Terrorist Threat: Human Factor

Vasily Popovich, Manfred Schrenk, Vasily Osipov, Filipp Galyano

(Prof. Vasily Popovich, SPIIRAS, 14 Liniya, 39, St. Petersburg, Russia, popovich@mail.iias.spb.su)
(Manfred Schrenk, CEIT - Central European Institute of Technology & CEIT ALANOVA Department for Urbanism, Transport, Environment and Information Society, Am Concorde-Park 2, Gebäude F, A-2320 Schwechat, m.schrenk@ceit.at)
(Prof. Vasily Osipov, SPIIRAS, 14 Liniya, 39, St. Petersburg, Russia, osipov_vasiliy@mail.ru)
(Filipp Galyano, SPIIRAS, 14 Liniya, 39, St. Petersburg, Russia, galyano@oogis.ru)

1 ABSTRACT

This paper is the summarizing research activities done during recent years by St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). The object of research is a terrorist’s treat as a common problem and research subject is a human aspect as a key factor of a terrorist phenomena. Research activities have been focused mainly on a theoretical investigation. Proposed research can be interpreted as a special kind of a theory that includes some principles and methods for study such complex subject domain.

Developed theoretical aspects will be a basis for investigation practical solutions in a very wide sense: from different monitoring systems to counter-terrorist’s operations. Also some statistical results are presented in the paper.

Acknowledgement: Authors extend their appreciation to ONRG for support project CRDF: RUE2 – 1649 – ST – 06

2 INTRODUCTION

Start point of the research was to discover most important limitations or borders of the theory. This problem should be done in accordance with practical realization of the theory. According to those two most important processes have been selected as follows: global monitoring of individuals and fast analysis of each individual “profile”. Let us to examine them in short.

a. Global monitoring of individuals within a risk zone and/or upon a query (cellular and/or other communications, electronic mail, transportation record (airplanes, trains, etc) and transportation and/or housing geography, etc.

Monitoring concept is often discussed in specialized literature, many other publications and mass media. In this paper we will give consideration to monitoring systems related to human environment in the interests of supporting some parameters of potenthial terrorist’s profile (TP) and terrorist’s track (TT). Supporting systems of different transportation, communication and other means can be assigned to such systems.

Specific but not complete characteristics of alike schemes are:

- the need to use different measuring elements for TP&TT parameters and hasardous matherial and devices (HM&D) parameters with given discontinuity;
- the need to use diverse information of different nature from different sources, as a rule, of great dimension;
- the need to make decisions on a real time basis.

For the above systems the concepts of information (data) harmonization, integration and fusion have certain and determined sense. In this regard the determining value is information fusion aimed at:

- reduction of data dimension (volume reduction);
- increase of data accuracy and reliability (reduction of uncertainty);
- enhancement of data stability (errors correction).

Challenges of the contemporary world such as nets of terrorist’s organizations, refugees and immigrant’s flows, drug and HM&D traffic – these entire phenomena have a strong spatial constituent. Their adequate study requires spatial modeling and simulation. Purely mathematical simulation of spatial processes not provided with visual representation, that may be called “blind”, essentially decreases effectiveness of experts in discovering the patterns regulating these phenomena. For visual representation of spatial simulation in the earth-scale Geo-information systems (GIS) are suited. GIS rapidly evolve from simple browsers of electronic digital maps into complex software applications capable of visual simulation of the real world spatial processes. As level of the GIS applications is rising and their inner logic becomes more complex the use of conventional programming languages in their development turns into a restraining factor. They urgently need means of intelligent support, such as rule-based systems, ontologies, multiagent systems and other AI techniques, especially in the web applications.

b. Fast analysis of each individual "profile" by specialized software technologies (immune computing, pattern recognition, cluster analysis and other).

The available sources analysis reveals that situation awareness (SA) for various subject domains is described realized through different approaches and techniques. SA role and importance were conceptually by the authors of JDL information integration model. The proposal is to implement artificial intelligence techniques, and in this sense the above approach can possibly suggest a universal solution both conceptually and technologically. The approach gives a terrorists' situation (TS) definition, its formal representation and specifies TSA for the above problems. Then the approach gives an analysis of currently known techniques and SA algorithms. It is proposed to consider the SA process and process of the TSA as identical ones. At that use of one or several statistical and/or mathematical algorithms cannot solve the problem on the whole. This is why an approach incorporating expert systems technology is proposed as a core in SA complex model. The above described idea can be realized computationally through the CLIPS shell (Open Source).

As a core for recognition algorithm an immune computing (IC) approach is selected.

IC proposes the following new approaches to AI problems as a new type of calculations:

- pattern recognition and data analysis based on molecular recognition principles;
- language representation and tasks solving based on analogues between words and bio molecules;
- natural and technical systems modeling based on bio molecules interactions.

At the current stage the following parts of the theoretical approach are obtained.

1. Formal description of the subject area, analysis of the problem-related available printed matter: information sources of the governmental, public, and private sectors; study of the past terrorists' actions, forming "profiles" of potential criminals and their activities' "tracks".

2. Determining the potential of the currently existing and developing information sources (cellular communications, Internet, various forms of registration, booking, using credit cards, etc.) for "tracks" detection.

3. Development of mathematic methods to identify terrorist situations, classification of potentially vulnerable locations, "profiles" and "tracks" identification.

3 FORMAL DESCRIPTION OF THE SUBJECT AREA

3.1 Ontology

Initially "ontology" term denoted a philosophy sub-discipline dealing with fundamentals of being [1]. An ontology is a formal explicit description of concepts in a domain of discourse (classes), properties of each concept describing various features and attributes of the concept (slots), and constraints on slots (facets).[1]

[1] According to [2] matter used from now forth.

