International Journal of Engineering & Technology, 7 (2.28) (2017) 187-190

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Criminal intelligence analysis based on ontological knowledge representation model

M. A. Jalil, F. Mohd*, C. P. Ling, N. M. M. Noor

School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Kuala Terengganu, Terengganu, Malaysia.
**Corresponding author E-mail: mpfatihah@umt.edu.my*

Abstract

Nowadays, community security is an issue which is given higher priority by all agencies, aiming to reduce crime incidence. As knowledge representation is the appropriate way to apply on complex crime analysis information, hence ontology-based case matching model is proposed to represent the relationships among the knowledge. Therefore, in this study, the ontology model is developed using semantic web modelling tool, TopBraid Composer Standard Edition in order to represent the crime information with the well-defined classes and relationships. The advantage of TopBraid is the ability in offering comprehensive supports for building, managing and testing the configuration of ontology and Resource Description Framework (RDF) graphs.

Keywords: *Crime analysis, Knowledge representation, Ontology, TopBraid.*

1. Introduction

The growing number of useful information and the complexity and dynamic characteristic of crime investigation area call for new ways of information and knowledge representation. Hence, applying an appropriate method of crime analysis information is needed. Crime analysis is used to extract data patterns of the criminal events and its investigation [1,2]. According to INTERPOL, criminal intelligence analysis is defined as the "study [of] criminals, crime suspects, incidents, issues and trends.. [to] identify relationships or connections between different crimes in different places" [3]. During a crime investigation process, a hidden relationship between different crimes to the previous crimes are extracted. It involved the matching process on how a crime is related to another crime and modus operandi used between that crimes.

In this study, knowledge presentation based on ontology is used to represent the relationships between data, such as modus operandi and crime scene in a structured and meaningful way. This method is used to represent the data that could assist police in identifying and analyzing crime patterns to reduce further occurrences of similar incidence and provide information to reduce the crime. According to Dzemydiene and Kazemikaitiene, the current system of a recording nature only keeps the criminal records, card files and databases [4]. It does not construct the information in a meaningful way and therefore it would be hard for the investigation officer (IO) to analyze the related information and make a proper decision during the investigation process. The key information of the criminal case report such as the modus operandi, case number, report summary, motorcycle type, crime scene and time is stored independently in separate folders without any relationships among the key information. In order to overcome the problems described by [4], there is a great need to develop an ontological-based case matching model where the IO could first acquire the related motorcycle theft information from the ontology model in a well-structured form and understand the relationships between the information easily.

Therefore, an ontology model is being proposed as a knowledge representation model to represent these relationships. In this study two problems have become the focus: (i) how to develop a knowledge representation model, (ii) to represent the crime investigation information and how could we make good use of the model's output to assist the investigating officers (IO) in aiming the targeted suspects based on the previous case records. So that, at the end of this study, the ontological-based case matching model applied to the crime investigation domain is presented. The rest of the paper is organized as follows. Section 2 discusses the materials and methods used, containing development of a framework and model used in this study. Section 3, the results produced are presented and discussed. Finally, the study is concluded in Section 4.

2. Materials and methods

This section explains the methodology of developing an ontological-based case matching model. A prototype named CrimeAnalysis based on an ontological-based case matching model is introduced. The ontological modelling of CrimeAnalysis presented the relationship between data, such as modus operandi and crime locations in a structured and meaningful way. The relationship between classes created in the ontology model will be applied in the case matching mechanism of the prototype in order to support the theft investigation process of the IO. The prototype would only represent the information of the motorcycle theft cases and the related output would be applied in the case matching process of the same domain.

Copyright © 2016 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2.1. CrimeAnalysis framework

The crime dataset used in the study applied the population and the sum of crime variables considered violent crimes in every state. Based on the domain being identified earlier, a framework of ontological-based case matching model for motorcycle the theft analysis is proposed as shown in Figure 1. It starts with the identification of key information to be fed into the model as the new case, development of an ontology model, implementation of the crime analysis prototype, and lastly performing case matching to produce final output, CaseID and case summary report.

