Contents

I. Introduction

In order to indicate the origin of terrorism problems, the following background is worth mentioning.

1.1 Problem Background

Terrorism is the implementer of the non-armed organized use of violence or threats of violence, certain objects by placing terror to achieve certain political objectives behavior. The international community in some organizations or individuals to take kidnappings, assassinations, explosions, air hijacking, hostage-taking and other terrorist means, seek ideas and actions to achieve their political objectives or a specific requirement. [1]

Contemporary terrorism began to take shape in the late 1960s, usually in a series of explosions academia, assassinations, kidnappings, hijackings and other terrorist incidents occurred in Europe in 1968 as the starting point of contemporary terrorism, and in the decades that followed the years, terrorist activities intensified, until the '9.11' incident reached a peak. Currently, the definition of terrorism, the more common definition is: Terrorism is the use of violence or threat of use of violence; terrorism has a clear political purpose, in order to produce an atmosphere of terror, usually harm innocent civilians objects: Terrorism is an organized activity systems typically have extraordinary resistance, anti-moral, arbitrary, concealment, uncertainty and unexpected features.[2]

With the development of the network, terrorism has gradually developed a complex, subtle dynamic network model, and derive a series of new features, which makes for a more difficult fight against the terrorist organization. Unlike traditional networks, terrorist networks do not have a strict hierarchical structure, contact between members having arbitrary and hidden, so that terrorist networks showing loose, casual and covert organizational structure. Redundancy and flexibility of the network has become the main features of terrorism. In addition, even if the central node in the network of terror is removed, usually the network will not be destroyed. Because the redundancy and flexibility of terrorist networks so that it can quickly rebuild without losing its functionality. Moreover, concealment and lack of data also allows for information to combat terrorist networks become a problem.[2]

The fight against terrorism is imminent. We need to use scientific methods, the use of scientific weapons to fight the inhuman terrorist jihad. According to the terrorists we monitor e-mail exchanges between the peer network behavior, we use data mining technology network communication diagram between them build out and take advantage of big data processing techniques to analyze these data, to grasp their movements, help us prevent and combat terrorist attacks. Moreover, we can also use our data in the hands of the collapse of their correspondence, provide support and help for the complete removal of a terrorist organization.

1.2 Previous Research

The analysis of the terrorism has been a universal concern. Moreover, there have been a lot of research into the studies of risk index. For the risk index and the data mining of suspicious person, which are considered as important factors in solving terrorism, there are a lot of previous research that can facilitate our understanding of the model.

For example, In the winter of 2001, "the international social network analysis network" (INSNA) specialized online journal "Connection" for the problem would benefit several network analysis and terrorism has been specifically discussed.[3] Valdis E.Kreb is the first one to build the event of the 9/11 terrorist network model, his use of publicly available information to build the 19 hijackers were of personal relationships, and found a relationship between them.[4] Nasrullah Memon et al., rebuilt this model in 2007, were factional, sectarian and K- cluster computing and characters between events analyzed by the model, found that 911 key people and key events in the subgroups.[5] In 2007,D.M.Akbar Hussain proposes a new social network analysis methods to analyze the characters through a variety of attributes, building pluralistic property values for tasks, enabling a more comprehensive evaluation of the characteristics of a terrorist, and found a key figure. In the same year December, D.M.Akbar Hussain applied this model to 911 events of a terrorist organization, which found that a key figure in the terrorist network, and analyzed easily destroy terrorist networks and vulnerability.[6]

1.3 Our Works

We are asked to build a realistic, effective, and useful model to solve terrorism. Our model should not only consider the monitored people’s internet use and categorize the big data we get from the monitored people, but also other inessential factors that we need to consider, such as their social activities.

