

An Easy-Using Evaluation System Model in Terrorism Based on Individual Risk Index and Classification

Jing Tang

College of Information Science and Engineering
Shandong University of Science and Technology,
Qingdao, Shandong, China
493117148@qq.com

Ning Zhang

College of Information Science and Engineering
Shandong University of Science and Technology,
Qingdao, Shandong, China
zning@sdust.edu.cn

Abstract: In this paper, we first establish the evaluation model of the individual risk index based on the social network analysis. The attributes of the individual through the behavior of the performance is divided into social attributes (including psychological attributes) and network attributes, and then to the social attributes of refinement, and that with some social attributes of people more dangerous. In which, we analyze the psychological factors, and the comprehensive evaluation of the factors, the network attributes are linked to the second tasks in the model. Then, we establish the model of double layer classification based on fuzzy clustering model and weighted Bias classification model. So as to make a preliminary classification of a large number of data, and finally get the specific behavior of the classification table "The Definition Sample of Cut Set of Fuzzy Matrix ". Then the weighted Naive Bayesian classifier is used to classify the individual behavior sets, so that each behavior is obtained from the correlation with terrorism. Then, we can use the weight and then return to the analysis of the individual can be derived from each person's expectations, the expectations of everyone, and applied to the first task of the risk index evaluation.

Keywords: Combat Terrorism; Individual Risk Index; Evaluation Model; Fuzzy Clustering; Weighted Bias Classification; Double Classification Model;

1. INTRODUCTION



Figure 1. Terrorism is becoming the focus of attention (Photo from Internet)

Since 9.11, terrorism has become one of the main factors that disrupts the international order, threatens national security and social stability. [1] With the development of network, modern terrorism has presented a series of new features and it has put a serious challenge to the national antiterrorist activities.

Recently, a group of the extreme religious jihad who calling itself ISIS (of Islamic State of Iraq and Short al-Sham) is increasingly rampant and anti-terrorism situation is imminent. Under the current situation, the terrorist activities have become more wanton rampant rather than Low-key. Today, in the Internet boom, they are distorted by instant messaging, social networks and other forms of propaganda and make more people join them. [2] [3] But a lot of people, who have psychological or mental problems, listen to their propaganda and then engage in terrorist activities. How can we find them, and make timely intervention? This is a proactive approach. Our paper will focus on this phenomenon to predict their behavior through the

network and data flow so that we can identify the people who real need "repentance".

2. THE DESCRIPTION OF THE PROBLEM

In our paper, we will analyze the following contents:

1. For first task, we will establish an evaluation model based on social network analysis of individual risk index.

By analyzing the social attributes (including psychological attributes) and network properties, we will determine whether a person is a terrorist and assess his (or her) acceptance of terrorism. Based on the outcome, we will give his (or her) proposed measures accordingly.

2. For the second task, we will establish a data classification model based on Naive Bayes classification. Based on the original big data provided by experts, we carried out a preliminary screening classification and obtain the status of each person's behavioral characteristics. At the same time, we can get the situation of terrorism in the region.

We believe that the information obtained by the task two will be helpful for the analysis of the individual risk index in the task one. And the specific role of this part of the mission will be specifically explained in the model of task one.

3. In the end, we will propose a reasonable suggestion to President Obama in response to the rampant ISIS in the near future. And at the end of the article, we will evaluate and summarize our model.

3. THE HYPOTHESIS OF PROBLEM

In the section, we use some symbols for constructing the model as follows.

We assume that:

1. The data that experts have provided is very useful for us. And they provide the Internet keyword about terrorist organization or terrorist activities.
2. We believe that the individual of the given data will not make physical and psychological changes within the time limit (72 hours).

4. THE DESCRIPTION OF SYMBOL

Table 1. The Description and Explanation of Symbol in this Paper

Symbol	Description and Explanation
$C_d(u)$	Point degree centrality
$r(u, v)$	It represents the node u, whether there is a link with v
$C_c(u)$	Approaching Centrality
$C_B(k)$	Mediation Centrality
$C'_B(k)$	relatively Centrality
σ_{ij}	Number of the shortest path between the node B and S
$\sigma_{ij}(k)$	Number between any two points i and j with k shortest path through the junction
a_{ij}	Relationship between nodes i and j
xi	Feature vector Centrality
λ	The maximum eigenvalue
μ	Scaling factor
C_{eff}	Graph Theory (network), the average efficiency
$E(G)$	Network Efficiency
d_{ij}	The shortest distance between nodes i and j
I	Individual Risk Index

5. MODELING

5.1 An Evaluation Model in Individual Risk Index Base on the Social Network Analysis

In recent years, with the dynamic social network analysis, we can find complex covert network, and through a series of algorithms to estimates people who have been in contact with terrorism individuals risk index, in order to determine a key

figure what we need. And then take corresponding measures to solve it.