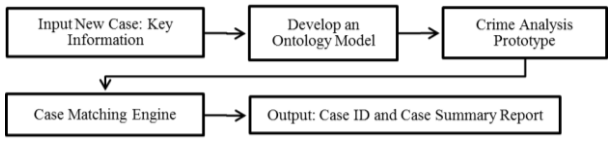


Fig. 1: Example of an image with acceptable resolution

The detail of CrimeAnalysis framework is elaborated in the next section. The proposed architecture of CrimeAnalysis model which is composed of pre-processing phase, case matching engine, database and ontology model is as illustrated in Figure 2. The ontology model needs to be constructed first to represent the classes and the relationships between the classes of the motorcycle theft information. Later, a prototype is developed to prompt the IO to input the data of the newly reported case of motorcycle theft. The case matching engine is triggered to match the newly input data with the previous data records that have been saved earlier in the ontology model. As a result, it would return the matched CaseId and display the case summary report to the IO as a reference during the investigation.

2.2. Pre-processing

A pre-processing process which includes tokenization, stop words removal and stemming process is added into the prototype in order to help the IO in retrieving the key information of cases in an automatic mode. For example, when a police receives a report from the public, the reported case would be recorded in a writing form referred to as the criminal case report. The report will then be stored in a database. Before storing the report, a pre-processing will take place to extract the keywords from the content of the report. The pre- processing steps are tokenization and stop words removal of the report text and stemming process.

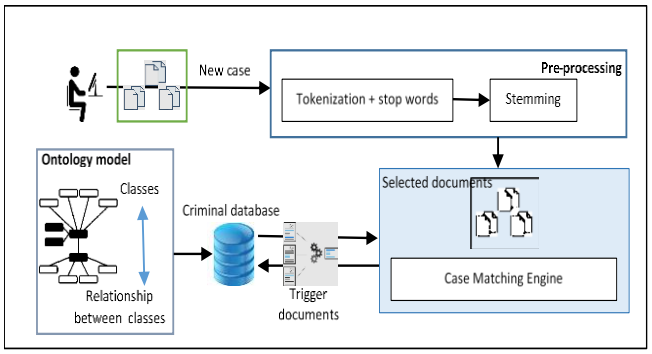


Fig.2: CrimeAnalysis model architecture

The tokenizer will take the raw text input, segments it into word tokens and identifies the boundaries of the sentences [5]. In short, tokenization occurs to extract the related keywords. While the stop words are words which can be excluded from the search and are being removed from the text [6]. Therefore, stop words removal is used to eliminate the unnecessary words. Lastly, a stemmer is used to extract the root words from the keywords by eliminating the affixes.

2.3. Developing a CrimeAnalysis model

Ontology model is developed to define the relationship and attributes, as well as classes of things that are characterized by certain relationship and attributes. In this study, the ontology model is built using the semantic modelling tool named TopBraid Composer [7, 8] (see Figure 3).

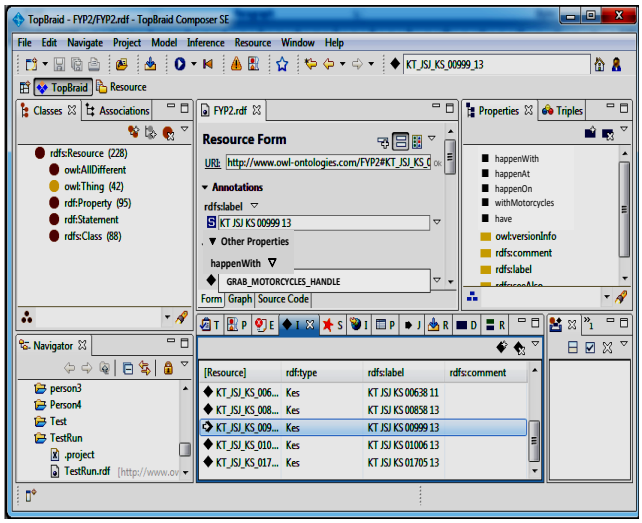


Fig. 3: Creating crime analysis model using TopBraid

Before developing the ontology model, the study of the key information in the criminal case report is done. Later, four key elements in the criminal case report that are vital in the investigation process are determined, namely modus operandi, motorcycle type, crime scene and time. These key elements would be defined as a different classes in the model during the development process. In this study, six main classes related to the motorcycle theft cases are created using the TopBraid tool, namely MotorcyclesBrand, Case, Time, ModusOperandi, CaseSummary and CrimePlace. After the classes have been defined, the object properties would be set. Object properties represent connections between Resource Description Framework (RDF) resource and another resources. Five objects properties created are “happenedWith”, “happenedAt”, “happenedOn”, “withMotorcycles” and “have”. The instances and relationships for each of the classes is also defined in this process. Figure 4 shows an example of the relationship between different instances of the classes that is represented as graph using TopBraid.

