INTRODUCTION AND BACKGROUND

Previously, a knowledge management (KM) framework has been proposed for the CrimeFighter toolbox. The framework identifies a number of processes that support humans in harvesting, filtering, storing, managing, analyzing, structuring, mining, interpreting and visualizing information about criminal and terrorist activities. These KM processes are individually or severally supported by a number of different tools as shown in Figure 1.

Our research is positioned within this framework and deals with the SNA Tools and Visualization Tools components. A previous prototype (CrimeFighter Assistant, Figure 2) to some extent supports the network analysis and visualization functions.

We have identified a number of shortcomings in existing theory and methods in complex network visualization and analysis:

- **Little focus on links (relationships) as first-class entities:** Links are fundamental to network structure and evolution, and provide at least as much relevant information about the network as nodes. In most network analysis methods, links are often considered secondary to nodes, and the rich information attached to them is filtered out.

- **Network models are too abstract or limited:** Networks have traditionally been studied in terms of very simple but useful models. However, when additional levels of abstraction are considered, most network measures generally don’t apply.

- **Imperfect data sets:** Covert network datasets are especially plagued by a number of issues—with fuzzy boundaries, they are nearly always assumed to be incomplete, some of the available data may not be completely reliable, and sometimes deliberately misleading and incorrect data creeps in. Network metrics are sensitive to imperfections in data, and often the erroneous addition or omission of a few links in the network can easily throw off the analysis results.

- **Little support for typed nodes and links:** Real-life networks are essentially heterogeneous with a lot of semantic information attached to nodes and links. However, in practice networks are often stripped of all this information for the ease of analysis. We believe that this information is vital, and should be retained to provide a more informed analysis of the network.

- **Limited support for prediction tasks:** Most prediction methods in complex networks are relatively new and address simple cases. There is little support for prediction in multi-relational networks.

RESEARCH FOCUS AND OBJECTIVES

Based on the shortcomings mentioned earlier, this research would specifically focus on the following tasks and associated questions:

- **Support network analysis with weighted links:** Considering links as first-class entities, formulate more holistic model of a network to capture variation in links operationalized as weights. How can network analysis take advantage of link weights (e.g., in determining key nodes)?

- **Mitigate effects of imperfect data on network analysis:** Which factors affect the reliability of network analysis (e.g., centrality measures) and to what extent? How can existing network analysis methods be improved to be more robust when some data is missing or erroneous?

- **Support analysis of heterogeneous networks:** Develop a more realistic model of a complex network capable of capturing heterogeneity at various levels of network composition—actors (node), relationships (links), and their attributes and semantics. What data model best addresses this requirement?

- **Support visual and exploratory network analysis:** With the ability to retain type information on nodes and semantic information on relations, provide the user with additional dimensions in visual and exploratory analysis of the network, including graphically querying heterogeneous networks. Does this scale well with large-scale networks?

- **Support prediction tasks in multi-relational networks:** Enhance existing algorithms for link-prediction in social networks to scale up to multi-relational networks. A number of single-relation link prediction methods exist based on a well-defined intuition; can one or more be adapted to work with multi-relational networks? Or we need a different intuition?