In terms of the key figures digging, Dr. Nasrullah Mammon proposed concept of Network Efficiency, Location of Role Centers and Dependence centrality based on Traditional Center index. It provided calculating individual risk index with a new set of indicators. Dr. D. M. Akbar Hussein proposed a new method to calculate individual risk index, these are of high academic value in the key figure in the fight against terrorist networks.

5.1.1 The constitution of Individual Risk Index Assessment

Individual risk index is divided into two parts, social attributes and network attributes. By calculating the two parts of the property value, we can get an individual preliminary risk index of who have been in contact with terrorism.

1. Social attributes, social attributes are the measurement indicators based on the basic situation of the individual with the person who has been exposed to terrorism. Generally considered, individuals with some social attributes are more threatening than others, so it could potentially be a key figure. According to the analysis of the relevant terrorist information display [4] [7] [10] [11] [14], people who have had contact with terrorism have a greater similarity in the following 14 attributes. Among them, the psychological factor accounted for a large proportion of the comprehensive measurement.

Therefore, we mainly select these 14 social attributes, the specific description of the table as follow:

Table 2. The selected social attributes (Total 14)

Attribute name	Explain
Name	Name appeared in the international or national terrorist files are more threatening
Age	In general, young people are more enthusiastic and aggressive, more than the elderly threat
Sex	Men are often more threatening than women.
Marriage	Married people tend to be stable, so the threat is relatively small.
Nationality	People who come from the countries of terrorism are more threatening.
Religion	People who have faith and believe in religion are more threatening than those who are not.
Education	Under normal circumstances, the higher education level, the greater the threat.
Occupation	People engaged in blasting, chemical, physical, transportation, computer and finance and other industries have professional skills, able to provide technical support for terrorist organizations, funds and other support, and therefore may be more threatening

Attribute name	Explain
Crime history	It can be considered that the people who have the history of crime are more threatening.
Psychological index	People who have a certain degree of psychological problems (For example, depression, or dying) are easier access to terrorism, the greater its threat
Comprehensive measurement	According to the data collected by the intelligence, According to the comprehensive index of its behavior, from this comprehensive indicator of their daily behavior, then give the impact factor
Participate in the number of terrorist activities	People involved in many terrorist activities the more experienced, therefore more threatening
Social features	If a person's social appearance is bad, compared the good person, can be considered as a threat to the greater
Recruiting ability	If someone has the ability to recruit terrorists, it is more threatening.

2. Network attributes. Network attribute is the measurement index determined by a series of social network analysis method. Social attributes are based on the personal characteristics of the individual with the contact with terrorism, there are some subjective and one-sided. [5] [9] [12] Here, we analyze the network properties based on social network analysis, to be more objective, comprehensive evaluation of the risk of exposure to the individual risk index. In social network analysis, the center is one of the most important indicators. Therefore, we determine the network properties based on central degree. The main network attributes include: Center of point, Proximity Center, Center of media, Eigenvector centrality and network efficiency.

1) Point degree center. The center of the point is the typical representative of the center; it refers to the degree of concentration as determined by each node in the network diagram. Index of individual activity, it is the most direct and simple way to measure individual centrality. This concept comes from the concept of "Star" in the social metrology. In terrorist activities, terrorists are often more important to have a relatively high degree of centrality. If a person has been exposed to terrorism has a direct connection with many other members of the terrorist activities, then it can be considered as an important platform for the members to obtain and exchange information in the network.

The most common definition of the center of the point is:

$$C_d(u) = \sum r(u, v)$$

Among them, $r(U, V)$ is a two variable, which is used to indicate whether there is a link between u and v .

