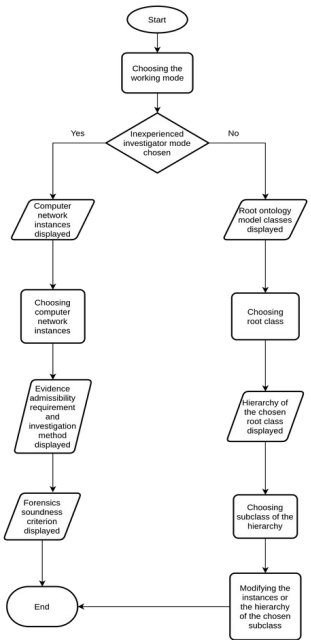

Figure 4. Flowchart of the web application functioning mechanism

V. CASE STUDY

The use case presented in this section comprises the web application working mode for both inexperienced and experienced digital forensic investigator.

We suppose that experienced investigator wants to enrich the ontology with knowledge about sources of digital forensic evidence related to the TCP/IP network model. That evidence includes different log files such as authentication logs, application logs, operating system logs, and network device logs. For each type of log file, he/she wants to note some specific log files.

For example, authentication logs are usually recorded by authentication servers such as RADIUS, TACACS, and Kerberos server.

Application logs comprise FTP transfer logs, web access logs, and e-mail server logs.

Some of the log files on the UNIX operating system are "Aculog", "Syslog", and "xferlog". "Aculog" contains a record of when the modems were used to establish a connection. Routers and firewalls are usually configured to add their log to the "Syslog" file on a remote logging server. "Xferlog" contains a record of all files that were transferred using the FTP protocol.

NetFlow is an example of a network device log file.

According to the Admissible Network Forensic Correlation Model [24], formats of collected log files must be unified and duplication of the single event within different log files must be reduced. After the collection of log files, the process of aggregation and correlation of events from log files should be performed.

Besides those forensic investigation methods, an expert investigator would note the investigation process phases related to specified methods, such as the examination and analysis.

In the remainder of this section, we describe the way an expert forensic investigator can upgrade the ontology with that information through the web application, and also we propose the way an inexperienced investigator can get insight into all information provided by an expert investigator.

After an expert investigator approaches the web application, all root classes of the ontology are being presented, such that an expert can choose one of them. Figure 5 depicts the UI of the web application used as an expert investigator.

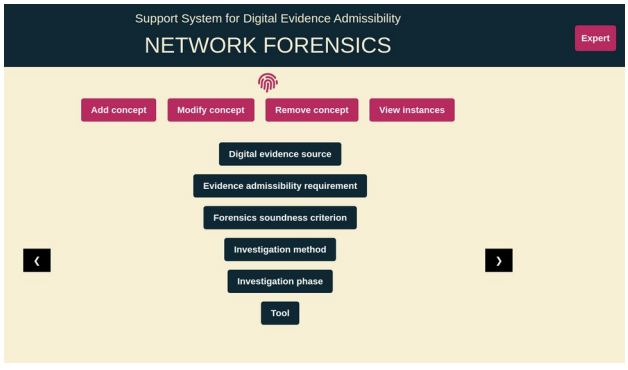

Figure 5. Web application's UI (expert investigator)

If he/she wants to improve the knowledge about digital evidence sources, he/she chooses the DigitalEvidenceSource class. Then he/she adds the LogFile subclass and LogFile specializations like ApplicationLog, AuthenticationLog, NetworkDeviceLog, and OperatingSystemLog. Those specializations have particular instances, hence an expert investigator creates them like follows: "Aculog", "Syslog", and "xferlog" as instances of OperatingSystemLog class, and "FTP_transfer_logs", "web_access_logs", and "e-mail_server_logs" as instances of ApplicationLog class.

As DigitalEvidenceClass is related to all classes representing network concepts, an expert has the ability to relate AuthenticationLog class with the AuthenticationServer class and its instances ("RADIUS" and "TACACS").

EvidenceAdmissibilityRequirement class is related to the DigitalEvidenceSource class, hence all the subclasses and instances of the EvidenceAdmissibilityRequirement class are presented to an expert. We divided the admissibility requirements into two groups – legal and technical requirements, thus an expert chooses one of the subclasses in order to enrich it.