№	Class name	Subclass name	Slot name
1	Object under Monitoring	-	Name
			Scale
			Shape
			Object vulnerability out of monitoring agency
2	Subject under Monitoring	Terrorist profile	-age -birth place -nationality -religion -party affiliation -participation in social organizations -convictions -being searched by federal or international police,etc.
		Terrorist track	- location - arrival time - departure time - subscriber’s telephone number - call time - call duration credit card purchase designation - purchase time,etc.
			Individual aggregative indices
	…	…	…
			…

Table 1: Basic Ontology Classes and their characteristics

3.2 Terrorist’s profile (TP) and terrorist’s track (TT) identification (description)

Preliminary verbal definition of the terms “profile” and “track” of individual (“individual profile” and “individual track”) is:

(1) “individual profile” – a complex of biography parameters fixing important facts of individual history (place and date of birth, education, religion, religion changes, ethnic group and nationality, etc.);

(2) “individual track” – a complex of parameters characterizing individual behavior with a linkage of these events to time and place of realization (time and place of departure to a certain destination, time and subscriber of the last mobile phone calls, etc.).

Therefore, though differing by its content and means of gathering, individual profile and track data conform to the same mathematical form – a form of a vector x=(x1,…,xn) of parameters x1,…,xn, where each parameter is measured on the certain scale (nominal, order, numeric, etc.). Consequently, using the same mathematical methods for processing both individual profile and individual track data is reasonable.

This relevant data condition could be stated as follows: every parameter x1,…,xn, combined into profile and/or track is necessary, and all of them taken together are sufficient for the definition of implication in terrorist activity.

If the mentioned condition of the parameters x1,…,xn relevancy is met, then one could consider an individual terrorist profile and individual terrorist track, or in other words, the person's terrorist profile and the person's terrorist track.

The stated similarity of the mathematical form of individual terrorist profile and terrorist track representation as a vector of values of a certain characteristics set allows combining the methods of individual terrorist's "track" identification and individual terrorist's "profile" identification.

3.3 Terrorist situations (TS) formalization

A *tactic situation* is regarded as a combination of some parameters, clearly or by implication defining the explored system state at a given moment of time. A *terrorist situation* is regarded as a tactic situation in the system of defense against terrorist threat (threat of HM&D using).

Initial basic TS could be identified in different ways:

- using experts;
- using some theoretical footing;
- choosing some basis for classification.

It is possible to assume that TS could be associated with some operation which, in turn, has an analytical description and, of course, solution.

Thus, TS identification could be performed using the following chain of activities.

TS class identification.

As a rule, TS classes are identified on the basis of main problems solved by a specific system (in particular, counter- HM&D system). Often TS class is defined by the information from the higher counter- HM&D forces control, coming in the form of a task formulated in an order.

Search of a specific tactic situation variant in the identified class

Current modeling step results in:

(1) specific TS hypothesis definition (Terrorist situation Hypothesis – TSH) – in the range of TS class defined (identified) at the previous step;

(2) special calculations (determination or specification of direction and speed of dislocation, object location forecasting, necessary calculations for changing the operating modes of searching facilities etc) and development of recommendations for counter-terrorism forces control in arising TS.

The next and, in this particular case, the final step is to prepare offers of TS resolution for counter-terrorism forces actions control and to prepare necessary control actions to realize them.

Attainment of the goals put in the project assumes accessing a great number of heterogeneous (of different data presentation model) data sources. It is necessary to substantiate and chose universal data carrier ensuring communication with any information source.

Data model should be, in the first place, uniform for all data sources; in the second place, its medium should provide convenience of exchange through any automated communication channels; and, finally, its medium should allow easy development of applications to use it.

Extensible markup language (XML) is the most preferable one for solving the formalized data presentation problem (including data forming TP and TT, and also learning and recognizable information about TS and TSH) and its exchange between separate system components.

3.4 Information source formalization

Search for information sources is a rather complicated problem. Obviously there will be heterogeneous sources using different data presentation formats. However, it is necessary to duly classify possible information sources and describe them formally, as reliability, efficiency and trustworthiness of the received information determines successful solution of a problem of timely detection of the terrorist threat.

The following main groups of information sources for developing knowledge and data base, describing individual terrorist profile and individual terrorist track can be identified:

- security services reports (corporate, national (CIA, FBI, etc.) and also world counter-terror organization);
- technical supervision, control or reconnaissance data;
- information about migration flows (national migration services data, passport and visa information, tickets information, etc.);
- information from agents penetrated into terrorist groups;
- Criminalist study of individual appearance.

Terrorists themselves are often information sources about terrorist groups.

It is important to take into consideration the influence of an information source and a communication channel on its reliability and accuracy and their interference, when analyzing each specific message. The information source influence depends on its characteristics. The information source properties are characterized using the following quality indices:

- scope or service area;
- situation representation completeness;
- accuracy and detail (resolution);
- situation representation reliability.

All information received from the sources *(Information sources - SI)* could be sorted into two types: statistic *(Statistical Data – SD)* and operational *(Operational Data – OD)* data:

$$SI = \{SD; OD\}$$

Statistical data are usually received from public services and counter-terrorist activity research and development organizations. This data define potential individual terrorist profiles and also could be used to define individual terrorist track:

$$SD = \{BD, BiD, MW, TC, SC\},$$

where BD - biographic data; BiD - biometric data; MW - data of weapon used in terrorist acts; TC - terrorist acts realization time constraints data; SC - terrorist acts realization spatial constraints data. Each of these statistic data types is set by a vector characterizing current data type. Mathematical form of a vector describing individual terrorist profile features is the same for all of them:

$$x = (x_1^P, ..., x_m^P),$$

where m - number of features describing individual terrorist profile; P - index of current feature value type.

Operational data are the dynamically changing data that could be received from different sources:

$$OD = \{MF, IS, RI, CC\},$$

where MF - data received during cellular communication systems analysis; IS - data from Internet; RI - data from registration forms and CC - data about credit cards. This data is also set by features vector $\vec{x}$.

3.5 Monitoring System description

Structure elements of global (local) terrorist acts monitoring system functioning in a region (at an object) are:

-subsystem of information sources about terrorist threat;

-subsystem of identification of terrorist act;

-subsystem of control;

-geo-information system.

Subsystem of information sources provides gaining data characterizing the terrorist threat level:

-data characterizing individual terrorist profile (*TP*);

-data characterizing individual terrorist track (*TT*);

-entourage data (*TI*) which influence on the terrorist threat level evaluation.