(2) Close to the center. Close to the center of the network is based on the distance between a node and other nodes. The shorter the total distance, the higher the degree of the network. Close to the center of the information to measure how long it can spread from one node to the other nodes in the network,

shows the close degree between the actors and the other actors. In terrorist activities, the individuals who are close to the central level are in the central position of the contact with terrorism, they can get in touch with other members of the terrorist activities through a minimum of intermediate position, so that more independent of the other terrorists.

For the most common definition of proximity to the center:

$$C_c(u) = \frac{1}{\sum_{v \in V} d(u, v)}$$

Among them, $\sum_{v \in V} d(u, v)$ representing the distance of any point $u \in V$ in graph theory $G = (V, E)$. It should be noted that, unlike the point of the center of the degree, the proximity of the center is a measure of the overall network.

(3) The center of the media. Medium centrality is the frequency that a node appears on the other nodes to contact the path, and measure the degree of the connection between the nodes as the other nodes. Terrorist activities, a node that has a low degree but it contacts with a terrorist usually has a high degree of mediation, such a person plays the role of a broker or a goalkeeper to control the communication and exchange of information. [13] [15] [16] [17] But in the real environment, communication is often not strictly in accordance with the network path, in view of this situation, some scholars put forward two kinds of methods to measure the Intermediate Center. The first one is based on the calculation of all the possible paths between a pair of nodes, and the second is based on the random path calculation.

At present, the most common definition of the center of the center is:

$$C_B(k) = \sum_{i \neq j \neq k \in V} \frac{\sigma_{ij}(k)}{\sigma_{ij}}$$

Among them, σ_{ij} is the number of shortest paths between B and S , $\sigma_{ij}(k)$ is the shortest path through the number of nodes K between i and j .

However, in the real world, a pair of nodes often has many paths, which makes the calculation of the center of the media becomes very complicated. Similarly, the computation of $C_B(k)$ is also influenced by the network size. By dividing the K by dividing the number of nodes outside the $(n-1)(n-2)$ can be standardized, the relative center degree is obtained. Freeman gives the definition of relative centrality:

$$C'_B(k) = \frac{C_B(k)}{(n^2 - 3n + 2) / 2}$$

(4) Feature vector center degree. The concept of feature vector centrality is firstly proposed by Bonfish. Subsequently, the PageRank algorithm of Google improved it. The nodes defined by the center of the feature vector are the nodes which are in the central position in the central position. Let express an Adjacency matrix of $n \times n$. If the node i and node j by the line directly connected, then $a_{ij} = 1$; Conversely, $a_{ij} = 0$. Because A is a symmetric matrix, so its eigenvalues are real numbers, the feature vectors are orthogonal and can be used to be of the right. Then the feature vector of the center node x^i is defined as the maximum eigenvector of the A in the i standard. The λ

for maximum eigenvalue, the corresponding eigenvector x , then $Ax = \lambda x$, then $x = \frac{1}{\lambda} Ax$, and as follows:

$$xi = \mu \sum_{j=1}^n a_{ij} x_j$$

Proportional coefficient $\mu = \frac{1}{\lambda}$, This xi is proportional to the center of the eigenvector of the other nodes connected to it.

(5) Network efficiency. Complex networks can be understood from the point of view of efficiency, For example, it is believed that a network with a small world character is very efficient in information transmission. Network efficiency $E(G)$ is a measure of the nodes in a network to exchange information efficiency indicators. In order to define the efficiency of M , we first calculate the shortest distance d_{ij} between nodes J and I . Assume that each node sends a message through their contact. The communication efficiency of the node i and j is inversely proportional to its shortest distance.

When there is no path between the nodes 'i' and 'j', $d_{ij} = +\infty$, and the network efficiency is 0. Suppose N is the number of nodes in the network size or graph theory. The average efficiency of graph theory (Network) G can be defined as:

$$C_{eff} = E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} (C_{eff} E \in [0,1])$$

From the above formula can be obtained in the $[0,1]$ interval value.

5.1.2 Individual Risk Index Calculation in who have been in contact with terrorist

The traditional social network analysis method often evaluates the key figures only through a series of network indicators, and do not consider the individual with the exposure of terrorism has its own characteristics.

This is a good way to ensure the objectivity of the evaluation results, but in the practical application, it will cause the one-sided and inaccurate. We learn from key figures evaluation method proposed by D. M. Akbar Hussein. On this basis, we introduce the degree coefficient of event participation and the individual risk index evaluation system to measure and calculate the risk index, and then evaluate the key figures. [19]

1. Property classification and evaluation methods. Different methods for classifying and assigning attributes are the key to build individual risk index evaluation system. First of all, the classification method of the attributes is explained by the religious attributes (See table below).