3. Results and discussion

This section presents the implementation of CrimeAnalysis prototype and the finding of the study.

3.1. Implementing a CrimeAnalysis model

To apply the information that has been well-defined by the ontology model, the prototype is developed. When the investigation officer (IO) input the data of a newly reported case in the prototype, it would automatically trigger the case matching mechanism. The newly input data is being matched with the data records saved in the ontology model. If matched, the CaseId of any related cases would be returned to the IO. Based on the returned CaseId, the IO could access similar past cases as a reference to the case investigation process (see Figure 2). The purpose of the prototype is to evaluate the accuracy of the framework. The prototype is expected to represent the relationships between the information in a structured form and the output produced would be applied in the case-matching mechanism of the web application in order to support the motorcycle theft investigation process.

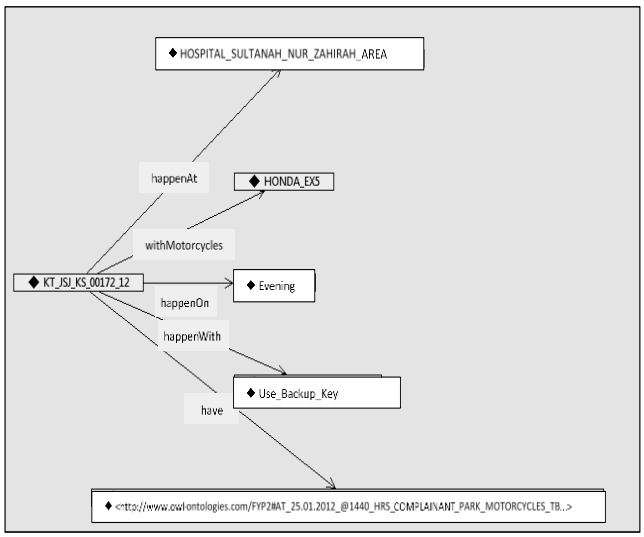


Fig. 4: Graph representation of crime analysis model using TopBraid

There are two main parts in the prototype (see Figure 5). The left hand side column consists of four combo boxes that prompt the user, IO, to choose the key information related to the newly reported motorcycle theft case which are the modus operandi, time, crime scene, and motorcycle type. While the right hand side column will prompt the IO to choose the case id in the combo box in order to display the desired case report summary. After the IO has done all the four selections: ModusOperandi, Time, CrimePlace, and MotorcyclesBrand, the system will display the summary report after being filtered in the filtering process (see Figure 6).

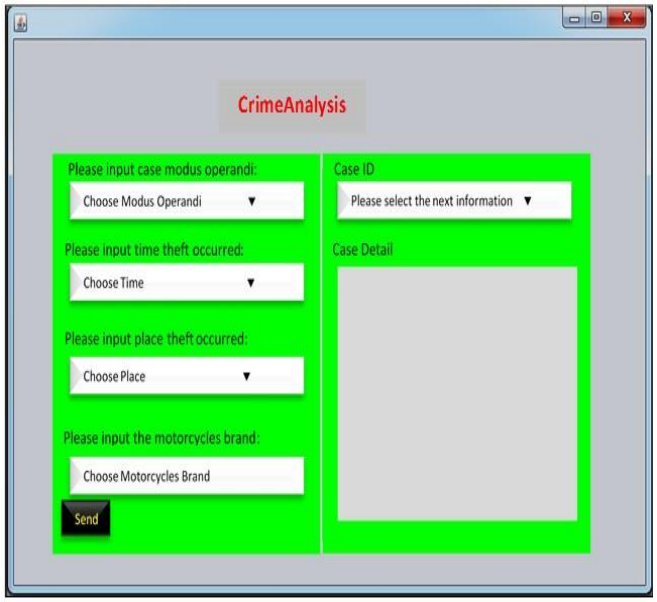


Fig. 5: CrimeAnalysis prototype

3.2. CrimeAnalysis evaluation

The efficiency of the proposed model is evaluated by running the tests repeatedly with a benchmark of motorcycle theft cases. The testing process are repeated with 30 different cases as the benchmark of this project in order to test the consistency of results gained. As a result, a 30 over 30 cases (100%) correct result was gained which could indicate that the case matching engine is reliable and it meets the purpose of the research. The crime dataset used in the study supplied by Director of Criminal Investigation Department, Royal Police Department Malaysia (RPM) and it is considered the population and the sum of crime variables considered violent crimes in every state of Terengganu.