Data received by the subsystem of information sources enter the subsystem of identification quantized by time in the form of:

-initial characteristics of individual terrorist profile vector:

$$x^{OP} = (x_1^{OP}, ..., x_k^{OP});$$

-initial characteristics of individual terrorist track vector:

$$x^{OT} = (x_1^{OT}, ..., x_m^{OT});$$

-initial characteristics of the terrorist threat level vector:

$x^{OI} = (x_1^{OI},...,x_n^{OI})$.

Geo-information system supports location representation of the information sources, their functioning results and the functioning results of the subsystem of identification.

Following indices characterize terrorist activity degree:

-combined index of individual implication in the terrorist activity $\bar{Q}_i^T$ (i=1…N), characterizing the degree value of possible implication of each individual from N individuals under control into the terrorist activity;

-terrorist threat index J_{TS} characterizing numerically the terrorist act probability in a region (at an object);

-terrorist threat class K_{TS} in a region (at an object).

The subsystem of identification structure needs to be developed in order to provide:

-evaluation of the combined index of individual implication in the terrorist activity $\bar{Q}_i^T$;

-evaluation of the terrorist threat index J_{TS}

-determination of the terrorist threat class K_{TS} in a region (at an object).

Subsystem of the identification information model briefly reflects the interference of input data, output data and data formed inside the subsystem. A simplified scheme of the information model is depicted in Fig.1.

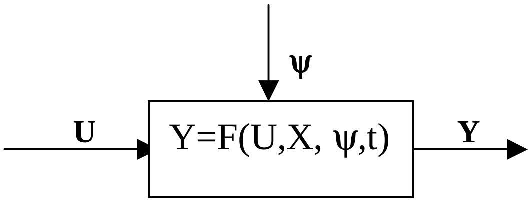

Fig. 1: Subsystem of the identification information model

U – input data vector, i.e. data received from subsystem of information sources about terrorist threat:

y - external impact on the input data, usually accidental;

X – inner parameters of the subsystem or parameters of its inner processes;

Y – vector of output parameters. Values of the vector Y components could be considered as a reaction of the subsystem to the input data U and distortions y;

F- conversion operator correlating a given set of arguments U, X and y in a discrete time moment t, linked to a moment of new data incoming, with the vector of output parameters Y. Operator F should provide such input data conversion that provides evaluation of the system output parameters.

Analysis of the subsystem input data shows their peculiarities:

- large data volume;
- vagueness of some data values;
- data heterogeneity;
- dependency of data volume and type of the monitoring system scale and knowledge received during system functioning.

It is important that there is no unambiguous correspondence between input and output data of the system. This prevents one from using strict data processing algorithms as a conversion operator for the identification problems solution.

The stated peculiarities determine the necessity of expert systems usage as a development basis for the subsystem of the terrorist act identification.

Initial data entering the subsystem of identification could be measured by following scale types:

1. Logical scale.

2. Nominal (categorical, qualitative) scale.

3. Linear order scale.

4. Numerical scale.

There are some more scale types, for example:

- relations scale which values are invariant to positive linear conversions, shift and stretching;
- partial order scale with partial order defined in a usual way on the value set of the scale, etc. However, they are negligible in the frames of application problem concerned and therefore are not discussed further.

The presence of identification problem's initial data measured on polytypic scales does not allow using same conversions for all data. Specific processing methods are developed for each data type to solve tasks of detecting regularities.

Subsystem input data are:

-features data of individual terrorist profile (features vector *TP*);

$$x^{OP} = (x_1^{OP},...,x_k^{OP})$$

-features data of individual terrorist track (features vector *TT*);

$$x^{OT} = (x_1^{OT},...,x_m^{OT})$$

-features data of terrorist threat (features vector *TI*);

$$x^{OI} = (x_1^{OI},...,x_n^{OI})$$

Features data enter scale transformation units. Scale transformation units converse input data measured on the scales providing their further processing by other subsystem units:

$$x^{'OP} = (x_1^{'OP},...,x_k^{'OP})$$

$$x^{'OT} = (x_1^{'OT},...,x_m^{'OT})$$

$$x^{'OI} = (x_1^{'OI},...,x_n^{'OI})$$

Feature data of an individual under identification enter the evaluation algorithm of the aggregated index of individual implication in the terrorist activity $\bar{Q}_i^T$. Algorithm functioning results are the development of the generalized individual characteristic in a form of the aggregated index of individual implication in the terrorist activity $\bar{Q}_i^T$, and also calculation of accuracy S_i and reliability P_i of aggregated index evaluation.

Aggregated index of individual implication in the terrorist activity $\bar{Q}_i^T$ is the input parameter for the algorithm of the terrorist threat classification. Besides that, parameters of vector of the features of terrorist threat level $x^{'OI} = (x_1^{'OI},...,x_m^{'OI})$ are input parameters for the algorithm of the terrorist threat classification.

Algorithm of the terrorist act classification provides composition of the terrorist threat index J_{TS} and the terrorist threat class in a region (at an object) K_{TS}, which are output parameters of the subsystem of identification together with the generalized individual characteristic.

Subsystem initial and consequent learning is provided by learning samples formation units. Units offer the user the possibility to correct weight coefficients $w_1,...,w_p$ used in algorithm of degree evaluation of the individual implication in terrorist activity and learning sample $x_{o_1}^{OI},...,x_{or}^{OI}$ in the algorithm of the terrorist act classification.

Database stores learning samples, input data vectors and results of individual and terrorist situation identification.

3.6 Environment formalization

An environment is regarded here as external factors having influence on the monitoring system (MS) functioning. External factors influence is regarded here mainly as constraints they impose on the MS functioning.

Formal representation of the monitoring system environment O could be formalized as:

$$O = \{PDZ, SZ, \vec{X}'', \vec{Y}''\}.$$

where PDZ – potentially dangerous zone, zone with individuals flow beyond the control system coverage;

$\vec{X}''$ – input flow of individuals which implication in the terrorist activity is analyzed;

$\vec{Y}^1$ - output flow of individuals of potential terrorist threat;

$\vec{Y}^2$ - output flow of individuals of no potential terrorist threat;

SZ - security zone entering by an individual after MS check in case he's decided to be of no potential terrorist threat.

Potentially dangerous zone in the monitoring system environment is formalized as:

$$PDZ = \{a, b, S, \gamma\},$$

where a - potentially dangerous zone width;

b - potentially dangerous zone length;

S - potentially dangerous zone square;

γ - set of zone additional characteristics.

