Determining possible criminal behaviour of mobile phone users by means of analysing the location tracking data

Boris Galitsky

School of Computer Science and Information Systems
Birkbeck College, University of London
Malet Street, London WC1E 7HX, UK
galitsky@dcs.bbk.ac.uk

Alexander Miller

Profilium, Inc. 152 Rue Notre Dame East, Suite 300
Montreal, Quebec Canada, H2Y 3P6
amiller@profilium.com

Abstract

The possibility to detect a criminal behavior of mobile phone user, analyzing the location tracking data is considered. A hybrid reasoning system is applied to extract behavior patterns from a group of agents who coordinate their locations via mobile phones

The problem outline

Nowadays, more than 80% percent of adult population use mobile phones. This form of telecommunication service is heavily dependent on the accurate determination of the handset locations to promptly switch from one service station to another. Telecommunication servers accumulate huge amount of data that includes the recording of locations of handsets at certain time intervals. Also, the phone numbers of both callers and call addressees are recorded. Such data is a tremendous resource of information that can be used for a wide variety of purposes, including, in particular, targeted advertisement, based on accumulated customer profiles (Profilium 2005).

Currently, to the best of our knowledge, this kind of data is not used to detect suspicious forms of behavior. A special pattern recognition and reasoning technology is required to automatically process the location data (see e.g. Wherify 2005); its manual browsing is unaffordable. The lack of adequate technology leads to disregarding a rich body of data of potential value in filtering out criminal behavior. Crimes might be prevented and networks of criminals groups with peculiar inter-connections (Saferstein 1990) identified if it were possible to discover sets of unusual patterns of coordinated movement for groups of mobile phones. Also, it is important to develop a special location tracking services for anonymous phones, where only the trajectory of movement may identify an individual.

Extracting the behavior patterns

To extract the behavioral patterns from the data that is derived from the formal description of human activity (in particular, the data on locations and calls), it is necessary to have a computational model of participating (interacting) agents that simulates their behavior. We must reproduce the reasoning of wireless subscribers to hypothesize on their movements and calls to judge on the possibility of a criminal behavior for a selected group. Having obtained the behavioral patterns as results of supervised machine learning, it is possible to apply them to location and call data in real time.

The organizers of 2005 AAAI Spring Symposium on AI Technologies for Homeland Security have raised the question whether AI technologies can augment the ability of human analysts to objectively analyze large quantities of complex data while simultaneously reducing the impact of their personal biases. Interactive analysis of the location-based services (LBS) data may serve as an example of such AI system, combining such fields as temporal and special databases, data mining and primarily, reasoning about agents' attitudes (Galitsky 2003).

Assigning mental states to interacting wireless subscribers

The raw data for our analysis includes the series of absolute locations (detected with certain accuracy and time intervals) for wireless subscribers (agents) and the selected locations where these agents are making a call or a receiving a call. A scenario includes the set of mental states of the participating agents (who want to inform or to give an order, or want to be informed or being given an order).

The mental state for each agent can be fully or partially reconstructed. Scenario does not contain the data on conversation content because of the privacy issues and unreliability of speech recognition of the wireless signals. Scenarios where the calls do not affect the future locations (or change the movement directions) are irrelevant to our consideration.

We intend, for example, to detect the following scenarios:

1. Following a target vehicle by one or multiple vehicles when the target vehicle intends to escape.
2. Escaping a vehicle by a single or multiple vehicles that wants to reach its target.
3. Observation of a landmark by one agent and transmission of information about this landmark to another agent (surveillance).
4. Involving a group of agents with hierarchical subordination structure.
5. Approaching a meeting point by two or more vehicles in a manner that is not straight-forward.

These scenarios should be distinguished from the normal ones where a set of agents exchange calls to meet each other. Also, for two communicating agents we intend to understand who is leading and who is following, to analyze the other behavior peculiarities and distinguish normal from criminal patterns. Note that such scenarios as frequent calling to one agent by another, frequent exchange of calls by the parties that intend to meet, coordination of one vehicle by another vehicle using the mobile phone, taken separately, are the normal scenarios. Table 1 contains the sample semantic interpretation of LBS data for two agents.

