Event-based Approach to Money Laundering Data Analysis and Visualization

Tat-Man Cheong
Faculty of Science and Technology
University of Macau
ma46514@umac.mo

Yain-Whar Si
Faculty of Science and Technology
University of Macau
fstasp@umac.mo

ABSTRACT
Crime specific event patterns are crucial in detecting potential relationships among suspects in criminal networks. However, current link analysis tools commonly used in detection do not utilize such patterns for detecting various types of crimes. These analysis tools usually provide generic functions for all types of crimes and heavily rely on the user's expertise on the domain knowledge of the crime for successful detection. As a result, they are less effective in detecting patterns in certain crimes. In addition, substantial effort is also required for analyzing vast amount of crime data and visualizing the structural views of the entire criminal network. In order to alleviate these problems, an event-based approach to money laundering data analysis and visualization is proposed in this paper. The effectiveness of the proposed method is demonstrated on a money laundering case from Taiwan.

Keywords
Crime detection; money laundering; event-based data analysis, visualization.

1. INTRODUCTION
In criminal networks, offenders play different roles for various illegal activities [19]. These offenders are usually connected via various relationships such as co-workers, friends, business partner and kinship [12]. Criminal networks usually involve in organized crimes such as armed robbery, drug trafficking and money laundering, etc.

Link analysis tools [11] are often used in investigation to uncover the information among criminals or criminal networks. Link analysis [28] may uncover the offenders who belong to the same supply chain in illicit drug supply networks. The main function of link analysis is to retrieve information from raw data which is related to entities such as bank accounts, telephone records, and criminal reports. The data is then processed and presented in a structured format. Link analysis further analyzes the information and identifies the relationship between the entities.

Vinci'10,Sep.28-29, 2010,Beijing, China.

In recent years, a number of researches on link analysis systems have been undertaken. These system includes NETMAP [21], Analyst's Notebook [1], COPLINK [7][13][14], etc. Most of these systems are originated from academic researches. Some of them have been extended into commercial products and are being used by various investigation departments. Majority of these systems use proprietary data formats for analysis.

Link analysis can be time-consuming and requires substantial amount of manual work. This includes ad-hoc tasks such as searching heterogeneous databases, analyzing reports on crime, and reviewing clues from criminal networks. Although some commercial software are labeled as link analysis tools, these software only provide functions for visualizing manually constructed criminal networks. In these systems, associations between entities are often required to be input by the user manually.

In addition, association analysis still encounters many challenging issues. For instance, it is extremely time consuming in analyzing vast amount of available crime data and visualizing structural views of the entire criminal network. As a result, these tools are less effective in detecting patterns in certain crimes. In order to alleviate these problems, an event-based approach to criminal data analysis and visualization is proposed in this paper. The proposed approach specifically targets in detecting money laundering cases.

The overall system utilizes a structured event-based database to store crime related records. The system also includes modules for detection of clusters, modules for calculation of degree of suspicion and degree of association, and a module for criminal networks visualization. A prototype system based on the proposed techniques is being designed and implemented in Microsoft Visual Studio [23].

The paper is structured as follows. In Section 2, we review related work on link analysis, visualization tools, and event-based techniques. In Section 3, different corruption and money laundering methods and trends are introduced. In Section 4, we briefly describe the overall system design. In Section 5, a brief introduction on heuristics for detecting money laundering is discussed. In section 6, algorithms for proposed crime detection system is detailed. In Section 7, a case study on a money laundering case from Taiwan is given. In Section 8, we conclude the paper with future work.

2. BACKGROUD AND RELATED WORK
In criminal investigation, association analysis tools are often used to identify/search the links connecting various entities from the dataset of large criminal networks.

2.1 Link Analysis and Visualization

NetMap Analytics [21] is used by intelligence agencies around the world for analyzing millions of items of data to detect fraudulent activities. NetMap uses the techniques that process across a whole spectrum of criminal cases to identify the murderers, drug rings, insurance fraud rings, terrorists, and the supporters. NetMap can link thousands of pieces of data and reveal invisible connections among entities.

At present, NetMap assists the government investigators from Australia, UK, USA and Europe to cull useful information from the large volumes of data for analyzing the criminal behavior. Two successful cases of such applications in Australia are the "Backpacker Murders" and the "Mystery TNT Options Trader" [21].

Analyst's Notebook [1] is an analysis and visualization tool for analyzing criminal data and fraudulent activities. The investigators can use Analyst's Notebook to identify connections, patterns and trends between criminal networks. The Analyst's Notebook can be combined with other compatible products via the i2 operation platform to acquire data from disparate data sources. It also allows users to search, match and analyze data from the various products. For instance, Analyst's Notebook users can generate a chart by dragging and dropping entities into the forms.

COPLINK [7] integrates multiple data sources from local, regional, and national police departments for detecting the association among the crime entities. The COPLINK prototype has been designed with knowledge management technology to consolidate, share, and identify relationships from criminal records. NetMap Analytics, Analyst's Notebook, and COPLINK [13] are all designed with the functions for visualizing the association among entities. A common goal of these systems is to analyze and discover hidden relationships and interconnections from seemingly unrelated data.

2.2 Event-based Database

In this section, we review existing event-based systems based on a number of criteria. An event is defined as a significant occurrence to present activities at a single point in space-time [15]. These events usually involve components such as occurred time and location related to the activities. Event-based database stores records involving various parameters and dimensions of events. The event time and location are the primary attributes [16] when representing a criminal case. A spatio-temporal database [4] is a system which primarily handles both space and time information [26]. These databases are also used in processing information such as crime records [27], patients' history, and records of traffic accidents etc. Table 1 shows a summary on event-based techniques from [6].

