Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2004 Proceedings

Americas Conference on Information Systems (AMCIS)

December 2004

Criminal Relation Exploration Tool for Law Enforcement Knowledge Management

Jennifer Xu
University of Arizona

Criminal Relation Exploration Tool for Law Enforcement Knowledge Management

Jennifer Xu
University of Arizona
jxu@eller.arizona.edu

ABSTRACT

Understanding relations between offenders in organized crimes is important for law enforcement and intelligence agencies to detect, prevent, and respond to crimes and terrorism events. However, several challenges (e.g., information overload, difficulty of relation search and extraction of relation patterns) face the task of criminal relation exploration. In this paper we present our ongoing research that develops an intelligent knowledge management system to help identify criminal relations from large amounts of crime incident records, search for the strongest relation paths between criminals, and extract and browse relation patterns and structural properties of a criminal network. We employ the concept space approach, shortest-path algorithms, and social network analysis techniques to achieve these goals. The preliminary results from evaluation studies show that our system can achieve a higher efficiency in criminal relation exploration than existing systems. Future work will be done to extend the system evaluation and analyze the dynamics of criminal networks.

Keywords

Criminal relation, relation search, relation browsing, law enforcement, knowledge management.

INTRODUCTION

It has been believed, especially after the tragic events of 9/11, that law enforcement and intelligence agencies need sophisticated knowledge management systems to solve crimes and prevent terrorist attacks to enhance national security. Especially, the investigation of organized crimes, such as terrorism, narcotics violation, and gang-related crimes, are dependent on effective and efficient information technologies that link new incidents to previous crimes, uncover relations between criminals, and extract patterns of criminal relations from data.

However, several challenges face the task of uncovering criminal relations. First, the information overload problem (Blair, 1985) resulted from large data volume may cost investigators much time and effort. Because criminal relations usually are not directly available in criminal-justice data, investigators may have to perform extensive database search and read a number of incident reports to seek clues of relations between offenders. Second, even if criminal relation data are available, it is difficult to manually search for relation paths between criminals who are not directly connected. One can get lost easily during a relation search when a criminal has connections with multiple persons who are also related to multiple other persons. Third, to fight organized crimes investigators must not only focus on individual crimes but also integrate data from multiple, serial crime incidents and try to extract patterns of criminal relations from data.

Although important and challenging, the development of knowledge management systems for uncovering criminal relations has not received much attention from academic researchers. We present in this paper our ongoing research which is aimed at developing techniques and systems to explore relations between offenders in organized crimes. Our system provides three major types of functionality: relation identification, relation search, and relation pattern browsing. Using the concept space approach (Chen and Lynch, 1992), our system can automatically extract criminal relations from crime incident data. These relations form a network representation in which a node represents a criminal and a link represents a relation found between two criminals. The relation search functionality is designed to identify the strongest relation paths among two or more criminals in the network using the shortest path algorithm (Dijkstra, 1959). The browsing functionality employs social network analysis (SNA) approaches and measures (Wasserman and Faust, 1994) to help identify from a criminal network the key persons, criminal groups, and patterns of interactions between groups. We also report our preliminary results from a system efficiency evaluation. The reason we focus on the efficiency issue is because it is essential to crime fighting (Hauck et al., 2002). A prompt resolution of crimes may mean fewer victims and less financial loss. The Sniper case occurred in Fall 2002 has shown that efficiency is a highly desired feature of law enforcement knowledge management systems. In our future studies, we will evaluate our system in terms of other metrics such as effectiveness, accuracy, and usability.

The rest of the paper is organized as follows. Section 2 briefly reviews related work of criminal relation exploration. Section 3 presents the proposed approaches and the system architecture. Evaluation results are presented and discussed in section 4. Section 5 concludes the paper and discusses our plan for future work.

RELATED WORK

In this section we review existing techniques and systems for criminal relation exploration.