We suppose that he/she wants to note some process models for log file manipulation and that he/she creates the LogFileModelRequirement subclass of the ModelRequirement class. Then he/she can instantiate this subclass with the instances representing the following requirements: duplication of the single event within different log files must be reduced; formats of collected log files must be unified; events from log files must be aggregated and correlated.

One of the forensics soundness criteria is that forensic investigators should be experienced. This is achieved in our specific case if the forensic analyst performs the appropriate process model. Accordingly, an expert investigator can upgrade that part of ontology creating the AppropriateProcessModelUsed subclass, which instance could be the representation of the Admissible Network Forensic Correlation Model. To note some investigation method that can be applied in this case, an expert adds the LogFileAnalysisMethod subclass within the analysis investigation method and instantiate it with instances "data_visualization_techniques", "machine_learning_techniques", "support_vector_machine_techniques, and "statistical_analysis".

The most important thing an expert investigator should do is to define the added subclasses to enable the HermiT reasoner to deduce the classes the instance belongs to.

Besides enriching the knowledge base, an experienced investigator must provide information about himself/herself aiming to prove competency for engagement in enhancing the knowledge in the realm of digital forensics. This is achieved by annotations for all classes and instances added by an expert investigator, such that an inexperienced investigator can trace that reference and assure himself/herself in the credibility of an expert investigator.

The web application provides an inexperienced

investigator with the ability to choose the relevant network forensics concepts or instances. Figure 6 depicts the UI of the web application used as an inexperienced investigator.

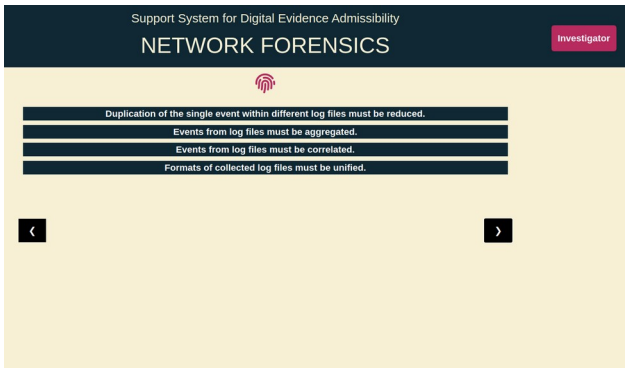

Figure 6. Web application's UI (inexperienced investigator)

We suppose that he/she wants to know what are the sources of TCP/IP-related digital evidence and what are the admissibility requirements and forensics soundness criteria that must be satisfied performing some investigation method and using some forensic tool. Therefore, an investigator chooses the TCP/IP class, and the system in the background creates the instance of that class and the anonymous instance of DigitalEvidenceSource class. The fact that an expert investigator previously fed the ontology with the LogFile class and subclasses and connected it with some TCP/IP instances the system is able to deduce that anonymous instance is of type LogFile. Then an investigator is provided with insight into all the subclasses and instances of the LogFile class. After that, the reasoning is continued, and types for all anonymous instances are found. Thus, an investigator is aware of admissibility requirements related to log files, the appropriate process model for manipulation with log files, and analysis methods that can be applied to log files.

VI. Conclusion

In this paper, we proposed the way of collaboration among experienced and inexperienced forensic investigators aiming to ensure the production of admissible digital evidence. We restricted the digital forensics field to network forensics and we presented network forensics concepts within the ontology. An experienced investigator has the ability to enhance the ontology model and populate it through our web application, and an inexperienced investigator has insight into the structure of the ontology and its instances, which is useful guidance during the investigation process.

An advantage of this approach in comparing with live meetings of digital forensics investigators or with forums is that our web application provides the user with a proof of the credibility of authors who were being part of the knowledge base creation and also our web application is based on the ontology, which provides automatic reasoning, such that the system itself can enrich the knowledge base.

Current implementation covers only network forensics concepts.