At first, we searched a big number of paper that studied the solution of terrorism and outliers to help us deepen the understanding of the problem. Validis E.Kreb is the first one to build network model, and then Nasrullah Memon improved that model in 2007.Later, we found out that the traditional network model cannot evaluate the terrorists’ risk index completely. A person educated, nurturing environment and religious factors will affect his future choice of values. So consider terrorists community property is a very necessary content. For example, many terrorist attacks are carried out in the name of religion, therefore investigate whether the monitored person's religious beliefs and terrorist organizations have countless ties to evaluate the person's risk index has a very important significance. Again, an unmarried staff required social responsibility and family responsibilities is less than a married person, the consequences of carrying out terrorist activities when he needs to be considered is small, and therefore more dangerous.

Based on the idea of big data processing have been obtained for data classification analysis and exception data processing, we selected KNN algorithms. KNN algorithm in 1968 proposed by Cover and Hart, is a relatively mature theoretical algorithms, and its working principle is: the existence of a set of sample data, and the data are present in each sample set label that we know the relationship between each data sample belongs classification. Enter the new data without a label, will each feature a new centralized data and sample data corresponding characteristics were compared, and algorithms to extract classification label features the most similar sample set of data. In general, we select only the sample data set before the k most similar data, which is the k- nearest neighbor algorithm source k, k is an integer of usually greater than 20. Finally, select the maximum number of classification k most similar data appear as a classification of new data. Those data, we call outliers, is to be classified out of the data. These anomalies need to focus on monitoring data, it is willing to be a sign that the terrorists carry out terrorist activities, which also provides technical support for our prediction and prevention of terrorist and criminal activities.

After analyzing all the data, we have come to the specific method how to evaluate human terror risk index and search suspects deliberate behavior. Finally, we wrote a letter to US President Barack Obama, the letter to the model we put forward on the basis of previous research, we recommend that the president should set the perfect terrorist risk index evaluation system, and to detect anomalies terrorist activity data, its strike, to prevent the tragedy from happening again.

2. General Assumptions

- We collected terrorist character of information and data related to terrorist attack s through public reports and databases.
- We are continuously monitoring the terrorists, the situation that information isn't collected does not exist.
- We don't consider the encryption information and don't take the special data into account.
- Some of the data is held in our hands can be clearly demarcate Category, the situa tion that the data is too ambiguous to classify isn't exist.
- The terrorist information is true, this model does not identify the authenticity of t he data.

3. Symbol Description

u, v are any two nodes of the graph theory.
$\mathrm{r}(u, v)$ is a binary variable that represents whether there is a link between the node u and the node v.
σ_{ij} is the number of shortest paths between i and j.

$\sigma_{ij}(K)$ is the shortest path between i and j through number of nodes K.
A is a n*n's adjacency matrix.
α is the proportional parameter.
β is the attenuation parameter.
N is the number of nodes in the network size and graph theory.
R is the relational matrix.
$\zeta(u,v)(W)$ is the dependent factor node u reach any other node by node v.
$occurence(u,v)$ is the number of node by node u to communicate with other nodes v.
$path(u,w)$ is the number of the shortest distance between the node u and the node w.
$d(u,v)$ is the distance between node u and node v.
k_i is the number of nodes adjacent nodes.
e_{ij} is the nearest paths between node i and node j
$AT_{i,T}$ is the matrix values of the members i of the task T.
|T| is the total number of tasks that the network members undertake.
W_r is the standard score of the monitored people get in religion
W_n is the standard score of the monitored people get in name
W_a is the standard score of the monitored people get in age
W_g is the standard score of the monitored people get in gender
W_m is the standard score of the monitored people get in marriage
W_{na} is the standard score of the monitored people get in nationality
W_e is the standard score of the monitored people get in duration
W_j is the standard score of the monitored people get in job
W_c is the standard score of the monitored people get in criminal history
W_s is the standard score of the monitored people get in social presence
W_{ab} is the standard score of the monitored people get in recruiting ability

4. Risk Index

4.1Traditional Evaluation Model

4.1.1The TREI Model

Network terrorist threat index of the monitored people can be solve by the following set of equations: