Action Stories for Counter Terrorism

Elizabeth Guest [E.Guest@leedsmet.ac.uk], John Elliott [J.Elliott@leedsmet.ac.uk], Samantha Dixon [S.J.Dixon@leedsmet.ac.uk], Ah Lian Kor [A.Kor@leedsmet.ac.uk], Duncan Mullier [D.Mullier@leedsmet.ac.uk]

Leeds Metropolitan University

Leeds, LS6 3QS, England, UK

Extended Abstract

Introduction

Due to the raised terrorist threat worldwide, there is an urgent need to research that assists security and police services to protect the public and key assets and to prevent attacks from taking place. Successful protection and prevention may require potential and known suspects to be monitored or arrested. These operations are high risk because inappropriate surveillance, interview or arrest may have damaging political, public relations and intelligence effects. In addition to better tracking information on which to base suspicions, the security and police services need to have confidence that operations will yield evidence that can demonstrate conclusively that a deceptive activity such as a terrorist attack was in the process of being planned or executed before an operation takes place.

Many methods for mining data for identifying potential suspects have been proposed. This includes combining event calculus with UML (Unified Modelling Language) to manage a complex network in such as way that the user only requires domain knowledge (Ling et al, 2007), clustering methods, graph matching, latent semantic indexing of texts, and various kinds of machine learning techniques (Popp and Yen, 2006).

At LeedsMet we are developing novel methods based on combining a range of techniques from Artificial Intelligence and elsewhere to identify potential suspects from their “scent trails”. These trails are made up of data from location tracking, communications, and transactions. Our data is obtained by paying volunteers to play a location based game. This simplifies the problem, but much effort has gone into the game to make it as realistic as possible. The data obtained from this game is sufficiently complex that it should be relatively easy for the security services to adapt our method to the data they use.

Instead of applying data mining techniques to the data, we propose to read it into a spatial OLAP and store it in 3D format. Each player, team, and location will have an “action story” that can be extracted from this data. These stories will not contain all the data, but will consist of “features” designed to compress data that is not needed. For example if there are a few common routes used to get between A and B, we do not need to keep all those data points. However, if someone deviates from these, it may be worthwhile to record this.

Hybrid Reasoning Architecture

In capturing relevant and useful information for the purposes of modelling and extracting categorising behaviour, an integrated multi-dimensional space-time schema is being developed. The underlying rationales for aspects of this taxonomy draw from human cognitive and reasoning templates, which, in part, draw from biological systems adapted for Artificial Intelligence; often being fused or adapted beyond their conventionally prescribed use. The input data for this version of the system is a constrained representation of behaviour but, nevertheless, adequately complex for testing and deriving a system scalable for functioning with unconstrained real world information.

Sensory information forms the myriad of contextual clues that 'colour' and shade our thoughts, which result in our understanding of events occurring in our environment. Therefore, these Micronemes also form the primary stage of our quasi-hierarchical system that ultimately is designed to integrate processed information – akin to the Parietal Lobe – for the purposes of behavioural categorisation and constructing narrative from an n-dimensional vector space.

Constituent secondary stages of the system involve mapping data in space-time and creating behavioural k-line (knowledge line) representations. These k-lines are, in essence, a form of reductionism but like our own evolved method of memory store, hold the important components that recall and represent complex knowledge and events. These threads not only represent a complex linear 'process' but can point to and activate 'agents' of other stored representations that recall and add further pertinent detail. Such a system allows for reduced principle component modelling, whilst embedding common and reusable state frames. These embedded 'agents' then provide a network of associations and hierarchical knowledge that allow the capacity for a range of complexity, comprising varying levels of inheritance and descriptors. It is also submitted that this mechanism provides the facility for identifying subtle but important information significant in identifying abnormal behaviour.