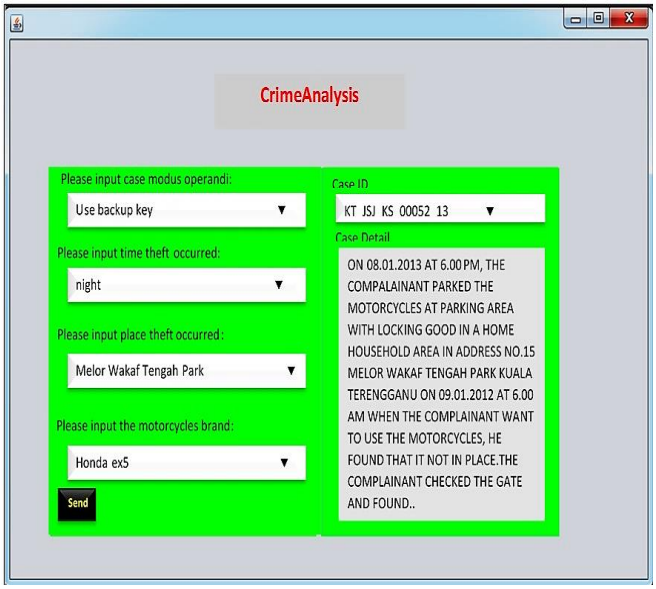


Fig. 6: Screenshot of summary report

4. Conclusion

The ontology model called CrimeAnalysis is implemented based on a framework of ontological-based case matching mechanism. The results show that CrimeAnalysis can be used to match the newly reported cases with the similar past case records and display the related case summary report to assist the IO in the investigation process. With the suggested pre-processing process being added into the model, it will help to save the IO's time and effort in retrieving the key information about the newly reported case in an automatic mode without requiring the IO to pick up the key information manually as with the current situation. In the age of Industrial Revolution 4.0 (IR 4.0), crime investigation should be more secured, fully or semi-autonomous, combined with artificial intelligence and fully integrated into organizational strategy. Thus, the ontology model proposed in this study, would help to save the security agencies' effort in aiming and targeting the possible suspect within the shortest time interval.

Acknowledgement

This study has been supported in part of the Fundamental Research Grant Scheme (FRGS) with vot number 59394, under the Malaysia Ministry of Higher Education (MOHE) and Universiti Malaysia Terengganu (UMT). The authors would also like to thank for continuous supports given by Director of Criminal Investigation Department, Royal Police Department Malaysia (RPM).

References

[1] Krishnamurthy R, & Kumar JS (2012), Survey of data mining techniques on crime data analysis. *International Journal of Data Mining Techniques and Applications* 1(2), 117-120.
[2] Wortley R, & Townsley M. (Eds.) (2016), Environmental criminology and crime analysis. *Taylor & Francis*, 1-25.
[3] INTERPOL. (2016), Criminal Intelligence Analysis. http://www.interpol.int/INTERPOL-expertise/ CriminalIntelligence-Analysis.
[4] Dzemydiene D, & Kazemikaitiene E (2005), Ontology-based Decision Support System for Crime Investigation Process. *Springer*, United States, US.

[5] Baldwin T, & Su'ad A, "Open source corpus analysis tools for Malay", Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC2006), (2006), pp:2212-5.
[6] Don ZM, (2010), Processing natural Malay texts: a data-driven approach. *Trames* 14(1), 90-103.
[7] Ramakrishnan S, & Vijayan A, (2014), A study on development of cognitive support features in recent ontology visualization tools. *Artificial Intelligence Review* 41(4), 595-623.
[8] Alatrish ES, (2013), Comparison some of ontology. *Journal of Management Information Systems* 8(2), 018-024.