Table 1. Event dimensions [6]

Dimension	Explanation/metrics	Analysis tools
Time-related	The event dimensions are associated with day, date, minute, second, week, month, year components etc.	ReCAP [2] COPLINK [3]
Space-related	The event dimensions are associated with region and position information.	ReCAP [2] CrimeStat [17] COPLINK [3] Snap [10]
Content-related	The event dimensions are associated with reports, document details, and case description event details.	CrimeStat [17] COPLINK [3] Snap [10]
Person-related	The event dimensions are associated with personal information such as occupation, name, date of birth, address, nationality, contact information, etc.	COPLINK[3]

3. MONEY LAUNDERING METHODS AND TRENDS

Money laundering is the process of creating an appearance that large amounts of money obtained from illegitimate business originated from a legitimate source. Money laundering cases usually involve directing or collecting large amount of money coming to/from unknown origins. In such cases, a number of bank accounts may also be setup for temporary use by the offenders. The main purpose is to make the money "clean" or to hide the ambiguous source of funds. The process may involve concealing large amounts of money acquired from serious crimes, such as trafficking firearms, drugs, high-interest loan, smuggling, prostitution, fraud, corruption, tax evasion, extortion, kidnapping and robbery, etc. In money laundering cases, offenders usually tend to avoid traditional payment systems such as credit cards and checks to conceal financial transactions. Although cash payments are widely used in money laundering cases, it is hard to deal with huge amount of cash since it is bulky and difficult to be transferred.

In summary, money laundering can be also considered as a process of concealing illegal money transactions. Money laundering includes three money laundering stages namely placement, layering and integration [9].

3.1 Placement Stage

Placement is the first stage in the money laundering cycle [9] [18] [20]. In this stage, the suspects may attempt to avoid detection of illegal benefits by transforming physical currency into other assets. For instance, huge amount of (physical) money from the illegal activities is laundered through the financial system via business activities such as cash smuggling, investing in real estates, buying valuable items, currency trading, buying of securities and futures, depositing into bank accounts or mixing with the legal money, remittance, and retailing.

3.2 Layering Stage

In the layering stage, the suspect will carry out a series of activities in an attempt to obscure the trail of the illicit source of the funds [9] [18] [20]. The launderer may establish complex financial transaction layers between the funds in order to conceal the audit trail and the ownership of the funds. The layer generally consists of bank-to-bank, bank-to-remittance agent, money changer-to-bank, etc. It may also involve transferring of funds to offshore bank accounts related to shell companies in tax haven countries, resale of goods/assets, and frequent deposits/transfers/transections.

3.3 Integration Stage

The final stage in the laundering cycle is to gather (aggregate) the illicit funds from various legitimate commercial activities and financial systems. These activities involve selling of property assets, valuable goods, and financial instruments [9] [18] [20].

The dispersed illicit funds appear as legal assets after the integration process.

An example money laundering scenario between three users (A, B, and C) is shown in Figure 1. In this process, user A attempts to confuse, hide and interrupt the tracking of funds by passing through different channels. The illicit funds could be transferred via a remittance agent in order to avoid the scrutiny by law enforcement or bank regulatory authorities. The funds may also be divided into numerous amounts which are lower than the threshold of government reporting requirement. The money launderer may also perform a number of deposits or withdrawals in small amount to avoid detection by the authorities.

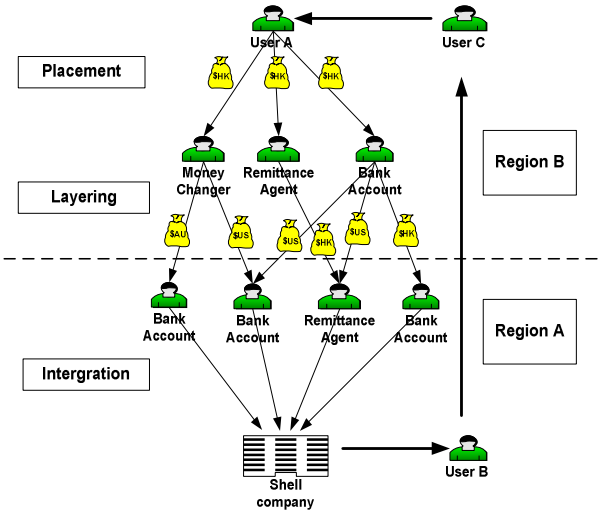

Figure 1. Money laundering procedure

4. SYSTEM DESIGN

The design of the proposed crime analysis system is depicted in Figure 2. The system consists of four components which include Event-Based Database for Criminal Records, Data Preprocessing for Criminal Detection, Clusters and Association Detection, and Visualization. The proposed system can be partly mapped to the notional model of sensemaking loop for intelligence analysis [25].

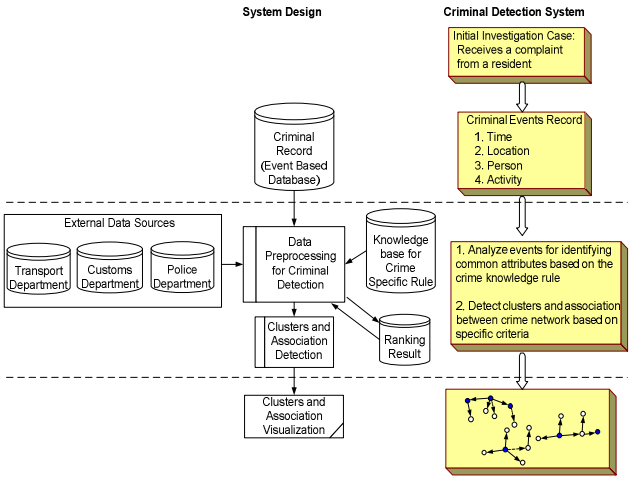

Figure 2. System design

In our system, crime related information is stored in the criminal record database. The database contains basic information such as time, locations, persons, and crime activities. In addition, a knowledge base is used to store domain specific heuristics for identifying links. The cluster and association detection module in Figure 2 is responsible for finding criminal networks and the degree of association between two or more criminals. The uncovered clusters and associations are then displayed in visualization module.

4.1 Event-based Data Analysis

We use an event-based database model [7] to store the records about criminal relationship network analysis. In this model, crime entities are represented as objects. Each object includes information such as time, locations, persons and activities. In addition, associations are defined as links between two or more objects. An example of event-based method to represent activities within a specified period is shown in Figure 3.