In addition to a range of contextually related information, semantically associated information also forms part of any intelligent system. Again using rationales from biologically inspired architectures, our system design incorporates an artificial Parietal Lobe, associating and co-ordinating the different types of sensory input. From this co-ordinated hub of sensory modality, current information can be then related to established knowledge, inheriting and assessing events in context. This data fusion is essential for us as intelligent agents and therefore demands similar functionality in an artificially intelligent system. In any initial phase of 'training' or knowledge population, a feedback loop will create required 'world' knowledge from observation, forming base rules. These rules will then comprise initial truths from which inference mechanisms can later provide exception handling and refinement.

Complementing these upper levels of intelligent categorisation and interpretation of observations, the state changes that occur in space-time across the sensory modality, facilitate a further and crucial aspect of scene understanding; central to this is language and the facility of Trans-Frames. Our internal dialogue, which ultimately associates meaning to observed action and thought processes, labels all known, and even unknown, phenomena with degrees of probability through association. This 'labelling', in conjunction with inherited knowledge and attributed values of uncertainty [belief and plausibility], then form a system dialogue that is designed to contribute to an action story line that narrates the observed events. The range of comprising slots within Trans-Frames represents effects ranging from action, who performed what and why, purpose, resources utilised, and intentionality.

Data Analysis Methodology

The location-based phase of the 'game' will comprise three sets of data:

- Mobile phone communication data
- GPS location-based data with a time stamp (spatio-temporal data)
- Transaction data (includes the transfer of goods or services, e.g. building

Data collected will be organised according to the following data structure (schema):

- Communication data
    - <date, time, caller, receiver, start_time, end_time, call_duration>

- Spatio-temporal data
  - <date, time, member, latitude, longitude>
- Transaction data
  - <date, time, member, place, quantity> (note, if it is a transfer of services then, the default value for quantity will be zero)

Four sets of action stories will be abstracted from the raw data

- Communication-related action stories
- Spatio-temporal action stories
- Activity-related action stories
- Overlay of communication, spatio-temporal , and activity-related action stories

The social focal points (or actors) of the action stories will be:

- Individual dynamics
  - Individual only
  - Individual-individual
- Group Dynamics
  - Individual-group
  - Group-individual
  - Group-group

Supporting Theories and Techniques for the Analysis of Data Social Network Analysis for Communication Data also form an important cornerstone of our methodology. Krebs (2005) is a pioneer of Social Network Analysis (SNA) methodology which maps and measures interactions in the form of relationships, social ties or flows between individuals, groups, organisations, computers, or other information/knowledge processing entities. In this research, the SNA will provide both a visual and a mathematical analysis of mobile communication patters among the members of the observed population. In accordance to Cross and colleagues' (2002) recommendations, features that will be investigated on are:

- What is the communication flow within the observed population?
- Is there any peripheral member in the network?
- What are the sub-networks?
- What are the central nodes in a network that provide the only connection between the sub-networks?
- What are the isolation parts of the network (peripherals that are remote within the network)?

Other communication features that this research will look into are:

- Intensity of links between every two nodes in the network. The two dimensions of intensity that will be considered are:
  - Frequency count for the number of calls made/received between two nodes throughout the observed period
  - Aggregated duration of calls made/received between two nodes throughout the observed period
  - Peaks and troughs when frequency count is plotted against various temporal granularities (e.g. hour, day, week, month, quarter, year, etc.)
  - Peaks and troughs when aggregated call duration is plotted against various temporal granularities (e.g. hour, day, week, month, quarter, year, etc.)

Individual trajectories will be mapped to semantics of places – to provide inference about the possible activities (e.g. if part of an individual trajectory overlays with the highway then it could be inferred that the individual is driving, etc.). It is proposed that two laws will be employed to describe individual trajectories. Firstly, the Law of Contiguity in Qualitative Physics (principle whereby nothing passes from one state to another without passing through all the intermediate states) applied to space and time (Forbus, 1984). Secondly, the commonsense Law of Inertia in Cognitive Robotics (Shanahan, 1997) will be employed to solve the frame problem which is first described by McCarthy and Hayes (1969) when describing effects of actions or events using logic. Based on this law, only actions that effect significant changes will be considered and analysed.