Scenario	Interpretation	Information transmission direction
D D		
→	Obtains advice	→
T	Chooses advised directions	
←	Asks for directions	→
T	Chooses advised directions	
←	Neutral call or asking for	↔ or →
D	order / direction	
→	Advice on direction	→
→	Call to someone else	
T	Turn in accordance	

Table 1: The scenarios and their possible assignments with mental states. In the left column, ← → are the cell phone calls from one agent to another, D- driving without changing the direction, T-turn. The right column presents the detected direction of information transmission.

The system architecture

Revealing the behaviour of a group of wireless subscribers includes the pre-processing step, two consecutive rule-based subsystems and the prediction subsystem (Fig.1). Pre-processing step inputs the log file that is generated by the LBS. The information necessary for our analysis includes the locations at certain time intervals and the transactions of incoming / outgoing calls.

Extracting the nontrivial scenarios unit matches subscribers' logs against each other to reveal the repetitive mutual transactions. If a selected set of agents keep calling each other within a certain time framework, and some of them perform correlated movements and meet, this unit tries to form a scenario for a shorter time span. If movements and calls are mostly correlated within this reduced time span, the scenario is considered as a nontrivial.

Assigning mental states to extracted behaviour unit involves the rule-based system that acts in the format similar to the one presented in Table 1. Each call is assigned a direction of information transmission: from given agent, to given agent, and both ways; belief states are assigned accordingly. Intention states are assigned given the information about who is a caller. If an agent *wants* to achieve a mental state of another agent (including being informed, or generation of an order by an addressee), she makes a call. A derived mental predicates such as *pretend, cheat, explain*, which are expressible in the basis of *want-believe* (extended BDI model, Galitsky 2003), can be assigned to a step of a scenario if there is sufficient evidence.

Multiagent Mental Simulator (Galitsky 2003), which is capable of yielding the abstract consecutive mental states, verifies the consistency among the mental states assigned to the scenario by the *rules.*

Detection of criminal scenarios unit applies the rules, which relate a scenario to a class of normal and criminal scenarios, revealing specific mental formulas (Fig.2). These formulas, as well as the templates for the criminal scenarios, are manually selected and evaluated by security personnel before the real-time functioning of the system. The template database is automatically filled by the *Detection of criminal scenario* unit to feed the machine learning - based predictor, which matches current scenarios against the templates to specify the probable future moves of the wireless subscribers.

Conclusions

In this paper we suggested a new way of extracting criminal behavior: mining the data of location-based services. As preliminary evaluation has shown, the algorithms of reasoning about mental states (Galitsky 2004), situation calculus and deterministic machine learning (Galitsky 2003), implemented as logic programs, are adequate to process the scenarios of the interactions of wireless subscribers to distinguish the patterns of their normal and criminal behaviors. Artificially generated data helped to create the initial set of (wireless) domain specific rules. Since the generic (domain-independent) reasoning and data processing subsystems have been deployed in a variety of applications and have been thoroughly verified, we believe that the development process of building the suggested system for the real-time data and embedding it in the infrastructure of security personnel will take rather short time period.

The model serves to deduce whether the behavior of a mobile phone user is normal or criminal, based on the totality of evidence collected (phone call location, duration, frequency), identified at a particular point in time.

Analyzing the mobility and geography in serious crimes has been verified (Saferstein 1990) to assist investigators in identifying the offender's home, place of work and recreation. Building efficient detectors of criminal behavior can reduce loss in all aspects of our life.