Time is a critical factor in detecting criminal cases. Investigators usually specify the start and end date of the detection period to filter some important clues. For instance, in some corruption cases, the suspect used to have certain decision making privileges in an organization during a specific period. For this type of cases, the investigator will be interested in collecting the related information about the suspect such as activities of his/her colleagues and family members during that period in which the suspect has special privileges. Such investigation can be found in the corruption case of former Taiwanese's President, Mr. Chen. He was involved in money laundering during his time as the president of Taiwan.

In the prototype system, an event-based model is designed to record all the activities of the criminal entities together with the respective time attributes. For instance, as shown in Figure 3, the system stores an event regarding establishing of a company at the time of 2007/12/10 by a person called Dick. The "Officer" attribute records the origin of the information. For example, it records whether the information is provided by a witness, or an informant, or by a law enforcement officer. The attribute "RPT 012/09" provides a reference to identify the information elsewhere in a database, or a filename. Attribute "L1" (level 1) is used to reflect the reliability of the information.

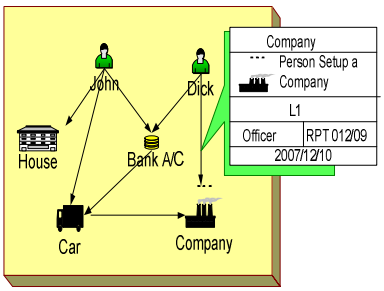

Figure 3. Example of an event

4.2 Data Preprocessing for Criminal Detection

Money laundering is one of the most reported crimes in recent years and usually connected with smuggling and corruption cases. Although money laundering occurs in various parts of the world, it has some common characteristics. These characteristics become

recognizable signatures in detecting money laundering cases. In the proposed model, data preprocessing for criminal detection module employs a knowledge base for storing crime specific heuristics. These heuristics are described in the Section 5.

If necessary, the proposed system could be extended to interface with other data sources. For instance, the system may need to acquire additional data from disparate data sources such as police, customs and transport departments. In our system, a ranking database is also designed to store temporary search result obtained from the application of crime specific rules. The ranking is used to assist in data preprocessing to filter the unrelated information.

5. HEURISTICS FOR DETECTING MONEY LAUNDERING

For identifying associations, we have used some of the investigation methods proposed in [5]. Attributes considered for detecting criminal associations in money laundering cases are identified as follows:

Corruption Network: In general, a suspect may use a network of trusted people such as family members, colleagues or friends for laundering money. The suspect could be in charge of the treasury/accounting department or related to other staff in those types of department or holds important positions in government or public organizations or may have business connections or financial activities with other suspects.

Crime period: Offenses often take place when the suspect is appointed as a key officer.

Bank Transactions: Money laundering may involve frequent deposits or withdrawals of large cash transactions. Bank accounts may only be opened for a short period for deposit of money. In some countries, any remittance above a predefined level needs to be reported to the authorities. In some money laundering cases, instead of using large transactions, lower value transactions are often used to avoid detection. For instance, the Black Market Peso Exchange [33] uses structured transactions method to hamper any investigations. "U-turn" transactions are also used in money laundering cases. For instance, money from a person or a company will be transferred to another person or company, and then re-transferred the money back to the original owner or the company.

Suspicious entities: Suspects may launder unlawful money via shell companies, secretarial companies, company formation agents, simulation of international loans, capitalization of legitimate companies with illegal funds [8] in "Tax haven", and "Off-shore financial centers" which act as the authorized signatory of the bank account. The launderer usually establishes these anonymous companies to disguise the true ownership of the unlawful funds.

Suspicious regions: Events related to countries and regions which are not in compliance with anti-money laundering Financial Action Task Force (FATF) [8] Recommendations should be scrutinized.

Suspect's career: Low income occupations such as driver, hawker, waiter and student should not involve high value transactions.

Suspect's age: Young or elderly people should not involve high value transactions.

Suspect's salary: Based on the records of suspect's salary, possible amount of any remittance or deposit related to the related bank accounts can be estimated.

Transaction behavior: The average balance, the total number of transactions, and types of transactions of an account over a period of time can give an indication of the financial activities of the suspect. These activities could be either classified as normal or suspicious. Increase in any activities should be carefully examined by the investigators.

6. DETECTION ALGORITHMS

In this section, we describe the algorithms for calculating the degree of suspicion of a person and degree of association between two entities.

6.1 Algorithms for Calculating the Degree of Suspicion

All entities (including the suspect) are assigned with a degree to denote the level of suspicion (or possibility of committing the offense). In addition to the degree of suspicion for each member, we also define the degree of association for every link among entities within a cluster. This degree is used to represent the closeness between two entities. In this algorithm, if two persons share the same property (such as car, house, stock shares), or if they are found to be part of certain financial transactions, the degree of association of the link between them is increased. The algorithms for calculating the degree of suspicion of a person and degree of association between two entities are summarized in Table 2.

Table 2. Algorithms

	Summary
Algo. 1	Identifies the family members of the suspect during a given period.
Algo. 2	Identifies the colleagues of the suspect during a given period.
Algo. 3	Calculates the degree of suspicion of the suspect and its colleagues within the cluster based on the type/nature of occupations.
Algo. 4	Calculates the degree of suspicion of the suspect and its colleagues/family members within the cluster. In this algorithm, bank transactions between the suspect and its colleagues/family members are used for calculation.
Algo. 5	Calculates the degree of suspicion of all members within the cluster based on the members' bank activities.
Algo. 6	Calculates the degree of suspicion of all members within the cluster based on the existence of shell companies.
Algo. 7	Calculates the degree of association between two entities.

6.2 Algorithm for Calculating the Degree of Association

To calculate the degree of suspicion and degree of association, we define a scoring scheme as shown in Table 3. Whenever the corresponding condition is true, the degree of suspicion or association will be increased by the point defined in the table. For instance, in row 2 of degree of suspicion calculation, if the suspect is the head of the finance related department, the degree of suspicion will be increased by 1 point by the Algorithm 3. Note that the points defined in Table 3 can be different depending on the preference of the investigator or the nature of the crime.