Individual-individual Interaction
The type of individual-individual interaction will be of interest in our proposed analysis will be based on the Qualitative Trajectory Calculus (QTC) on networks (Bogaert et al, 2008), which will provide qualitative representations of trajectory pairs.

Group Dynamics
Features that will be investigated are: transient member/s of groups, links between subgroups that may assume a 'messenger' role, and time-based evolution of groups.

Data Fusion
A genetic algorithm will be employed to generate a terrorist suspect population and this will be the group that is most likely subjected to surveillance. Their respective communication and spatio-temporal data will be the focus of this project proposed data analyses (discussed above). The group/s that are abstracted from the SNA of communication data will be addressed as the 'virtually associated' group while the one/s from the spatio-temporal data will be known as 'physically associated' group. Two set operations that will be applied to these two groups are intersection and union. The intersection of these two will yield the inner core group while the complement of the intersection (note: the universal set in this context is the union of the two sets) will be regarded as the outer core group. The inner and outer core groups will be represented by the egg yolk and white respectively, which is adapted from Qualitative Spatial Reasoning (Cohn et al, 1997). Fuzzy logic will be employed to compute the degree of each membership (yolk or white) for ranking purposes.

Additionally, an artificial neural network (NN) will also form part of our hybrid solution. Experiments to date have been conducted using a multilayer back-propagation network with a sigmoid activation function. Many experiments were performed with variations on the type of neural engine, the NN architecture, the variables within the input patterns and the presentation order of these patterns, to ascertain the optimal solution; a ratio of 4:1 [Training :Test data] was used.

Tests on this NN indicate that it is successfully recognising and classifying non-terrorist patterns from which we can deduce terrorist behaviour. This method of classification is preferable as research indicates terrorist patterns change over time.

Conclusion

We realise that much of the terrorist behaviour will in fact, be normal behaviour so we will be looking for subtle clues that point towards deception; it is important to note that the system will not make the decisions. It is intended that this will provide a tool to assist the security services to sift through all the data to identify possible suspects. Individuals must be responsible for any decisions made using the system.

References

Bogaert, P., Van de Weghe, N., Cohn, A. G., Witlox, F., and De Maeyer, P. (2008). The Qualitative Trajectory Calculus on Networks. In: Barkowsky, Th, Knauff, M, Ligozat, G & Montello, D R (editors) **Spatial Cognition V - Reasoning, Action, Interaction**, pp. 20-38 Springer.

Cohn, A G; Gotts, N M. *The 'egg-yolk' representation of regions with indeterminate boundaries* in: Burrough, P & Frank, A M (editors) **Proceedings GISDATA Specialist Meeting on Spatial Objects with Undetermined Boundaries**, pp. 171-187 Francis Taylor. 1996.

Cross, R., Parker, A., and Borgatti, S. P. (2002). A Bird's Eye View: Using Social Network Analysis to Improve Knowledge Creation and Sharing. IBM Institute for Business Value. Retrieval Date: [03rd January, 2006], URL address: [http://www-1.ibm.com/services/us/imc/pdf/g510

-1669-00-a-birds-eye-view-using-social-network-analysis.pdf].

Forbus, K. D. (1984). Qualitative Process Theory. Journal of Artificial Intelligence, Vol. 24(1-3), pp.85-168.

Krebs, V. (2005). An Introduction to Social Network Analysis. Retrieval date: [3rd, January, 2006]. URL address: [http://www.orgnet.com/sna.html].

Ling, L et al (2007) Advances in intelligent information processing Source Information Systems archive Volume 32, Issue 7 pp 941-943 ISSN:0306-4379, Elsevier Science Ltd. Oxford, UK.

McCarthy, J., and Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, Vol.4, pp.463-502.

Popp, R. L., and Yen, J. eds. (2006) Emergent information technologies and enabling policies for counter-terrorism, IEEE Press.

Shanahan, M. (1997). Solving the Frame Problem. A Mathematical Investigation of the Common Sense **Law of Inertia**. Cambridge, MA: MIT Press.