As a and b parameters could be used either region geometry or geographic coordinates of the object under monitoring λ (latitude) and φ (longitude).

Security zone is

$$SZ = \{a, b, S, \gamma, SL\},$$

where SL - security zone limit.

$$SL = \{BT, DA, PG\},$$

where BT - security zone limit shape, set by vector of values. Confines of the object under monitoring with installed monitoring system are used as the current parameter. For example, confines of an airport, a railway station or marine passenger terminal, metropolitan, etc.;

DA - security zone limit type. MS type depends on this parameter (local or global MS);

PG - index of possibility of intrusion into the security zone escaping monitoring system.

3.7 Requirements for GIS

visualization of analysis results of terrorist threat features, detected from various information sources. Thus, GIS could be classified as the specialized information analysis geo-information system solving special-purpose problems.

To define requirements for the GIS of the stated class one should define requirements for its three main components:

- basic tools;
- base of spatial data;
- applications' specifications.

The basic tools should ensure realization of the following main operations:

- export-import of cartographic and thematic data;
- basic cartographic conversions;
- data storing and manipulating;
- measuring operations;

- spatial analysis;

Spatial data (geospatial data, geographic data, and geodata) are regarded as data of geographic objects which are the formalized digital models of tangible or ideal (abstract) objects of real or virtual world.

The following requirements are specified for the spatial data bases:

- accessibility to all system users;
- operative synchronized actualization;
- coverage completeness of the whole territory;
- standards compliance;
- creation priorities and consistency;
- completeness – necessary sufficiency and non-redundancy of data (absence of blank spots. A base of spatial data (BSD) doubling is acceptable in case of measuring taking place at different times and with different accuracy made for different parts of a base of spatial objects (BSO));
- logical coordination – upholding of constraints on attributes and topological geometry properties of objects and their sets;
- positioning accuracy – closeness to the real results of an object positioning in space;
- time accuracy – closeness of a fixed time of an object existence to real time;
- attributive accuracy (thematic accuracy) – closeness of actual attributes values to the real values.

The GIS should provide visualization of dislocation routes (tracks) of the objects of interest with the following characteristics:

- object identifier;
- initial and final date and time of the period of interest;
- coordinates of north-west and south-east corners of overlapping trapezium of the area of interest.

4 STUDY OF POTENTIAL INFORMATION SOURCES

4.1 Statistic data

One could set the following requirements on the sampling used by ISTS corresponding to the preliminary analysis stages:

1. Minimizing the sampling impurity, i.e. removing non-reliable data out of the sampling.
2. Checking stochastic independence of the sampling elements.
3. Evaluation (confidence intervals design) of the distribution parameters.
4. Preliminary evaluation of the chosen parameters significance.

Problems solved at the current stage are from mathematical statistics problems range.

4.2 Operational data

Operational data in the context of this report constitute dynamically changing information that could be used for TP and TT detection, but it defines better “row, sequence of events”, i.e. “track” of individual. “Individual track” is a complex of parameters operatively characterizing individual behavior with a linkage of these events to time and place of realization (time and place of departure to the given point, time and addressee of last mobile phone calls, etc.). This information could be received from sources of different types: cellular communication systems; Internet; registration services of airports, railway stations, hotels, etc.; information services of banks. Analysis of information received from the above-stated sources allows forming a set of features (feature vectors):

$$x_{OD}{}^{OP} = (x_{MF_1}{}^{OP},\ldots,x_{MF_{k1}}{}^{OP}; x_{IS_1}{}^{OP},\ldots,x_{IS_{k2}}{}^{OP}; x_{RI_1}{}^{OP},\ldots,x_{RI_{k3}}{}^{OP}; x_{CC_1}{}^{OP},\ldots,x_{CC_{k4}}{}^{OP})$$

$$x_{OD}{}^{OT} = (x_{MF_1}{}^{OT},\ldots,x_{MF_{m1}}{}^{OT}; x_{IS_1}{}^{OT},\ldots,x_{IS_{m2}}{}^{OT}; x_{RI_1}{}^{OT},\ldots,x_{RI_{m3}}{}^{OP}; x_{CC_1}{}^{OT},\ldots,x_{CC_{m4}}{}^{OT})$$

and a set of data (OD) characterizing the features of individual terrorist profile and individual terrorist track.

$$OD = \{MF, IS, RI, CC\},$$

where MF - subset of data received from cellular communication systems analysis; IS - subset of data from registration forms and CC – subset of data about credit cards use received from information services of banks.

The combined use of the operational data from various sources, set by vector

$$\vec{x} = \{X_1^{MF}, \ldots, X_{n1}^{MF}, X_1^{IS}, \ldots, X_{n2}^{IS}, X_1^{RI}, \ldots, X_{n3}^{RI}, X_1^{CC}, \ldots, X_{n4}^{CC}\}$$

allows designing *TP* and *TT* effectively to detect terrorist threat. At the same time, the impact on the citizens' civil rights and private life is minimal, as there are neither wide-range phone-tapping (which on its own is not too effective due to using "cover stories"), nor violation of secrecy of correspondence (from the point of disclosure of the message meaning content).

5 DEVELOPMENT OF MATHEMATICS METHODS

5.1 Terrorist's "Profile" and Terrorist's "Track" Identification

5.1.1 Weight-coefficients estimation based on mixed NNN-information about weight-coefficients and aggregated indices

In the method's framework it is supposed that all possible alternatives (synonyms: variants, solutions, courses of action, objects, etc.) of a decision are fixed by a decision-maker (DM). Also, it is assumed that some attributes (synonyms: characteristics, features, properties, parameters, etc.) are selected by the DM for the alternatives description. Thus, the alternatives of the decision-making may be named multi-attribute alternatives.

It is supposed that each of the constructed single preference criterion is necessary, and the whole set of them is sufficient for a numerical estimation of any alternative's preference. In other words, it is supposed that a numerical estimation of an entire alternative's preference is a numerical function of the set of all single preference criteria. Such numerical function of all single criteria of preference is named aggregated preference index, and is treated as an aggregated criterion of the alternatives' preference. Value of an aggregated preference index for a given alternative is its preference estimation which takes into account the whole set of single estimations of the alternative's preference.