Table 3. Degree of suspicion and association

Condition	Point(s)	Apply to Algorithm

Degree of suspicion		
Family member of the suspect	1	1
Head of finance related department	1	3
Staff member of department under investigation (suspicious department)	2	3
Head of department under investigation (suspicious department)	3	3
Transaction between suspects' bank accounts	2	4
Large bank transactions	1	5
Degree of association		
Colleague association	1	7
Family association	2	7
Bank activity association	2	7

6.3 Iterative Crime Detection and Link Identification

For each family and colleague member in cluster, all seven algorithms described above can be used iteratively to discover further suspects and new clusters. For instance, after applying algorithm 1 and 2, we can identify the family and colleague clusters associated with the suspect. In our prototype, yellow circles are used to indicate colleagues and green circles are used to indicate family members.

Next, algorithms 3, 4, 5 and 6 are used to calculate the degree of suspicion of each member in the cluster. If a suspect's degree of the suspicion is greater than a predefined threshold value, the node representing the suspect will be represented by a blue circle. The iterative detection can be preformed for the cluster members who have high degree of suspicion.

The final step in detection is the identification of strongest association among suspects. The algorithm 7 can be repeatedly used to find the association from a single source node to all the other nodes in a connected graph. If the degree of association is greater than a predefined threshold value, the link will be highlighted to represent a possible criminal connection.

Main algorithm

Begin

Initialize:

suspectValue {Set a threshold for degree of suspicion}

suspectLinkValue {Set a threshold for degree of association}

WA {Set of tuple for storing degree of suspicion}

LA {Set of tuple for storing degree of association}

Set P = root suspect

Call detectSuspect(P)

for *each suspect* A_i **do**

if WA_i > *suspectValue* **then**

Call detectSuspect(A_i)

endif

if LA_i > *suspectLinkValue* **then**

Highlight the link between P and A_i

endif

endfor

function *detectSuspect(P)*

F = Call algorithm1 {Return all the related family member}

C = Call algorithm2 {Return all the related colleagues}

Set A = F+C {The set of all family and colleague suspect list}

for *each suspect in the list i = 1 to Total (A)* **do**

while *detect degree of suspicion is end = false* **do**

WA_i = *Call (algorithm3, algorithm4, algorithm5, algorithm6) for* A_i

{Return the total degree of suspicion of each member A_i *in the cluster}*

LA_i = *Call algorithm7 for* A_i

{Return the degree of association between P and A_i *}*

endwhile

endfor

draw the result as a cluster

return (A, WA, LA)

endfunction

end

6.4 Visualization

Finally, the entities in the criminal network and all detected associations are visualized in graphical forms. In these graphs, each entity is assigned with a unique identity and different node colors are used for classification. The conceptualization of visualization process is depicted in Figure 4. At the current stage of the prototype implementation, we use Pajek [22] for visualizing identified clusters and associations. As for the future work, we are planning to use Microsoft Visual Studio [23] for the implementation of visualization module.

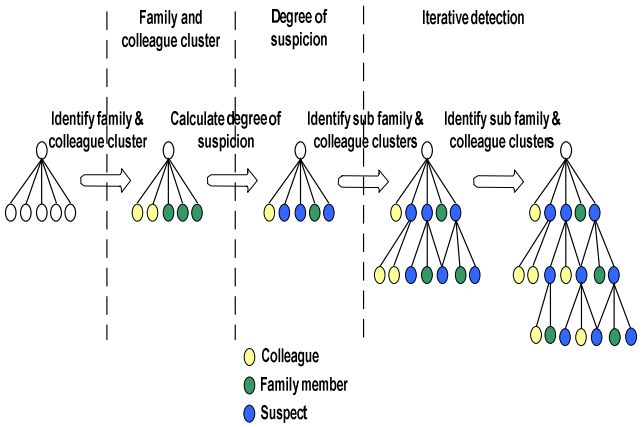

Figure 4. Visualization

7. CASE STUDY

In this section, we demonstrate the effectiveness of the proposed algorithms through the corruption case of Taiwan former president, Mr. Chen (Secret Presidential Funds) [30]. The Secret Presidential Funds scandal initially took place in 2006 in Taiwan, involving former President Mr. Chen and his family members. Mr. Chen, who was the president of Taiwan during 20 May 2000 to 19 May 2008, was indicted for corruption by The Public Prosecutor's Office of the Taiwan High Court on the alleged misuse of state funds involving fraudulent claims of 14.8m Taiwan Dollars (US$448,484). However, Mr. Chen was protected by constitution against criminal charges during his presidency and he was only prosecuted after he had stepped down. The first lady, WU, and three former presidential aides including former deputy secretary-general, MA, former Presidential Office director, LIN and HUI have been found guilty [29] by the prosecutors for inappropriately reimbursing personal expenses of the former first family using false receipts out of the president's state affairs fund between July 2002 and March 2006.