In comparison with those who have no faith, religion is more threatening.

Table 3. Religious attributes

Attribute name	Possibility	Fraction
Religion	A	4
	B	3
	C	2
	Faithless	1

In order to avoid controversy, we use A, B, C to replace different religions and specific religion can be selected

according to the specific circumstances in the actual analysis. In the above table, we divide our religious attributes into 4 levels, that is, A, B, C, and no religion. Based on previous experience of terrorist incidents, it can be found that the terrorists often carry out terrorist activities in the name of religion.

In comparison with those who have no faith, religion is more threatening. At the same time, in the case of specific terrorist incidents, such as religious conflicts caused by terrorist incidents, terrorists tend to have a common religious belief. In this case, people who believe in a particular religion are more likely to be terrorists than those who believe in other religions. Therefore, for a terrorist event, it can be considered that the people who have his belief of a religion (such as A) are more threatening and more likely to become a key figure. Therefore, for religious A, we give the highest score - 4 points. A or B except for C, we give the lower score-3 or 2, respectively. The people who have no faith are considered to be the least threatening, so it is given the lowest score of 1 point. [21] [22]

It should be noted that individuals engaged in terrorist activities are often concealed. Due to the impact of the national intelligence agencies, their information may be difficult to obtain. If the specific information of a terrorist is not obtained, then the attribute is given a minimum of 1 point. In another example, the table below grade social landscape attributes as follows:

Table 4. Social features

Attribute name	Possibility	Fraction
Social features	(suspected) Terrorism	5
	Violence / Racism / Drug / Crime	4
	Solitary / Independent	3
	General	2
	Good	1

Psychological indicators are shown in the following table:

Table 5. Psychological Index property

Attribute name	Possibility	Fraction
Psychological Index	Violent mental illness/ Serious violence obsessive-compulsive disorder / Mental illness delusions of persecution	4
	With excitement tendency of psychological disease/With mild, moderate excited state of mind	3
	Have depression type mental disease / Have other mild mental disease	2
	Mental normal / Mental disorders	1

The comprehensive measurement attributes are shown in the following table:

Table 6. Comprehensive measurement attribute

Attribute name	Possibility	Fraction
Comprehensive measurement	A lot of contact with the nature of the content related to terrorism	5
	Having contact with the contents of a terrorist nature	4
	Who do not have contact with the terrorist nature of the content / Who have contact minimal nature of the content.	1

The rest of the property will not go; their design principle is the same.

Of course, these selected properties are not absolute. Depending on terrorist activities and their structure, we can adjust the content of the above indicators and make it more accurate estimates.

2. Individual Risk Index calculation steps.

Individual risk index calculations are divided into the following steps:

- 1) The social attributes and network properties that we select are subdivided into several stages and we give a different score for each level. On this basis, we will build individual risk index evaluation system.
- 2) According to the information collected from the members of the terrorist activities, we will use the individual risk index evaluation system to score;
- 3) We can calculate the risk index of each node of the network by using the attribute values of each member to sum and multiplying its corresponding event participation degree coefficient. By comparison, we can determine the position and importance of each node in the network.

Expressed by the formula:

$$I = P \sum_{i=1}^n (S_i + N_i)$$

Finally, the key figures in the network can be obtained by comparing their individual risk index. It can be considered that the individual risk index value is the key person in the network. It is worth noting that they are not necessarily at the center of the network.

5.2 A Double Classification Data Model Based on Fuzzy Clustering Model and the Bayesian Classification Model with Weighted

From the point of view in behavior, the individual long-term network behavior has a certain stability. And may correspond to social reality people's social behavior. Therefore, analysis of network behavior is important. However, depending on the focus of research and application, the classification method is different. So, faced with a particular problem, we need to classify from a particular angle.

In the face of a large amount of data, we preprocess the data first, using small sample to get the distance matrix, and then the fuzzy clustering method for a large number of data in the network behavior of the initial classification. Then use the weighted Bias classification method to set up the key words in the sample, and then segment the data which has already been classified, and filter out the content we need. [23] [24]

5.2.1 Data pre-processing

Take out the data from the 100 IP addresses, from a pair of multiple perspectives, the corresponding behavior of each IP address to get out, a table as shown below. That each action corresponds to a value. The value should be determined according to the actual situation, for determining the classification of indicators.