Any process of alternatives' preference estimation with help of an aggregated preference index may be put into terminological shape of correspondent objects quality estimation by use of an aggregated quality index.

5.2 Terrorists' Situations Identification (TSI)

5.2.1 Description of pattern recognition basic algorithm using immunocomputing method.

In conventional tasks of pattern recognition and particularly in tasks of terrorist situation identification (TSI) the source data are multidimensional and allow representation in form of arrays (vectors) of real numbers and/or integers. Physical interpretation of IC method lies in a projection of arbitrary multidimensional data onto space of FIN, whose dimension is substantially less, and search of a problem solution, e.g., recognition and interpolation, is based upon points' proximity in this space of FIN (Fig. 2).

At that each coordinate in space of FIN determines a value of a so called binding energy between a corresponding basic vector (so called antibody-sample) and an arbitrary input vector.

In molecular biology the binding energy is a basic biophysical measure of biomolecular interaction (bonding). In molecular immunology proteins are the basic biomolecules of an immune system-antibodies, as well as any foreign protein – antigen. At that the essence of an immunologic test aimed at identifying (recognizing) an unknown antigen is reduced to a determination of a degree of its bonding with a known antibodies (samples) set.

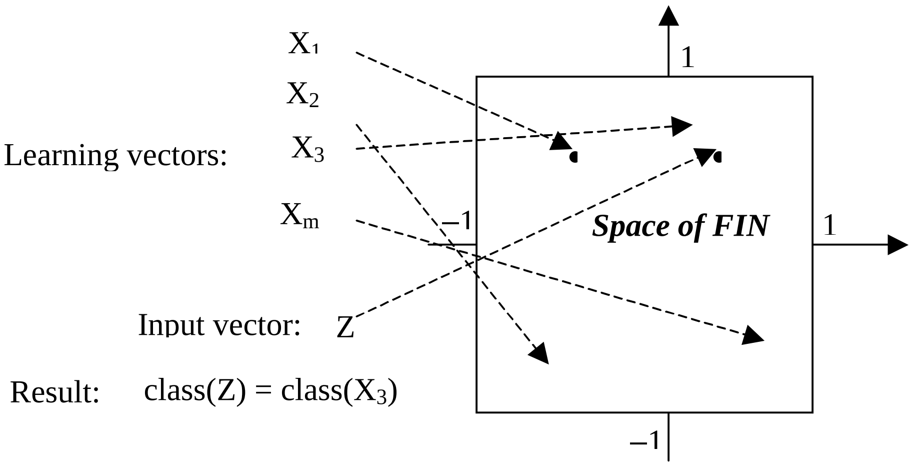

Fig. 2: Example of pattern recognition in 2D space of FIN

By definition the basic vectors in space of FIN are right singular vectors of the learning matrix composed of learning vectors; the above is related to a particular case of a so called *formal protein* used in IC for solving pattern recognition problems. A projection of arbitrary data onto space of FIN is one of basic peculiar properties of IC- pattern recognition algorithm. Such a transform possesses the following advantages; it

- has a rigorous mathematic substantiation in terms of matrices singular decomposition;
- significantly reduces data dimension (up to one- two- or three dimensional space of FIN);
- allows pictorially represent and visualize any situation as a point in one- two- or three-dimensional space.

5.2.2 Mathematical formalization of the problem

The pattern recognition problem could be formalized as follows:

Given:

- – Number of classes c;
- – set of m learning patterns: $X_1, \dots, X_m$;
- – class of any learning pattern: $f(X_1) = c_1, \dots, f(X_m) = c_m$;
- – arbitrary n-dimensional vector P.

Find:

Class of vector $P: f(P) = ?$

Let us consider capabilities of various intelligent information technologies to solve TS identification problem.

5.2.3 Comparative analysis of capabilities of intelligent information technologies when aimed at TS identification problem solving

Analysis of the most promising methods of TS identification problem solving allows to arrive to the following conclusion. ANN, GA and IC performance is regarded in terms of their learning time, rather than identification period. The point is that real practice applications (especially in real time) depend a lot on the system “flexibility”, i.e. its capability to “re-learn” under changing circumstances. In addition, regardless of the chosen ANN settings, its performance is determined by two main factors – disadvantages, widely known by ANN specialists:

1) exponent calculation necessity in activation functions for each neuron; and

2) slow convergence of BPE (back propagation of error) gradient method when weights are close to the local minimum of error function. Moreover, existence of multiple adjustable, not to mention obvious redundancy of interneural relations “each-to-all”, worsens essentially the ANN learning process. Thereby GA has even

more adjustable parameters, not to mention rather unnatural problem of data coding-decoding by the bit rows. As a result, ANN and GA using in many real-time applications proves to be simply impossible.

Based on the above it should be noted that IC is principally free of the considered ANN and GA disadvantages. Therefore, IC potential capabilities allow to reach a level of calculation stability, flexibility and performance, that is beyond of conventional and neural-computers capabilities.

Further, the basic algorithm of a terrorist situation identification on the basis of immunocomputing method is discussed.

5.2.4 Description of the developed algorithm of terrorist situation identification (TSI)

Algorithm of terrorist situation identification (TSI) is developed based on the basic IC-algorithm of pattern recognition.

```
//Standard interface module
 forming subject domain model of Situation
{
determine Situation as a set of parameters;
determine number coding of parameters; //parameters’ vector
Form learning matrix;
}
// Module "Situation"
learning //data mapping into FIN space
{
to receive a learning matrix;
calculate the SVD of the learning matrix;
store first three singular numbers and matching vectors;
}
recognition //data classification in space FIN
{
receive parameters’ vector of Situation; //pattern
to project the pattern into a point FIN [w1,w2,w3];
to find n closest points FIN; //n is given in interface module
to determine codes TS for these points; //classes TS
to determine probabilities TS for each point;
to forward results into interface module;
}
```

Adaptability of Terrorist Threat Identification Procedure in Various Functioning Modes of Monitoring System.