7.1 Case Analysis

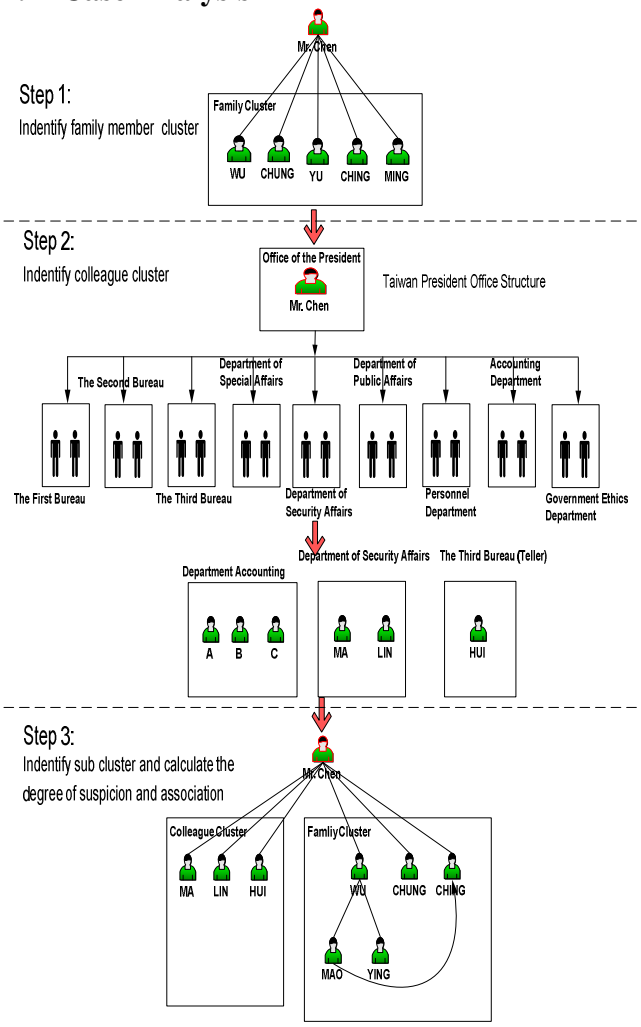

Figure 5. Detection workflow for the case of Mr. Chen

The investigation was initiated when Taiwanese prosecutor's office received some complaints about WU stealing purchase receipts for claiming from the Secret Presidential Fund. The fund is originally designated for handling diplomatic work overseas. According to the Taiwan's laws and regulations, a receipt should be provided to claim the related expense [29]. In the indictment from the Taiwan Taipei District Prosecutors Office, it was stated that WU, MA, LIN, and HUI are accountable by the corruption act and have been prosecuted based on the investigation result and evidence collected.

In this case study, we apply the algorithms from section 6 to the suspect and all related parties. Figure 5 shows the investigation procedure to identify the clusters.

- In step 1, algorithm 1 is used to identify the family cluster of Mr. Chen.
- In step 2, algorithm 2 is used to identify the colleague cluster of Mr. Chen.
- In step 3, algorithms 3, 4, 5, and 6 are used to calculate the degree of suspicion of each member in the family and colleague clusters. For each member in the sub cluster group, the algorithm 7 is used to calculate the degree of association. Both degrees are calculated iteratively for all members in the identified clusters.

Finally, by comparing both degrees with respective threshold values, suspicious targets and links from the clusters are identified and they are visualized as connected graphs.

7.2 Detecting Family Clusters

Family members, and closely related colleagues or friends are usually used as trusted persons to handle illegal funds during money laundering. Algorithm 1 is used to identify family members during the event period (2000/5–2008/5) from the family event records. For instance, in Table 4, a "marriage" event is used to link WU and Mr. Chen in 1975. In the algorithm 1, if the suspect is detected as the family members, the degree of suspicion will be increased by 1 point.

The algorithm discovers WU, CHUNG, YU, CHING, and MING as the family members of the suspect. The corresponding degree of suspicion for each member is shown in Table 5. The family cluster is shown as a connected graph in Figure 6.

Table 4. Family event records [29][31]

Event	Event Details	Related Person	Related Person	Event Date	Location
Marriage	WU got married with Mr. Chen	WU	Mr. Chen	20/2/1975	Taiwan
Birth	CHUNG was given birth by WU	CHUNG	WU	22/1/1979	Taiwan
Marriage	CHING got married with CHUNG	CHING	CHUNG	18/6/2005	Taiwan
Brother relationship	MAO is WU's brother	MAO	WU	11/7/1952	Taiwan

Table 5. Degree of suspicion for suspects from the family cluster

Suspect	Relationship	Degree of suspicion
WU	Family	1
CHUNG	Family	1
YU	Family	1
CHING	Family	1
MING	Family	1

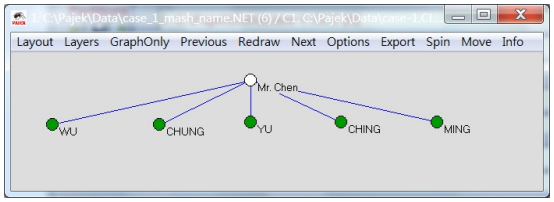

Figure 6. Family cluster

7.3 Detecting Colleague Clusters

The former President Mr. Chen exercised his powers in accordance with the Constitution [24] and established the Office of the President (see Figure 7) which includes: First Bureau, Second Bureau, Third Bureau, Department of Special Affairs, Department of Security Affairs, Department of Public Affairs, Personnel Department, Accounting Department, and Government Ethics Department.

Next, we apply Algorithm 2 to identify all of the occupations of suspicious person during the event period (2000/5–2008/5). The algorithm adds all related colleagues of each organization to the suspect list.

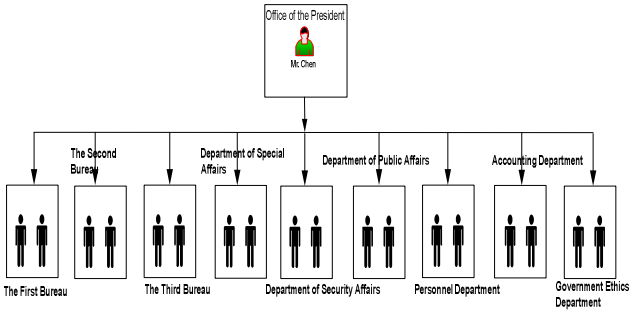

Figure 7. Taiwan president office structure [24]

Table 6. Occupation event records [29][31]