As shown in the table below, the $X_1, X_2, X_3 \dots X_n$ score table for different user behavior observation points:

Table 7. User score table for different behavior observation points

Behavior User	X_1	X_2	X_3	...	X_n
A	a_1	a_2	a_3	...	a_n
B	b_1	b_2	b_3	...	b_n
C	c_1	c_2	c_3	...	c_n
D	d_1	d_2	d_3	...	d_n

First, the average value of the sample of each behavior is calculated:

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ki}$$

And standard deviation:

$$S_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ki} - \bar{x}_k)^2}$$

Secondly, the standard value of each data XXX, and the standardization of data compression to YYY closed interval:

$$x_{ki}^* = \frac{x_{ki}' - x_{k \min}'}{x_{k \max}' - x_{k \min}'}$$

($x_{k \max}'$ and $x_{k \min}'$ said the maximum value and the minimum value respectively)

5.2.2 Fuzzy clustering behavior pattern classification

Now, x_i and x_j are two of the data that have been processed, and the function $d(x_i, x_j)$ satisfies the following conditions:

1. Non negative:
2. To all x_i and x_j , $d(x_i, x_j) = 0 \iff x_i = x_j$
3. Symmetry: for all x_i and x_j , $d(x_i, x_j) = d(x_j, x_i)$
4. Triangle inequality: for all x_i, x_j, x_k , $d(x_i, x_j) + d(x_j, x_k) \geq d(x_i, x_k)$

Then, in the p-dimensional Euclidean space, the Euclidean distance between any two elements I and j is defined as the square root of the difference of the vector:

$$d(x_i, x_j) = \left[\sum_{s=1}^p (x_{is} - x_{js})^2 \right]^{1/2}$$

In order to avoid the large difference between the results of a one-dimensional distance, it is necessary to carry out a linear transformation, the standard of all variables, each variable will be subtracted from the mean, then divided by the square root. After that, each variable has zero mean and unit variance, with the same weight to participate in the distance calculation.

In order to obtain the grouping of the vector of all network user behavior, the distance of all the vectors is usually considered, and the distance of the 100 samples is also expressed as the distance of the sample:

$$\begin{pmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ \dots & \dots & \dots & 0 & \\ d(100,1) & d(100,2) & d(100,3) & \dots & 0 \end{pmatrix}$$

Among them, $d(x_i, x_j)$ represents the distance between the behavior of i and the behavior of j.

After that, the cut sets of fuzzy matrices are defined, and the classification of the large numbers of data can be classified:

Table 8. The Definition Sample of Cut Set of Fuzzy Matrix

Category	Access to Information	Communication	Leisure and Entertainment	E-commerce	...
Body behavior	News, information acquisition, work related information, search engines, electronic journals, online search, browse the web, read information, online	Online chat, email, SMS, BBS forum, community, discussion group, free personal home page, Internet phone, alumni, RENREN, Wechat, Tieba...	Online games, multimedia entertainment, video viewing, live online, reading	Online education, software upload/download, online hospital, online banking, online design	...

Category	Access to Information	Communication	Leisure and Entertainment	E-commerce	...
	reading				
Keywords	a_1, \dots, a_n	b_1, \dots, b_n	c_1, \dots, c_n	d_1, \dots, d_n	...

5.2.3 Weighted Bayesian behavior pattern analysis

In this context, we preprocess the raw data and classify it based on fuzzy clustering. By the above classification, we can know what behavior is the user's risk behavior through the data acquired by the large data. On this basis, we will again through the Bias classifier to the specific behavior of the key words for a classification. This classification is mainly to have good user behavior and the abnormal behavior of the two classification. In order to effectively apply the data after the second classification, we will use the method based on the weighted Bias classification to analyze the behavior patterns that have been collected. The expression of the discriminant function is as follows:

$$e(x) = \arg \max_{1 \leq j \leq m} p(C_j) \prod_{k=1}^d p(x_k | C_j)^{x_k}$$

Based on the reparability judgment, we can obtain the final weights. We use JF(k) to represent the function, which means that the weight of the right of the classification of each feature.