Some techniques have been proposed to identify criminal relations from structured database records and unstructured textual data such as crime incident reports. For example, Goldberg and Senator (Goldberg and Senator, 1998) suggested to apply consolidation and link formation operations to uncover relations between individuals based on transactional data. This technique has been employed by the U.S. Department of the Treasury to detect money laundering transactions and activities (Goldberg and Wong, 1998). Lee (Lee, 1998) developed a technique to extract criminal relations from textual documents. Relying heavily on Natural Language Processing (NLP) techniques, this approach can extract entities and events from crime incident reports by applying large collections of predefined patterns.

For criminal relation search most existing systems allow only "single-level" search to identify relations between directly connected individuals. For example, the Watson system (Anderson et al., 1994) can identify possible relations between individuals by querying databases. Given a specific person's name, Watson automatically forms a database query to search for other related persons. The related persons found are linked to the given person and the result is presented in a link chart. The COPLINK Detect system (Hauck et al., 2002) also provides single-level relation search based on relations identified using the concept space approach (Chen and Lynch, 1992). However, neither Watson nor COPLINK Detect system allows users to search for relation paths connecting indirectly related criminals.

For criminal relation browsing several systems can automatically generate a network layout based on relational data. However, most of these systems cannot automatically identity network structural properties and relation patterns such as key persons, criminal groups, or interaction patterns between groups. The analysis burden is still on human investigators. For example, with the Anacapa charting system (Harper and Harris, 1975), an investigator first examines crime data to identify criminal relations and then constructs an association matrix to represent the discovered relations. Based on this association matrix, a link chart can be drawn for visualization purposes. An investigator may discover new investigative directions or confirm initial suspicions about specific suspects from the visual display of the network (Sparrow, 1991). Other systems such as Netmap (Goldberg and Wong, 1998) and Analyst's Notebook (Klerks, 2001) provide similar network visualization functions but cannot extract important structural properties and relation patterns in criminal networks.

In the next section, we prose thee techniques to help criminal relation exploration.

PROPOSED APPROACHES

Figure 1 presents the architecture of our system which takes crime incident records as input and provides three major types of functionality: criminal relation identification, relation path search, and relation pattern browsing. These three functions are closely related: the search and browsing functionality depends on the relations identified from criminal records; the search functionality provides a "local" view of relations among a few criminals while the browsing functionality provides a "global" network of relations among a large number of offenders. We propose the three techniques, namely, concept space approach, shortest path algorithms, and the social network analysis approaches, to fulfill the system functionality.