Event Details	Related Person	Start Date	End Date	Organization	Department	Position
Mr. Chen won the 2000 Presidential election	Mr. Chen	20/5/2000	19/5/2008	Office of the President	Office of the President	President
Mr. Chen took up the appointment as Mayor of Taipei	Mr. Chen	1994	1998	Office of Mayoralty	Office of Taipei Mayor	Mayor of Taipei
Assigned as the Secretary of department of Security Affairs	MA	20/5/2000	31/1/2005	Office of the President	Security Affairs	Secretary
Assigned as the Deputy Secretary-General to the President	MA	1/2/2005	4/6/2006	Office of the President	Office of the President	Deputy Secretary General
Assigned as the Secretary of Security Affairs Department	LIN	1/3/2005	19/5/2008	Office of the President	Security Affairs	Secretary
Assigned as the Accountant of Accounting Department	HUI	20/5/2000	28/2/2005	Office of the President	Accounting	Accountant

7.4 Iterative Detection

In the previous section, algorithm 1 and algorithm 2 are used to identify the family and colleague clusters of the suspect Mr. Chen. In order to narrow down the investigation scope, algorithm 3, 4, 5 and 6 are used to calculate the degree of suspicion for each member in the family and colleague cluster. Algorithm 3 is used to identify all colleagues who have ever been employed at the financial department and suspicious department (i.e. Security Affairs department). Specifically, the algorithm identifies the colleagues of the suspect who have ever worked at the Security Affairs department during the investigation period. In Figure 8, algorithm 3 has identified department heads A, B, C. The algorithm also identifies MA and LIN who are in charge of Security Affairs department. Please note that we have replaced the names of some of the suspects with anonymous letters to protect their identities. The resulted family and colleague clusters are shown in Figure 9.

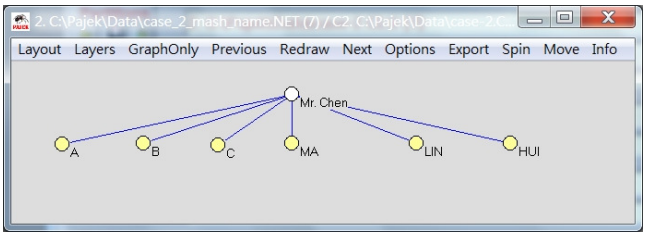

Figure 8. Colleague cluster (in yellow color)

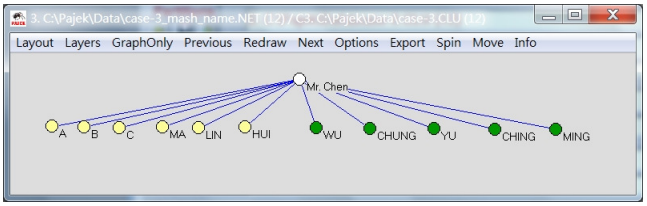

Figure 9. Family cluster (green) and colleague cluster (yellow)

Since MA and LIN, are found to be in charge of these selected departments, the degree of suspicion of these suspects is also increased. Based on Algorithm 3, the degree of suspicion for each person in Table 6 can be calculated. The algorithm calculates the degree of suspicion based on the predefined score points from Table 3. The resulted degree of suspicion for all suspects is shown in the Table 7.

Table 7. Degree of suspicion for colleague cluster

Suspect	Relationship	Degree of suspicion
A	Colleague	1
B	Colleague	1
C	Colleague	1
MA	Colleague	3
LIN	Colleague	3
HUI	Colleague	2

Next, Algorithm 4 and 5 are used to detect bank transactions between the suspect and its colleagues/family members. The

finance related events extracted from [31][32] are depicted in Table 8. The degree of suspicion for WU is increased to 4 since two transactions are detected in Table 8. In a similar way, the degree of suspicion for CHUNG and CHING can be calculated.

The updated degree of suspicion is shown in the Table 9. The main algorithm proceeds with the iterative detection for the suspects who have high degree of suspicion such as WU as shown in the Table 9.

In Figure 10, members who have high degree of suspicion (greater than 2) are displayed in blue circles.

Table 8. Finance related events [31][32]

Event Type	Event Details	Related Persons	Event Date
Finance	Diana Chen sent 10 million check to WU	Diana Chen; WU	1-6/4/2004
Finance	WU sold over 100 million worth of jewelry for money laundering in Hong Kong and remitted to CHING's bank account in Switzerland through Hong Kong and Singapore	CHING; WU	-
Finance	CHUNG used "Galahad Managements S.A." name to open accounts in "RBS Coutts Bank AG" of Switzerland for laundering 700 million	CHUNG	-
Finance	Mr. Chen was bribed in Longtan case for accepting 100 million Taiwan Dollars and transferred the unlawful money to overseas bank account of CHUNG	Mr. Chen; CHUNG	-
Finance	HUI assisted Mr. Chen to transfer around 5.55 million US Dollars to MAO's bank account in Singapore	Mr. Chen; HUI	Start from 10/1/2002

Table 9. Updated degree of suspicion

Suspect	Relationship	Degree of suspicion
WU	Family	4
CHUNG	Family	4
YU	Family	1
CHING	Family	3
MING	Family	1
A	Colleague	1
B	Colleague	1
C	Colleague	1
MA	Colleague	3
LIN	Colleague	3
HUI	Colleague	3

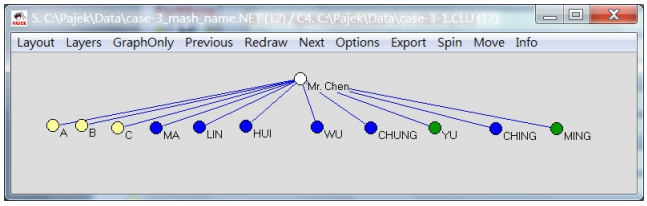

Figure 10. Persons with high degree of suspicion (in blue color)

The iterative detection continues to identify the family cluster of WU which includes YING and MAO (WU's brother). Through algorithm 4, the degree of suspicion is again increased for YING and MAO.

According to the records from Table 10, there were some transactions between MAO and CHING. According to records, money from MAO's bank account in Singapore was transferred to CHING's bank account in Switzerland. Finally, MAO was indicted [32] by The Public Prosecutor's Office of the Taiwan High Court and charged for assisting in money laundering as shown in Figure 11.