The IC method allows to process big volumes of the objective information. At the same time, precision of situation recognition when using IC method (as, however, it happens when using any other pattern recognition method) directly depends on the learning sample volume and quality. AIM is a kind of the IC method antipode. There can exist no learning sample at all (i.e. any data about characteristics of the analyzed tactic situation may not be available), but this fact will not prevent experts from adjusting AIM in accordance with their level of competence and obtain the searched index.

Based on the above, it is easy to make a conclusion that the real success in the TS analysis can be achieved through reasonable combination of the described methods. In case there are enough of learning data, or in other words when a big amount of well studied and described TS, similar to the analyzed one, is available the IC method is used for the situation assessment. If the learning sample size does not allow performing the correct TS recognition by the IC method (it happens when the analyzed situation is unique, not similar to any previously considered or rarely happens), the tactic situation analysis is performed by AIM based on expert knowledge. If volume of the information characterizing the tactic situation is sufficient for a correct implementation of both methods, AIM and IC methods are used in parallel in the unified system of the situation assessment aimed at mutual control and improvement of the recognition reliability.

6 STATISTICAL ANALYSIS

The reference data [12] were used as the facts reflecting biographical data of the terrorists, at that, ninety three biographies of the kind were subjected to the analysis, including biographies of terrorists from Great Britain, Germany, Greece, Egypt, India, Israel, Yemen, Lebanon, Palestine, Poland, Russia, Saudi Arabia, Syria, the USA, France, and time period of 120 years was covered. The terrorists' data were analysed, like age, education, social background, marital status, existence of personal tragedies, temper, goals they pursued. The times reflecting the beginning and the end of their implication in the terrorist activities were considered the age parameters. The educational levels were divided into elementary or incomplete secondary, secondary or incomplete higher and higher education. By their social background peasants, workers, intellectuals, business people (industrialists) were singled out. Under the existence of personal tragedies were understood the problems: with the law in the family where the terrorist was raised; of being deprived of parents' care; of losing the loved ones; of expulsion from a university, etc. The following personal character features as well balanced or hot tempered were separated.

Through the biographies' analysis and based on the above listed features the histograms were built, that further were processed to obtain the quantitative characteristics of terrorists' profiles, at that, the known methods of mathematical statistics were used.

The following time distributions reflecting the beginning and the end of the terrorists' implication in the related activities have been received (Fig. 3, Fig. 4).

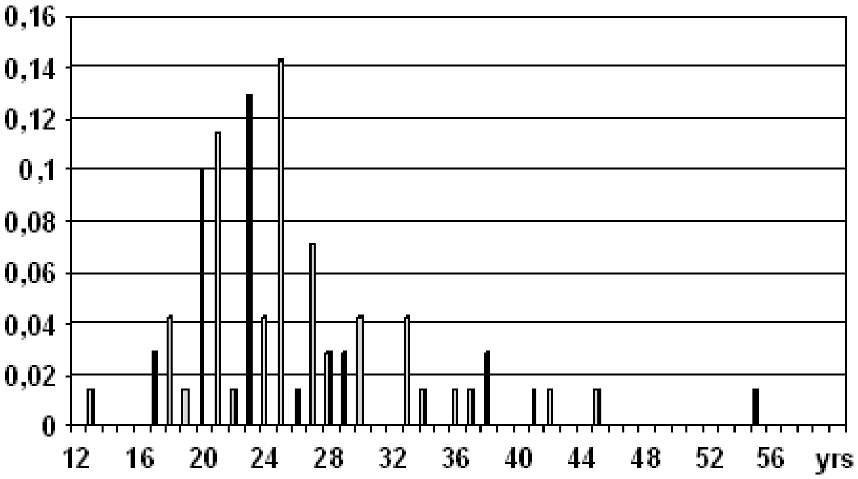

Fig. 3: Distribution of the beginning and the end of the terrorists' activities

In accordance with the first histogram (Fig. 1) the mathematical expectation of time (age) when the active terrorists' activity commenced was determined as $m_{x1} = 25.84 \pm 1.7$ years for the evaluated individuals.

Mean-root-square deviation $\sigma_{x,l}$ of this time from its mathematical expectation is equal to 7.2465 years.

In accordance with the second histogram (Fig. 2) the mathematical expectation m_{x2} of time distribution for the end of the terrorist activity equals 35.23 ± 2.84 years, and the respective mean-root-square deviation is σx2 = 12.18. The confidence level of the evaluated mathematical expectations falling into the above limits is 0.95 both in the first and second cases.

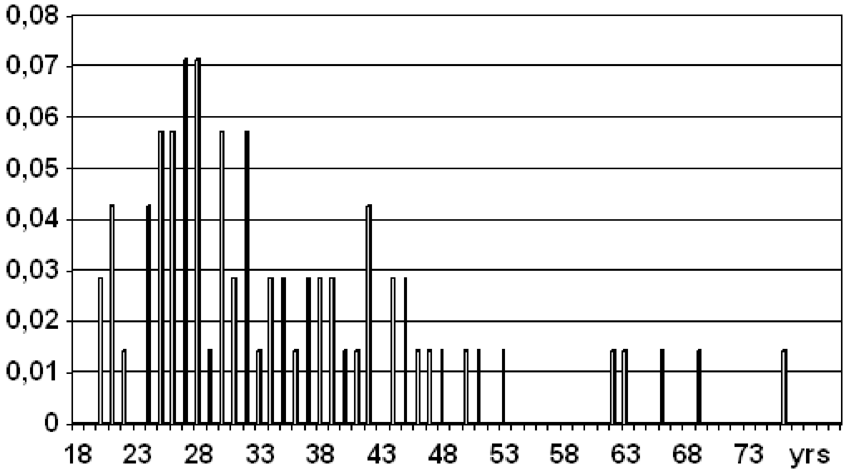

Fig. 4: Time distribution of the end of the terrorists' activity.