For future work we consider the integration of concepts belonging to other digital forensics disciplines. This means expert forensic investigators are able to integrate additional ontology modules through the web application. For example, an expert investigator could integrate the module that comprises mobile device concepts with the digital forensics module by mobile forensics concepts and instances.

Acknowledgement

We would like to acknowledge the support of the European Cooperation in Science and Technology (COST) action CA17124 "Digital Forensics: Evidence Analysis via Intelligent Systems and Practices (DigForASP).

References

[1] E. Casey, *Digital evidence and computer crime: Forensic science, computers, and the internet*, 3rd ed. Academic press, 2011.

[2] F. Amato, G. Cozzolino, V. Moscato, F. Moscato, "Analyse digital forensic evidences through a semantic-based methodology and NLP techniques", in *Future Generation Computer Systems*, 2019.

[3] R. McKemmish, "When is digital evidence forensically sound?", in *IFIP international conference on digital forensics*, Boston, 2008.

[4] V. Nguyen, "Ontologies and information systems: a literature survey", DSTO Defence Science and Technology Organisation, 2011.

[5] L. F. Sikos, "OWL ontologies in cybersecurity: conceptual modeling of cyber-knowledge", in *AI in Cybersecurity*, 2019.

[6] L. F. Sikos, "AI in digital forensics: Ontology engineering for cybercrime investigations", Wiley Interdisciplinary Reviews: Forensic Science, 2021.

[7] A. Akremi, H. Sallay, M. Rouached, M. F. Sriti, M. Abid, R. Bouaziz, "Towards a Built-In Digital Forensics-Aware Framework for Web Services", in *2015 11th Inter12national Conference on Computational Intelligence and Security (CIS)*, 2015.

[8] The Board of Trustees of the Leland Stanford Junior University. (2016-2020). *Protégé* [Online]. Available: https://protege.stanford.edu/

[9] M. Matijević. (2021, May 20). *Network Forensics Ontology* [Online]. Available: https://github.com/micamat/network-forensics-ontology

[10] *Model Standard Operation Procedures for Computer Forensics 3.0*, SWGDE policy

[11] *Recommended Guidelines for Validation Testing Version 2.0*, SWGDE policy

[12] A. Antwi-Boasiako, H. Venter, "A model for digital evidence admissibility assessment", in *IFIP International Conference on Digital Forensics*, 2017.

[13] R. C. Joshi, E. S. Pilli, *Fundamentals of Network Forensics*, Springer, 2016.

[14] N. Jaswal, *Hands-On Network Forensics: Investigate network attacks and find evidence using common network forensic tools*. Packt Publishing, 2019.

[15] T. V. Lillard, C. P. Garrison, C. A. Schiller, J. Steele, J. Murray, *Digital forensics for network, internet, and cloud computing. Syngress Publication Elsevier Inc*, 2010.

[16] L.F. Sikos, "OWL Ontologies in Cybersecurity: Conceptual Modeling of Cyber-Knowledge", in *AI in Cybersecurity*, 2018.

[17] D. Ellison, R. A. Ikuesan, H. S. Venter, "Ontology for Reactive Techniques in Digital Forensics", in *2019 IEEE Conference on Application, Information and Network Security*, 2019.

[18] R. Jones, *Internet Forensics: Using Digital Evidence to Solve Computer Crime*", O'Reilly Media, 2005.

[19] M. Matijević. (2021, May 20). Ontology-based Network Forensics Support System [Online]. Available: https://github.com/micamat/ontology-based-network-forensics-support-system.git

[20] Pallets, (2010). Flask web development, one drop at a time [Online]. Available: https://flask.palletsprojects.com/en/2.0.x/

[21] Python software foundation, (2021). Owlready2 [Online]. Available: https://pypi.org/project/Owlready2/

[22] J. B. Lamy, *Ontologies with Python: Programming OWL 2.0 Ontologies with Python and Owlready2*, Apress, 2021.

[23] Data & Knowledge Group, (2021). HermiT OWL Reasoner [Online]. Available: http://www.hermit-reasoner.com/

[24] A. Al-Mahrouqi, S. Abdalla, T. Kechadi, "Efficiency of network event logs as admissible digital evidence", in *2015 Science and Information Conference (SAI)*, 2015.