We use $D = (z_1, z_2, \dots, z_n)$ to represent the specific behavior of the keyword set (the following abbreviation for the sample word set). A_1, A_2, \dots, A_d represents a set of continuous variables. $C = \{C_1, C_2, \dots, C_m\}$ is used to represent the set of samples, and j is expressed by $(z_1^{(j)}, z_2^{(j)}, \dots, z_n^{(j)})$. And $j = 1, 2, 3, \dots, m$. For any attribute A_k , it is compared with the variable of the class object concentration. The expression of the C_j component of A is calculated as follows:

$$\mu_{jk} = \frac{1}{n_j} \sum_{j=1}^{n_j} z_{jk}^{(j)}$$

The total mean value is:

Table 9. Weighted Bias classification weight based on fuzzy classification of specific behavior

	a_1	a_2	...	a_n	b_1	b_2	...	b_n	c_1	c_2	...	c_n	d_1	d_2	...	d_n	...
Access to Information	w_1	w_2	...	w_n	0	0	...	0	0	0	...	0	0	0	...	0	...
Communication	0	0	...	0	w_{n+1}	w_{n+2}	...	w_{2n}	0	0	...	0	0	0	...	0	...
Leisure and Entertainment	0	0	...	0	0	0	...	0	w_{2n+1}	w_{2n+2}	...	w_{3n}	0	0	...	0	...
E-commerce	0	0	...	0	0	0	...	0	0	0	...	0	w_{3n+1}	w_{3n+2}	...	w_{4n}	...
...

$$\mu_k = \frac{1}{n} \sum_{j=1}^n z_{jk}$$

Class distance difference:

$$SB(k) = \sum_{j=1}^m n_j (\mu_{jk} - \mu_k)^2$$

Class within the difference:

$$SW(k) = \sum_{j=1}^m \sum_{i=1}^{n_j} (z_{ik}^{(j)} - \mu_{ik})^2$$

The value of JF(k) can be obtained:

$$JF(k) = \frac{SB(k)}{SW(k)}$$

In this function, SB (k) is used to describe the dissimilarity between the class t class and SW (k) is used to describe the dissimilarity of different samples within the same class. According to the definition of the criterion function, we can infer that the value of JF (k) is becoming more and larger and belongs to the same class. At the same time, the degree of similarity between different samples is lower. With the increase of the value of JF(k), the classification of its types is gradually enhanced.

For this function, when the value of JF is above the average, in order to improve the function of it, we should increase the weight to more than 1. When the value of JF is less than the average, the weight should be reduced to less than 1 in order to weaken the function of tithe method of establishing weights is given below:

Assuming the weight of Fisher is V_k :

$$V_k = \frac{JF(k) \cdot d}{\sum_{k=1}^d JF(k)}$$

So the definition of weighted Bias classifier based on Fisher criterion function is as follows:

$$P(C_j | y) = \frac{\alpha_j \prod_{k=1}^d \frac{1}{\sqrt{2\pi}\sigma_{jk}} \exp\left\{-\frac{(y_k - \mu_{jk})^2}{2\sigma_{jk}^2}\right\}^{V_k}}{\sum_{j=1}^m \alpha_j \prod_{k=1}^d \frac{1}{\sqrt{2\pi}\sigma_{jk}} \exp\left\{-\frac{(y_k - \mu_{jk})^2}{2\sigma_{jk}^2}\right\}^{V_k}}$$

By using the above model, we have made the two processing of the data that have been processed before the specific behavior of the key words. So we can get the results, and be able to judge the situation of the regional terrorism related

communication activities, as well as the impact of the key social or network factors.

In the table, the value $w_1, w_2, \dots, w_n, \dots, w_{4n}$ presents the weights given by each keyword after the classification has been classified by weighted Bias. Based on the weight of keywords, we can know what the specific content is similar to this kind of. In using this model, we will prepare a part of the similar sample data to be simulated and get the weights.

Therefore, the weights and the results from the previous training simulation are compared with the results of the previous training simulation, which can find out the individual and regional terrorist activities. **Title and Authors**

The title (Helvetica 18-point bold), authors' names (Helvetica 12-point) and affiliations (Helvetica 10-point) run across the full width of the page – one column wide. We also recommend e-mail address (Helvetica 12-point). See the top of this page for three addresses. If only one address is needed, center all address text. For two addresses, use two centered tabs, and so on. For three authors, you may have to improvise.