The analysis of Fig. 3, Fig. 4 shows that mean duration of the terrorists' activity makes about 10 years. At that, the received distributions can be

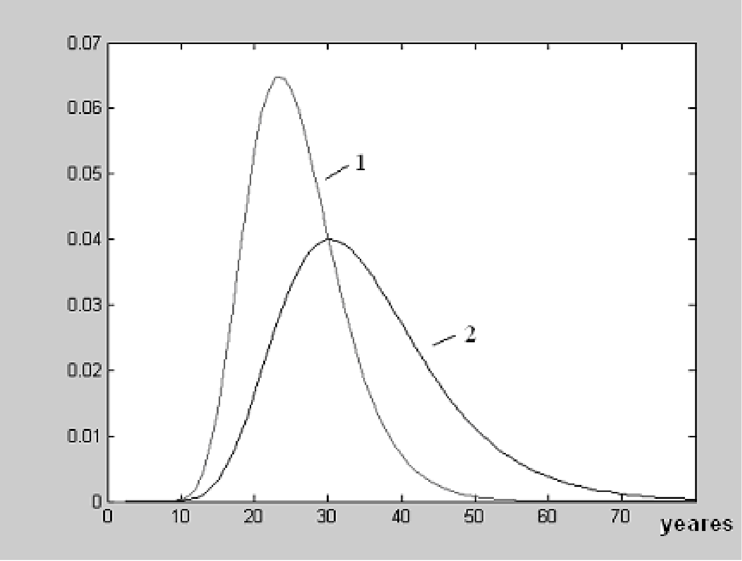

Fig. 5: Distribution densities for the aleatory variables obeying the lognormal law with the parameters: 1 — $m_1 = 3.2185$, $\sigma_1 = 0.2538$; 2 — $m_2 = 3.5108$, $\sigma_2 = 0.3130$.

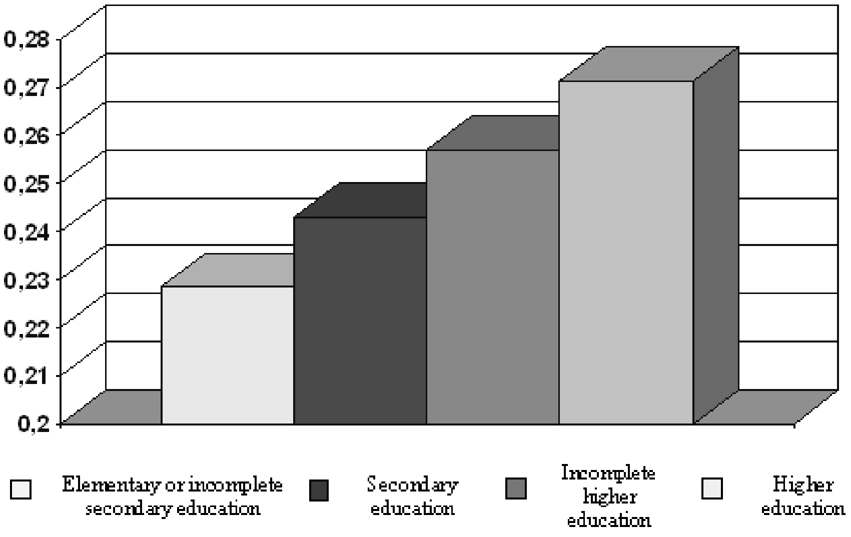

Fig. 6: Relative Terrorist Distribution by Education Levels

Fig. 6 relative distribution of terrorists by educational levels:

- elementary or incomplete secondary;
- secondary;
- incomplete higher;
- higher.

