Proposed Methodology / Approach

Firtsly, Understanding the composition and evolution of terrorist group networks: A rough set approach.

Here in the beginning

* Selecting relevant features

The idea underlying the proposed approach is conceptualizing terrorist groups by using knowledge on the attacks they have commited. Thus, it is needed to select a relevant subset of the features provided by the GTD. Such subset must be suitable to describe the behavior of the terrorist groups.

* Building rough conceptualizations of terrorist groups

Rough set theory has been a methodology of database mining or knowledge discovery in relational databases.

We use rough set approach to discover structural relationship within imprecise and noisy data. A rough conceptualization can be constructed by applying the rough sets operators over the set.

Given the constructed rough set (V[t1,t2] , V[t1,t2]), the lower approximation is a conceptualization of the characterizing behavior of the terrorist group g , the upper approximation represents both the characterizing behavior of g and behaviors that can be possibly associated to g (e.g., a behavior adopted also by other terrorist groups).

* Building boundary regions for terrorist groups

Once a terrorist group’s behavior is summarized by means a rough set, it is possible to apply the Three-way Decisions Theory to classify the attacks (of a given terrorist group) in three regions: positive, negative and boundary.

A **theory** of **three**-**way decisions** is constructed based on the notions of acceptance, rejection and noncommitment.

The POS region includes all events certainly commited by the considered terrorist group and that characterize its behavior;

The NEG region includes all events that surely have not been commited by the considered terrorist group and that cannot characterize the behavior of such a group

The BND region includes all events that could or could not be been commited by the considered terrorist group and that possibly express its behavior

* Building Similarity matrix

In order to construct a network of possible relations starting from a set of terrorist groups (g1, g2, . . . , gn, )it is needed to calculate similarities between each group, where the set of rough sets representing conceptualizations of all terrorist groups g1, g2, . . . , gn, i.e., Gi is the rough set corresponding to the concept gi . **Using the calculated similarity measures it is possible to build a similarity matrix:**

* Designing the terrorist groups network

The last step is constructing the terrorist groups’ network by using the similarity matrix. Asssume W = (V, E), where W is the terrorist groups’ network is created as combination of, V is the set of all nodes of the network and E ∈ V × V is the set of all edges of such a network.

Finding influential nodes in social networks based on neighborhood correlation coefficient.

In the proposed approach, we consider common hierarchy between a node and its neighbors rather than the number of their common neighbors. To this end, the network is first divided into different parts. Nodes that are removed in the same iterations of the k-shell algorithm can be assumed to be in the same hierarchy.

Consider the graph in Fig for instance. This graph is composed of 3 shells that are shown by different colors. Inspired by the k-shell decomposition, the following process is applied to determine the iteration (IT ) in which the nodes are removed. At First, IT counter is set at 1. The nodes with the lowest degree are removed from the graph; these nodes are considered in the first hierarchy, and IT = 1 is assigned to them. In the next stage, IT counter is increased by one, and again the nodes with the lowest degree are removed from the graph. Removed nodes in this stage are considered in the second hierarchy, and IT = 2 is assigned to them. This process is iterated until there remains no node in the graph, and at each iteration IT counter is increased by one and assigned to those nodes removed in that stage.

K-shell decomposition is the method in which we can divide nodes on the basis of the number of its degree like nodes with degree 1 in one bucket etc.

Many centrality measures have been proposed to rank nodes in networks. A simple one is degree centrality, namely, a node with larger degree is likely to have higher influence (e.g., as an initially infected node, it is expected to spread more quickly and broadly) than a node with smaller degree. However, in some cases, this method fails to identify influential nodes since it considers only very limited information

the correlation between node vi and every vj ∈ Ni is calculated , This process is repeated for all nodes. the Extended Cluster Coefficient Ranking Measure ECRMi is calculated for each vi ∈ V. Finally, nodes are sorted based on ECRM index to obtain the ranking.