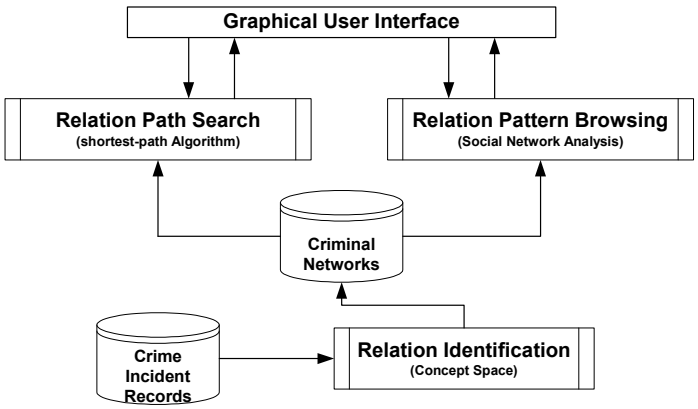


Figure 1. System Architecture

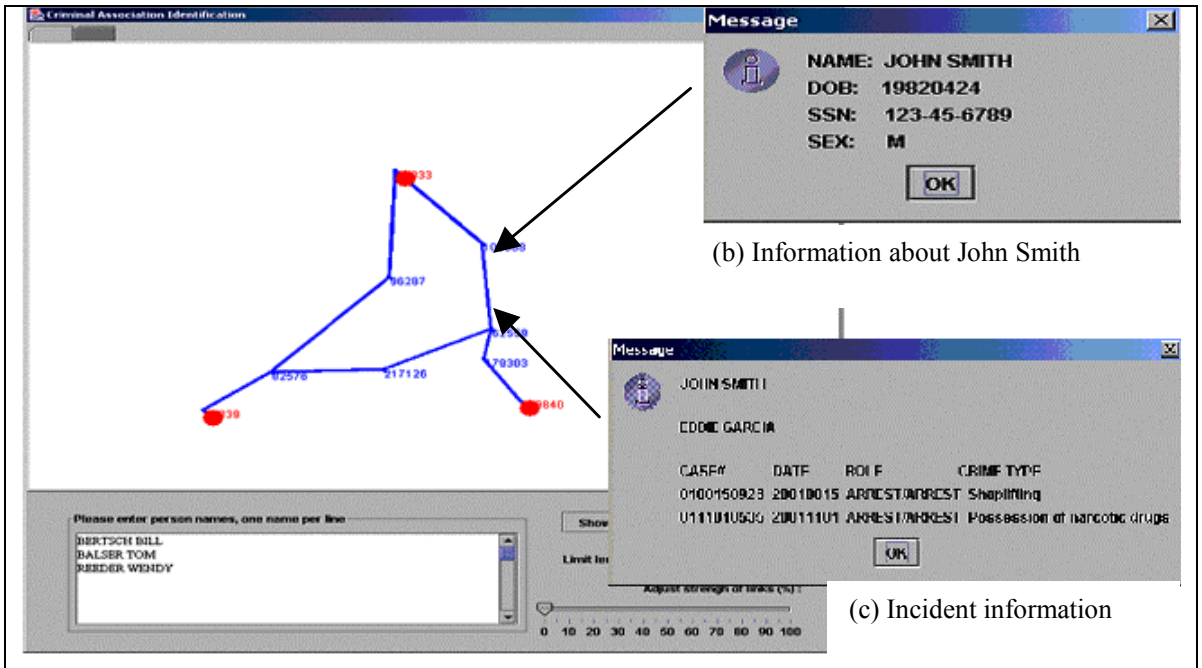
Relation Identification

Criminal records in law enforcement databases usually do not capture criminal relation information. We thus utilize crime incident records—structured database records specifying the date, location, persons involved, and other information about a specific crime—as the source for criminal relation information and use the concept space approach (Chen and Lynch, 1992) to identify the relations. This is based on the fact that criminals who commit crimes together usually are related, and the more often they appear together the more likely it would be that they are related. With this approach we treat each incident record as a document and each person's name as a phrase. Co-occurrence weights are calculated based on the frequency with which two individuals appear together in the same crime incident. As a result, the value of a co-occurrence weight not only indicates the presence of a relation between two criminals but also measures relational strength (Hauck et al., 2002). The relations identified are transformed into a network representation in which nodes represent individual criminals and co-occurrence weights of links represent relational strength.

Relation Path Search

Our system allows users to search for the strongest relation paths among criminals who do not appear to be related at the surface. We focus only on the strongest paths because each node in the network may have multiple links causing the retrieval of all possible relations among a set of criminals to be computationally prohibitive.

We choose the classic shortest-path algorithm (Dijkstra, 1959) to compute the strongest relation paths. The time complexity is $O(n\log n)$, where n is the total number of nodes in the graph (Cormen et al., 1991). Given a set of person names, the system computes the strongest relation paths between the input names and presents the search result on the graphical user interface. Figure 2 is an example illustrating a search for relations among three persons. All names are scrubbed or labeled with numbers to ensure data confidentiality. In Figure 2a, the large round nodes represent the three given persons. Other nodes represent the intermediate nodes connecting the three persons. A user can adjust the slider at the lower corner of the window and choose to display only links whose weights are greater than the value specified by the user. When the user clicks on a specific node the system shows the personal information details including his/her name, age, gender, SSN, etc. (Figure 2b). Clicking on a link between two nodes will bring up the corresponding records of the incidents in which the two persons are involved (Figure 2c).