Table 10. Bank account transaction events [29][31]

Account A holder	Bank account A	Direction	Account B holder	Bank account B	Amount	Type	Date
MAO	124709	A -> B	CHING	467683	1000000	Transaction	14/6/2005
MAO	124709	A -> B	CHING	467722	3000000	Transaction	19/6/2005

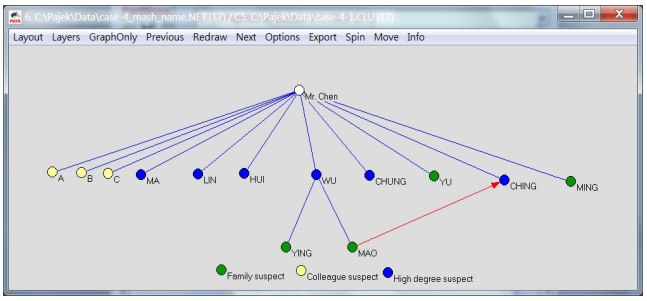

Figure 11. Identified clusters and links after iterative detection

7.5 Degree of Association

After the clusters have been identified, the algorithm 7 is used to determine the degree of association between two or more suspects. The algorithm applies scores defined in Table 3 for calculating the degree of association. The result of the calculation is shown in Table 11.

For instance, in row 1, last column of Table 11, degree of association between Mr. Chen and WU (denoted by link 2-1) is equal to 4 since (a) they are part of the suspect's family, and (b) bank transactions between them have also been detected by the investigators. Each condition results in score of 2 points and therefore the total degree of association is equal to 4. Degree of association for other suspects can be calculated in a similar way. The corresponding graph for Table 11 is depicted in Figure 12. Links with high degree of association are represented in red color.

Table 11. Degree of association for all suspects

Person ID	Suspect	Relationship	High suspicion	Link	Degree of association	Link	Degree of association
1	Mr. Chen	Family	Yes	1-1	0	2-1	4
2	WU	Family	Yes	1-2	4	2-2	0
3	CHUNG	Family	Yes	1-3	4	2-3	4
4	YU	Family	-	1-4	2	2-4	2
5	CHING	Family	Yes	1-5	4	2-5	4
6	MING	Family	-	1-6	2	2-6	2
7	A	Colleague	-	1-7	1	2-7	0
8	B	Colleague	-	1-8	1	2-8	0
9	C	Colleague	-	1-9	1	2-9	0
10	MA	Colleague	Yes	1-10	1	2-10	0
11	LIN	Colleague	Yes	1-11	1	2-11	0
12	HUI	Colleague	Yes	1-12	1	2-12	0
13	YING	Family	-	1-13	2	2-13	4
14	MAO	Family	-	1-14	2	2-14	4

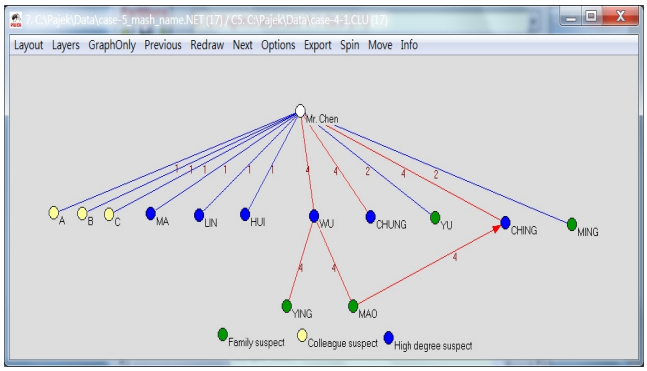

Figure 12. Visualization for degree of association (in red color links)

After the investigation, The Public Prosecutor's Office of the Taiwan High Court has indicted MA, LIN, HUI, WU, CHUNG and CHING. These suspects can be found in Figure 12.

8. CONCLUSION

In this paper, we propose an event-based approach to criminal data analysis and visualization. The overall system utilizes a structured event-based database to store criminal records. In this system, several algorithms are developed for identification of clusters (family and colleagues) and calculation of suspicious degree and association degree. The result of the calculation is then visualized as connected graphs. The effectiveness of the proposed method is demonstrated on a money laundering case from Taiwan.

The main contributions of our research work are two-fold; from the theoretical standpoint, we contribute to the analysis, design, and development of an event-based approach for money laundering data analysis and visualization. The proposed algorithms can (a) iteratively identify clusters, (b) calculate degree of suspicion and association based on the event records, and (c) visualize the analysis outcomes. Such automated processes presented in this paper can be extremely helpful for investigators since tedious manual network construction process is no longer required. From the practical standpoint, our research opens the door to the further development of mechanisms for detecting different types of crime. For instance, the heuristics described in section 5 can be extended for crimes such as drug trafficking and homicide.

As for the future work, we are extending the visualization module. We are also planning to extend the knowledge base for detecting other crimes.

Acknowledgments. This research is funded by the Research Committee, University of Macau.

9. REFERENCES

[1] Analyst's Notebook. http://www.i2.co.uk/. URL last accessed on 15 August 2010.

[2] Brown, D.E. 1998. The regional crime analysis program (RECAP): a framework for mining data to catch criminals. In: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics* (1998). IEEE, 2848-2853.

[3] Buetow, T., Chaboya, L., O' Toole, C., Cushna, T., Daspit, D., Petersen, T., Atabakhsh, H., and Chen, H. A spatio-temporal visualizer for law enforcement. 2003. In: *Proceedings of the First NSF/NIJ Symposium on Intelligence and Security Informatics (ISI 2003).* Springer, Heidelberg, Tucson, AZ, USA, 2003, 181-194.

[4] Chen, J., and Jiang, J. 1998. Event-based Spatio-temporal Database Design. *International Journal of Geographical. Information Systems*, 32, 4 (1998), 105-109.

[5] Chi, C. *Analysis of methods of investigation of corruption cases Respective sections: Prosecution theory.* http://www.ahjcy.gov.cn/jcy-news/newsinfo.jsp?id=9160. URL last accessed on 15 August 2010.