6. MODEL SUMMARY AND IMPROVEMENT

With the improvement of living standards, network security has become a hot issue of social concern, with the help of the network, the activities of terrorist organizations have been greatly expanded, and the research of network user behavior has become an important research topic. In the first question, we established the evaluation model of the individual risk index based on the social network analysis, and described the evaluation of the key figures in the terrorist network. In this second question, we first preprocess the data, and then we establish a fuzzy matrix model and a data classification model based on weighted Naive Bayesian classification to achieve a fast data classification, and we can get the information of each weights, laying the foundation for the next step in the analysis of personal information. This is of great significance to the management of network security. The advantages and disadvantages of the two models in this paper are as follows:

6.1 The advantage & disadvantage of Evaluation Model in Individual Risk Index

Advantages:

1. The personal attributes are classified into social attributes and network attributes, which are described, and the corresponding evaluation index system is established. We can take advantage of

the rank difference of two attributes; analyze what kind of activity that a person has to do, in order to make the concrete action to prepare.

2. In this model, a Meta network (human relationship network) and multi - Network (task network, skill network) are described.

Disadvantages:

1. This paper is based on the static network model, and it is from the perspective of the individual, it can be further extended by the key individuals to find a key group, from the perspective of dynamic network to construct a network model of terrorist groups.
2. Because this model relates to social and network two attributes, so in this based ranking, it is possible to weaken the influence of a certain attribute, and thus cause the error.

6.2 The advantage & disadvantage of Data Classification Model

Advantages:

1. The weighted Naive Bayesian algorithm can estimate the required parameters based on a small amount of training data. Based on these, we can classify a large number of data.
2. The weighted Naive Bayesian algorithm is the simplest algorithm in the Bias classification algorithm and it can be obtained by the data after the classification of the data on the basis of each data classification.
3. It has great flexibility for different types of data and different requirements of data classification. It also can establish a relatively stable classification model and not rigidly adhere to several data classification. And according to the actual needs of information classification, the model can be improved.
4. Bias network has a strong ability to deal with uncertainty problem. Bias network uses conditional probability to express the relationship between the various elements of information. It can learn and reason under the conditions of limited, incomplete and uncertain information. So we can effectively integrate information with the key words of our hypothesis.

Disadvantages:

1. In this model, it can be used to classify the data in the algorithm. So it is not possible to achieve the actual application. There will always be some errors.
2. The use of this model is to know a prior probability. It means that the model only can be obtained by a large number of data learning. This requires that a large amount of data is required to train the model to improve the accuracy of the model.
3. Classification decision error rate. The main reason for the induction of these errors is due to the problems in the training process. The main reasons

may be that the model design is unreasonable; the model training data is not sufficient or not typical and so on.

6.3 The advantages and disadvantages of fuzzy matrix classification model

Advantages:

1. Many concepts in real life are very difficult to distinguish between "with" and "no". Therefore, fuzzy clustering is used to describe the relationship between the behavior and the effect of the initial classification.

Disadvantages:

1. The selection of the fuzzy matrix has some errors, too large and too small can lead to classification error.

7. REFERENCES

- [1] Popp R, Poindexter J. Countering terrorism through information and privacy protection technologies [J]. Security & Privacy, IEEE, 2006, 4(6): 18-27.
- [2] Liu X. P. Modeling method for terrorist social network based on social network analysis [D]. Changsha: National University of Defense Technology, 2010.
- [3] Shang R., Sun D. Y. A key figure in the evaluation based on the individual threat index [J]. Journal of Wuhan Public Security Cadre's College, 2011, 25(4): 83-87.
- [4] Shang R. Individual Threat Index Based Key Member Detecting Research in Covert Network [D]. National University of Defense Technology, 2011.
- [5] Long J. A Study on Tag-based Folksonomy on the Internet [D]. Beijing: Peking University, 2007.
- [6] Zhang H. A Study on Terrorist Organization from the Perspective of Social Network Analysis [D]. Changsha: National University of Defense Technology, 2010.
- [7] Wang Q. W., Deng X. Z. A Study on Mental Models of Web Users' Information Behavior in Browsing [J]. Document, Information and Knowledge, 2010 (5): 93-96.
- [8] Rong G. J. Mathematical Models of Relationship between Psychology and Behaviors in Clothing Wearing and Purchasing [J]. Journal of Beijing Institute of Clothing Technology, 1989, 1: 006.
- [9] Liu T., Bu ZH. X., Wang G. H. Based on psychological factors and behavioral patterns of consumers brand choice model [J]. Enterprise economy, 2015, 7: 009.
- [10] GUO Z. Y. On the social Psychological model of the Psychological structure of morality [J]. Journal of Shaanxi Normal University (Philosophy and Social Sciences Edition), 2000, 29(2): 151-157.
- [11] GAO W. B., Chen Z. Y. A Study on Psychopathology and Psychotherapy of Internet Addiction [J]. Advances in Psychological Science, 2006, 14(4): 596-603.
- [12] Xue L. Study of Mental Pattern and Utilization of Social Support of Drug Addicts. [J]. China Journal of Health Psychology, 2007, 15(9): 833-835.
- [13] Wan H. Y. Research on Link-based Classification in Social Networks [D]. Beijing Jiaotong University, 2012.
- [14] DING H. J., CHEN D. J. ISIS's Cyber Terrorism Activity Impacts on China's Counter-terrorism and the