(b) Information about John Smith

(c) Incident information

(a) The strongest relation paths found among three persons

Figure 2. Relation Path Search Example

Relation Pattern Browsing

Our system allows not only to visualize relations between individuals in a criminal network but also to browse the relation patterns and structural properties extracted from a network, including criminal groups, central members in a group, and interaction patterns between groups.

To detect criminal groups, we employ a clustering technique to partition a criminal network into clusters based on relational strengths. A cluster is corresponding to a criminal group, whose members are more closely related than people outside of the group. The resulting partition is a cluster hierarchy in which two clusters merge into a larger cluster at higher levels of the hierarchy (Jain and Dubes, 1988).

To identify key persons and interaction patterns between groups, we use centrality measures and blockmodeling approaches in social network analysis (SNA) research (Wasserman and Faust, 1994). SNA has been studied for several decades. Our aim is not to innovate the SNA field but to apply SNA to facilitate knowledge management in an understudied area—the law enforcement and intelligence domain.

Three centrality measures (degree, betweenness, and closeness) (Freeman, 1979) are used for key member identification. *Degree* measures how active a particular node is. It is defined as the number of direct links a node has. *Betweenness* measures the extent to which a particular node lies between other nodes in a network and is defined as the number of geodesics (shortest paths between two nodes) passing through the node. *Closeness* is the sum of the length of geodesics between a particular node and all the other nodes in a network. These three centrality measures have different interpretation in criminal network context. For example, a very active individual with a hight degree may be the leader of a group.

The between-group interaction patterns in a criminal network are extracted using the blockmodel analysis (Arabie et al., 1978). Given groups identified in a network, blockmodel analysis determines the presence or absence of a relation between a pair of groups. By this means, blockmodeling integrate individual interactions into relations between groups so that the overall structure of the network becomes more prominent (Wasserman and Faust, 1994, White et al., 1976).

Figure 3 illustrates the relation pattern browsing functionality of our system. The network consists of 164 gang members (Figure 3a). To browse criminal groups and their interaction patterns, a user can adjust the slider at the bottom of the window and the system will display all groups, which are represented by circles, at a specific level of the cluster hierarchy (Figure 3b). Lines between circles represent between-group relations. The higher the between-group link density, the thicker the between-circle lines. A user can also view key members in a group by clicking on a circle. A small window will pop up showing a table of the rankings of each group member in terms of their centrality and the inner structure of the group (Figure 3c).