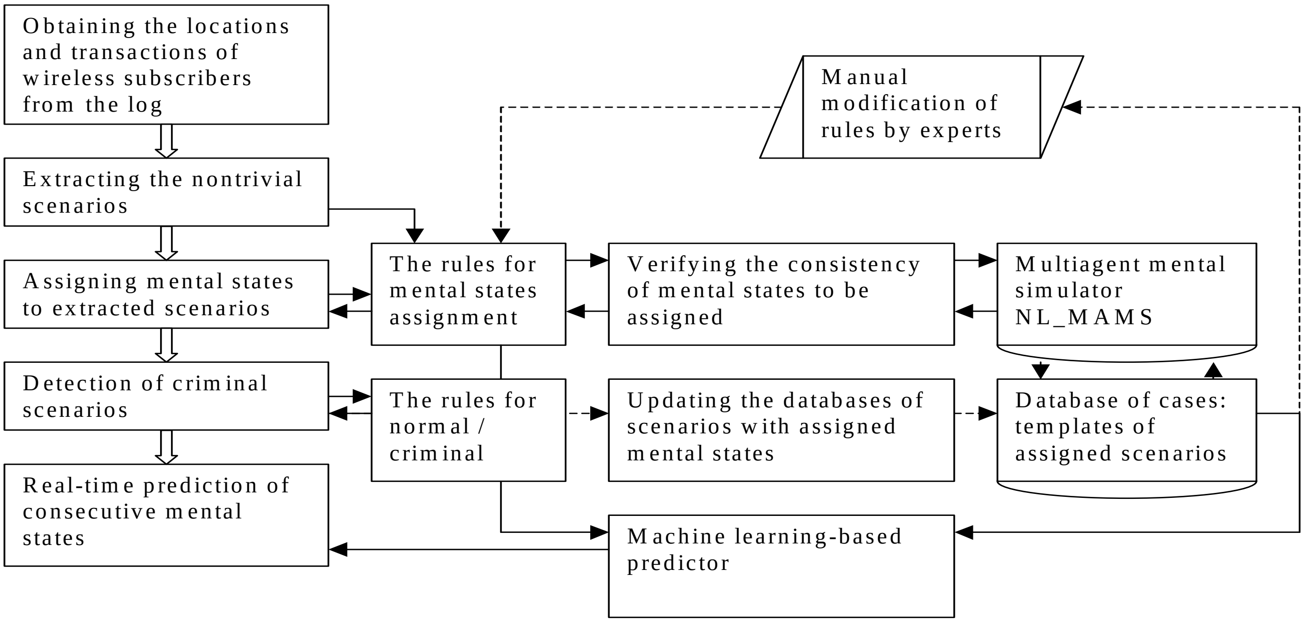
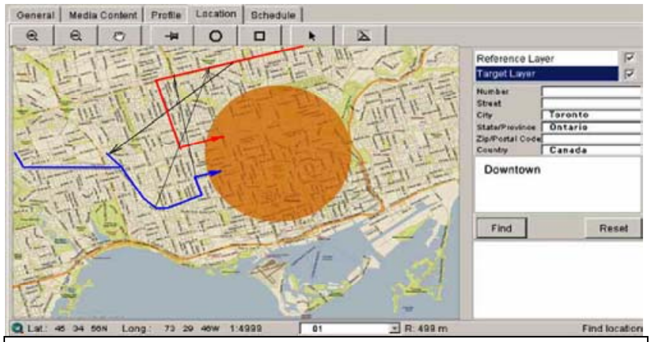

Obtaining the locations and transactions of wireless subscribers from the log
Extracting the nontrivial scenarios
Assigning mental states to extracted scenarios
Detection of criminal scenarios
Real-time prediction of consecutive mental states
The rules for mental states assignment
The rules for normal / criminal
Manual modification of rules by experts
Verifying the consistency of mental states to be assigned
Multiagent mental simulator NL_MAMS
Updating the databases of scenarios with assigned mental states
Database of cases: templates of assigned scenarios
Machine learning-based predictor

Fig. 1: Architecture of the criminal behavior detection system



want(car1, inform(car2, car1, directions)).
believe(car1, know (car2, directions)).
believe(car1, want(car2, meet(car1, car2))).
want(car2, not meet(car1, car2)).
want(car2, believe(car1, want(car2, meet(car1, car2)))).
Detected event: coordinated movement (leading and following cars)
Criminal behavior index: 3
Confidence level: 2

Fig. 2: Visualization of the scenario of multiagent interaction, mental state assignment and detected criminal behaviour. The interface is based on the Profilium (2005) software.

References

Profilium, Inc. 2005. http://www.profilium.com

Wherify, Inc. 2005. https://www.wherifywireless.com

Saferstein, R. 1990. Criminalistics: An Introduction to Forensic Science, 4th ed., Prentice Hall.

Galitsky, B. 2003. Natural Language Question Answering System: Technique of Semantic Headers. *Advanced Knowledge International*, Australia.

Galitsky, B. 2004. A Library of Behaviors: Implementing Commonsense Reasoning about Mental World. *8th Intl Conf on Knowledge-Based Intelligent Info Syst*. LNAI 3215 pp. 307-313.