- Countermeasures [J]. China Public Security. Academy Edition, 2015 (2): 95-97.
- [15] Conway M. Terrorist 'Use' of the Internet and Fighting Back [J]. Information and Security, 2006, 19: 9.
- [16] Kuang W. B. On characteristics of netizens' behavior in special times [J]. Journal of International press, 2003, 5: 004.
- [17] Zhou R. T., Zheng B. Folksonomy New Information Classification System in Internet Age [J]. New Technology of Library and Information Service, 2006 (3): 72-75.
- [18] Zhang N. Analysis of Folksonomy System Users' Behavioral Characteristics [D]. Taiyuan: Shanxi University, 2013.
- [19] Li Y. G., Li X. CH. Area information of Measure Model and its application based on variable weights [J]. Journal of Intelligence, 2006, 25(2): 107-109.
- [20] Li D. D. Research on the Methods of Network Traffic Classification Based on Behavior [D]. Harbin University of Science and Technology, 2015.
- [21] Ren W. J., Chen J. Y. Research of web-based user behavior analysis [J]. 2012.
- [22] Ma G., Du Y. G., An B., etc. Risk Evaluation of Complex Information System Base on Threat Propagation Sampling [J]. Journal of Computer Research and Development, 2015, 52(7): 1642-1659.
- [23] Song N., Fu J. L., Bao Q., etc. Cyber terrorism spreading and optimal intervention policies based on a scale-free network [J]. Systems Engineering—Theory & Practice, 2015, 35(3): 630-640.
- [24] Liu CH. Y., Hu X. F., Si G. Y., etc. Public Opinion Propagation Model Based on Small World Networks [J]. Journal of System Simulation, 2007, 18(12): 3608-3610.
- [25] Li Q., He H., Fang B. X., etc. Awareness of the Network Group Anomalous Behaviors Based on Network Trust [J]. CHINESE JOURNAL OF COMPUTERS, 2014, 37(1): 1-14.
- [26] Zou L. J. Research on Clustering and Classification Algorithm of Streaming Data [D]. Yangzhou: Yangzhou University, 2008.
- [27] Sui ZH. Research and Implication of Group Network Behaviors Model [D]. Sichuan: Sichuan University, 2006.
- [28] Li ZH. H., He J. L., Wu P. F. The Time Period Characteristic of Information Communication Model and its Management Strategies of Paroxysmal Public Crisis [J]. LIBRARY AND INFORMATION SERVICE, 2007, 51(10): 0-0.
- [29] Dong F. Q. Study on Network Users' Behaviors Analysis and its Application [D]. Xi'an: Xidian University, 2005.
- [30] Ji R. M. RESEARCH AND DEVELOPMENT OF WEBPAGE TEXT KEYWORDS FILTERING [D]. Inner Mongolia University, 2012.
- [31] Yao ZH. W. Dissemination of modern terrorism steering [J]. Journal of Yunnan Social Sciences, 2012, 2: 026.
- [32] Dai Y. RESEARCH AND DESIGN OF USER BEHAVIOR ANALYSIS AND CLASSIFICATION IN UNIVERSAL NETWORK [D]. Beijing University of Posts and Telecommunication, 2013.