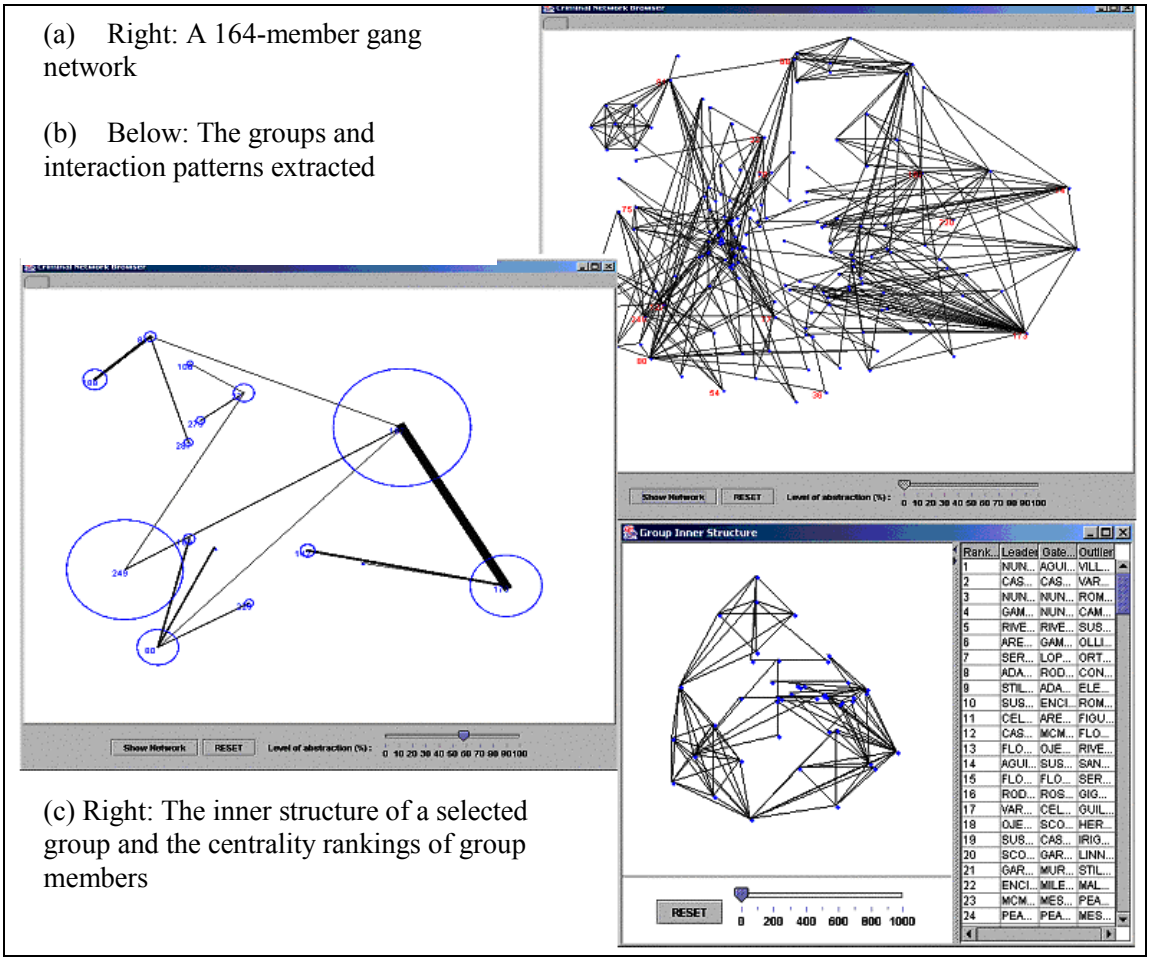


Figure 3. Relation Pattern Browsing Example

SYSTEM EVALUATION

In our previous research we evaluated the relation identification function and found that the concept space approach provided higher operating efficiency in crime investigation than the manual approach (Hauck et al., 2002). In this section we present the preliminary results from the evaluation of our system's relation path search functionality. We will evaluate the relation pattern browsing functionality in future studies.

This experiment was intended to ascertain whether the strongest relation path search was efficient compared to single-level relation search, which require manual exploration of all possible relation paths. We extracted from the Tucson Police Department (TPD) database 239,780 incident reports, in which 229,938 persons were involved. Ten experienced crime investigators at the TPD participated in the study. Each subject was asked to use COPLINK Detect, which provides only single-level relation search, and the shortest-path search function in our system to find the strongest relation paths among three given person names. The task completion time was recorded for each subject. Because the task required path search among multiple persons subjects must manually track the relations of each node using the COPLINK Detect, spending more time than using out system. The result shows that our system was faster than COPLINK Detect in term of average task completion time (COPLINK Detect: 24 minutes, Shortest-path: 0.7 minutes, $p < 0.0001$). This significant difference demonstrates that the shortest-path search function in our system can substantially improve efficiency of the search of the strongest relation paths between indirectly related criminals.

CONCLUSIONS AND FUTURE WORK

Sophisticated knowledge management systems can increase law enforcement agencies' abilities to detect, prevent, and respond to crimes and terrorism events, thereby enhancing national security and public safety. In this paper we present out ongoing research that is aimed at developing systems to help explore criminal relations.

Our system provides three types of functionality: identifying criminal relations from large amounts of crime incident records, searching for the strongest relation paths between criminals who do not appear to be related, and browsing of relation patterns and structural properties of a criminal network including key persons, criminal groups, and interaction patterns between groups. To achieve these functions, we employ the concept space approach, shortest-path algorithms, and SNA techniques. The preliminary results from the evaluation study show that our system can achieve a higher efficiency in criminal relation path search than existing systems.