[6] Chung, W., Chen, H., Chaboya, L.G., Toole, C.O., and Atabakhsh, H. 2005. Evaluating event visualization: a usability study of COPLINK spatio-temporal visualizer. *International Journal of Human-Computer Studies*, 62, 1 (January 2005), 127-157.

[7] COPLINK Analytics. http://www.i2group.com/us. URL last accessed on 15 August 2010.

[8] Financial Action Task Force (FATF). 2007. Complex Money Laundering Technique, a Regional View (23 February 2007). http://www.oecd.org/dataoecd/30/22/38418180.pdf. URL last accessed on 8 January 2010.

[9] Australian Transaction Reports and Analysis Centre, Australian Government. 2008. Introduction to Money Laundering (December 2008). http://www.austrac.gov.au/elearning/pdf/intro_amlctf_money_laundering.pdf. URL last accessed 15 August 2010.

[10] Fredrikson, A., North, C., Plaisant, C., and Shneiderman, B. 1999. Temporal, geographical and categorical aggregations viewed through coordinated displays: a case study with highway incident data. In: *Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth ACM international conference on Information and knowledge management* (1999), 26-34.

[11] Goldberg, H.G., and Senator, T.E. 1995. Restructuring databases for knowledge discovery by consolidation and link formation. In: *Proceedings of the First International Conference on Knowledge Discovery and Data Mining*, (1995). AAAI Press, 134-141.

[12] Harper, W.R., and Harris, D.H. 1975. The application of link analysis to police intelligence. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 17, 2 (April 1975), 157-164.

[13] Hauck, R.V., Atabakhsh, H., Ongvasith, P., Gupta, H., and Chen, H. 2002. Using Coplink to analyze criminal-justice data. *IEEE Computer*, 2002, 35, 3 (March 2002), 30-37.

[14] Hauck, R.V., Chau, M., and Chen, H. 2002. Coplink: arming law enforcement with new knowledge management technologies. In: *Advances in Digital Government: Technology, Human Factors, and Policy* (May 2002). Springer, 163-180.

[15] Jain, R. 2003. Experiential computing. *Commun. ACM* 46, 7 (Jul. 2003), 48-55.

[16] Jain, R. Events in heterogeneous environments. 2003. In: *Proceedings of the International Conference on Data*

Engineering (Bangalore, India, 2003). IEEE Computer Society Press, 8-21.

[17] Levine, N. 2010. CrimeStat: a spatial statistics program for the analysis of crime incident locations (v 3.03), *Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC,* (July 2010).

[18] Jost, P.M., and Sandhu, H.S. 2000. The hawala alternative remittance system and its role in money laundering. *International Criminal Police Organization - INTERPOL*, http://www.interpol.int/public/financialcrime/moneylaundering/hawala/default.asp. URL last accessed on 18 August 2010.

[19] McAndrew, D. 1999. The structural analysis of criminal networks. In: D. Canter, L. Alison (Eds.), *The Social Psychology of Crime: Groups, Teams, and Networks, Offender Profiling Series*, III, Aldershot, Dartmouth, (1999), 53-94.

[20] Bureau of International Narcotics and Law Enforcement Affair. 2004. Money Laundering Methods, Trends and Typologies. *International Narcotics Control Strategy Report*, (March 2004). http://www.state.gov/p/inl/rls/nrcrpt/2003/vol2/html/29910.htm URL last accessed on 15 August 2010.

[21] NetMap Analytics, http://www.netmap.com. URL last accessed on 15 August 2010.

[22] Networks/Pajek, Program for Large Network Analysis, http://vlado.fmf.uni-lj.si/pub/networks/pajek/. URL last accessed on 15 August 2010.

[23] Microsoft Visual Studio Developer Tools http://msdn.microsoft.com/en-us/vstudio/default.aspx. URL last accessed on 15 August 2010.

[24] Presidential Office Organization, Republic of China (Taiwan), http://www.president.gov.tw/1_structure/department_1.html. URL last accessed on 15 August 2010.

[25] Pirolli, P. and Card, S. 2005. The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis. In: *Proceedings of the International Conference on Intelligence Analysis* (May 2005), 2-4,

[26] Sadri, R., Zaniolo, C., Zarkesh, A.M., and Adibi, J. 2001. A Sequential Pattern Query Language for Supporting Instant Data Mining for e-Services. In: *Proceedings of the 27th International Conference on Very Large Data Bases* (2001), 653 - 656.

[27] Schwenke, S. 2002. Cross-Sector Analysis of Corruption: Summary Report, Sectoral Perspectives on Corruption (November 2002), http://pdf.usaid.gov/pdf_docs/PNACX009.pdf URL last accessed 15 August 2010.

[28] Sparrow, M.K. 1991. The application of network analysis to criminal intelligence: an assessment of the prospects. *Social Networks*, 13, 3 (September 1991), 251-274.

[29] Taiwan Taipei District Court Prosecutors Office Prosecutor's indictment. http://www.tpc.moj.gov.tw/public/Attachment/61131652147.pdf. URL last accessed on 15 August 2010.

[30] BBC News. Taiwan's Chen in corruption case. http://news.bbc.co.uk/2/hi/asia-pacific/6112668.stm. URL last accessed on 15 August 2010.

[31] Taiwan's Supreme Prosecutors Office, special investigation team of prosecutors on the additional indictment of Mr. Chen case http://www.tps.moj.gov.tw/public/Attachment/952516162998.TXT. URL last accessed on 15 August 2010.

[32] Taiwan's Supreme Prosecutors Office, special investigation team of prosecutors on the indictment of Mr. Chen case http://www.tps.moj.gov.tw/public/Data/8121219507655.PDF. URL last accessed on 15 August 2010.

[33] United States Department of the Treasury Financial Crimes Enforcement Network. 1997. FinCEN Advisory 9, (November 1997). http://www.fincen.gov/news_room/rp/advisory/pdf/advisu9.pdf. URL last accessed on 15 August 2010.