We plan to extend our current research in three directions. First, we will conduct systematic evaluations on the efficiency of our relation pattern browsing functionality and use more performance metrics such as effectiveness, accuracy, and usability. Second, we will compare our systems with benchmark systems such as Analyst's Notebook to ascertain the value of our system in facilitating knowledge management in the law enforcement and intelligence domain. Third, we will analyze the dynamics of criminal networks in terms of changes in network structures, group cohesion, and between-group interaction patterns over time.

ACKNOWLEDGMENTS

This project has primarily been funded by the National Science Foundation (NSF), Digital Government Program, "COPLINK Center: Information and Knowledge Management for Law Enforcement," (#9983304, July, 2000-June, 2003).

REFERENCES

1. Anderson, T., Arbetter, L., Benawides, A. and Longmore-Etheridge, A. (1994) Security Works, *Security Management,* 38**,** 17-20.
2. Arabie, P., Boorman, S. A. and Levitt, P. R. (1978) Constructing Blockmodels: How and Why, *Journal of Mathematical Psychology,* 17**,** 21-63.
3. Blair, D. C., Maron, M. E. (1985) An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, *Communications of the ACM,* 28**,** 289-299.
4. Chen, H. and Lynch, K. J. (1992) Automatic Construction of Networks of Concepts Characterizing Document Databases, *IEEE Transactions on Systems, Man and Cybernetics,* 22**,** 885-902.
5. Cormen, T. H., Leiserson, C. E. and Rivest, R. L. (1991) *Introduction to Algorithms,* The M. I. T. Press, Cambridge, MA.
6. Dijkstra, E. (1959) A Note on Two Problems in Connection with Graphs, *Numerische Mathematik,* 1**,** 269-271.
7. Freeman, L. C. (1979) Centrality in Social Networks: Conceptual Clarification, *Social Networks,* 1**,** 215-240.
8. Goldberg, H. G. and Senator, T. E. (1998). Restructuring Databases for Knowledge Discovery by Consolidation and Link Formation. In *Proceedings of 1998 AAAI Fall Symposium on Artificial Intelligence and Link Analysis*: AAAI Press.
9. Goldberg, H. G. and Wong, R. W. H. (1998). Restructuring Transactional Data for Link Analysis in the Fincen Ai System. In *Proceedings of 1998 AAAI Fall Symposium on Artificial Intelligence and Link Analysis*: AAAI Press.
10. Harper, W. R. and Harris, D. H. (1975) The Application of Link Analysis to Police Intelligence, *Human Factors,* 17**,** 157-164.
11. Hauck, R. V., Atabakhsh, H., Ongvasith, P., Gupta, H. and Chen, H. (2002) Using Coplink to Analyze Criminal-Justice Data, *IEEE Computer,* 35**,** 30-37.
12. Jain, A. K. and Dubes, R. C. (1988) *Algorithms for Clustering Data,* Prentice-Hall, Upper Saddle River, NJ.
13. Klerks, P. (2001) The Network Paradigm Applied to Criminal Organizations: Theoretical Nitpicking or a Relevant Doctrine for Investigators? Recent Developments in the Netherlands, *Connections,* 24**,** 53-65.
14. Lee, R. (1998). Automatic Information Extraction from Documents: A Tool for Intelligence and Law Enforcement Analysts. In *Proceedings of 1998 AAAI Fall Symposium on Artificial Intelligence and Link Analysis*: AAAI Press.
15. Sparrow, M. K. (1991) The Application of Network Analysis to Criminal Intelligence: An Assessment of the Prospects, *Social Networks,* 13**,** 251-274.
16. Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications,* Cambridge University Press, Cambridge.
17. White, H. C., Boorman, S. A. and Breiger, R. L. (1976) Social Structure from Multiple Networks: I. Blockmodels of Roles and Positions, *American Journal of Sociology,* 81**,** 730-780.