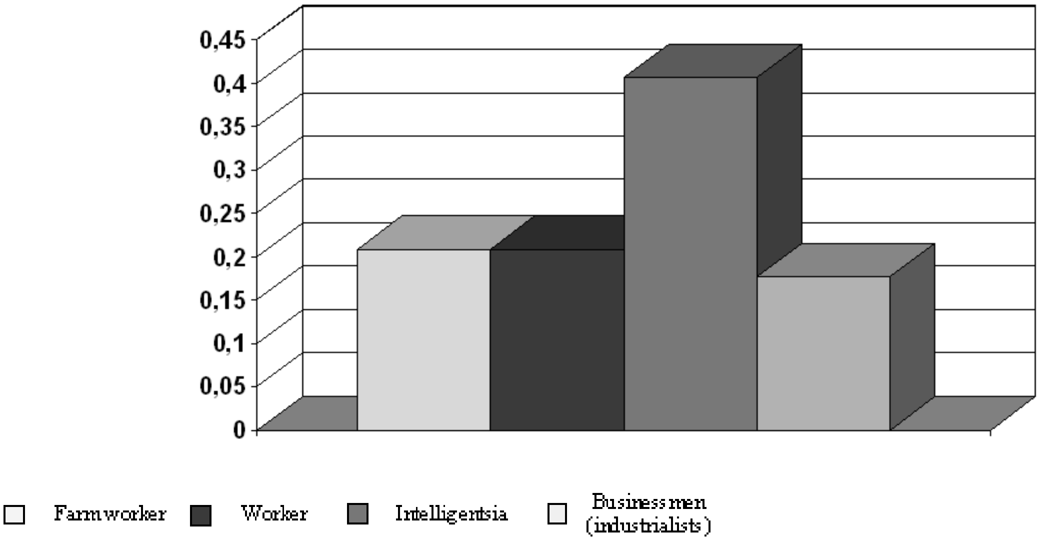

Fig. 7: Relative distribution of terrorists by social background

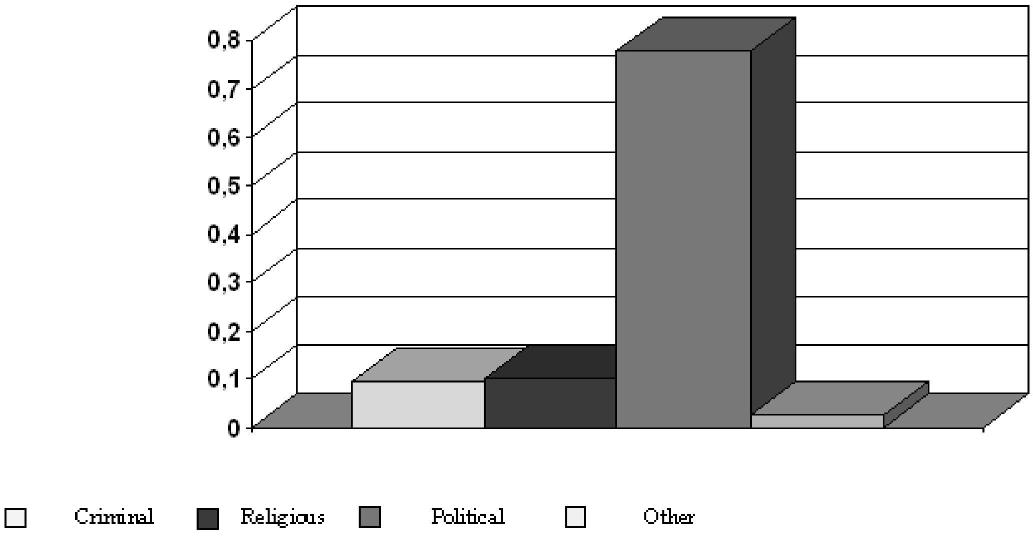

Fig. 8: Relative distribution of terrorists by pursued goals

7 CONCLUSIONS

Main parts of our theoretical research have been presented in the paper as follows:

- Formal description of the subject area, analysis of the problem-related available printed matter: information sources of the governmental, public, and private sectors; study of the past terrorists' actions, forming "profiles" of potential criminals and their activities' "tracks".
- Determining the potential of the currently existing and developing information sources (cellular communications, Internet, various forms of registration, booking, using credit cards, etc.) for "tracks" detection.
- Development of mathematic methods to identify terrorist situations, classification of potentially vulnerable locations, "profiles" and "tracks" identification.

The received statistical terrorist profile allows to analyze/evaluate certain individuals from different states, or confessions for their belonging to terrorist risk groups. The basis of this profile is made by seven common indices and individuals conflict ness indices depending on their countries and religion. It establishes that time of the beginning and the end of the terrorists' activity complies with lognormal law.

By means of the created profile certain individuals implicated in or inclined to terrorism can be recognized as well as the entire groups.

The received results can be used to create up-to-date or modify existing systems for increase the efficiency of providing the antiterrorist security measures in airports, at railway stations, at the documents' check points, and other counter-terrorist's systems.

8 REFERENCES

The Global Terrorism Database (2007). National Consortium for the Study of Terrorism and Responses to Terrorism START: A Center of Excellence of the U.S. Department of Homeland Security University of Maryland, College Park. http://www.start.umd.edu

Report on Terrorism (2008) USA. National Counterterrorism Center

Terrorism Review (2006) Israel Ministry of Foreign Affairs

Susan B, Glasser (2005) Global Terrorism Statistics Released. Clearinghouse Data Show Sharp Rise. Washington Post Staff Writer

Terrorism Statistics (2008). http://www.nationmaster.com.

Hetherington Cheryl L (2005) Modeling transnational terrorists' center of gravity: an elements of influence approach. Thesis. USAF: Air Force Institute of Technology. Graduate School of Engineering and Management

Kendall Shanece L (2008) A unified general framework of insurgency using a living systems approach. Thesis. Naval Postgraduate school. Monterey, California

Popovich V, Hovanov N, Hovanov K, Schrenk M, Prokaev A, Smirnova A (2008) Situation Assessment in Everyday Life. In: Manfred Schrenk (ed) 13th International Conference on Urban Planning and Regional Development in the Information Society, Vienna

Popovich V, Prokaev A, Schrenk M, Galiano F, Voronin M, Smirnova A (2008) Monitoring of Terrorist's Treats: a Theoretical Approach. In: Schrenk M (ed) 13th International Conference on Urban Planning and Regional Development in the Information Society, Vienna

Akimov VA (2004) Assessment and prediction of strategically risks Russia: theory and practice (in Russian). J Law and security 1(10)

Terrorist threats of Russia (2007) http://lukoyanov.novoemnenie.ru/articles2/16.html (in Russian)

Zharinov KV (1999) Terrorism and terrorists: History directory (in Russian). Taras AE (ed), Minsk, Harvest

Sokolova LV (2005) Computer software of the analysis, modeling and forecasting and possibility of their use for struggle against terrorism (in Russian). J Analytical bulletin 7 (259). Moscow, Russian Federation

Vasil'ev VI (1969) Distinguishing systems: handbook (in Russian). Kiev, Naukova Dumka

Mathematical methods of recognition of images (2007). In: 13th All–Russia conference: the Collection of reports. Moscow, MAX Press (in Russian)

The directory on the applied statistics (in Russian) (1990) vol 2: Lloyd E, Lederman U, Ajvazjan SA, Tjurin JN (eds). Moscow, Finance and statistics

Tsypkin JZ (1995) The information theory of identification (in Russian). Moscow, Nauka

Christopher J. Matheus, Mieczyslaw M. Kokar, Kenneth Baclawski. A Core Ontology for Situation Awareness. In Proceedings of Sixth International Conference on Information Fusion, pages 545-552, Cairns, Australia, July 2003.

Gabriel Jakobson, Jonn Buford, Lundy Lewis. Situation Management: Basic Concepts and Approaches. In Proceedings of the Third International Workshop: Information Fusion and Geographic Information Systems, St. Petersburg, Russia, May 2007.

OWL Web Ontology Language Overview. W3C Recommendation 10 February 2004. http://www.w3.org/TR/owl-features/.

Tarakanov A., Dasgupta D. A formal model of an artificial immune system. //BioSystems (Int. J. of Biological and Information Processing Sciences), 2000, 55(1-3), pp. 151-158.

Tarakanov A. O. Information security with formal immune networks, Information Assurance in Computer Networks (Gorodetsky V. I., Skormin V. A., and Popyack L. J. eds.), //LNCS 2052, Springer, Berlin, 2001, pp. 115-126.

Tarakanov A. O., Skormin V. A., Sokolova S. P. Immunocomputing: Principles and Applications, N.Y., Springer, 2003.

Dasgupta D., Krishna-Kumar K., Wong D., Berry M. Negative selection algorithm for aircraft fault detection. //Lecture Notes in Computer Science 3239, 2004, pp. 1-13.

Sokolova S. P., Abdullina V. Z., et. al. Artificial Immune System for the gerbil natural plague focus. / Edited by A.O. Tarakanov.– Almaty, 2002, 180p.

Melnikov Y., Tarakanov A. Immunocomputing model of intrusion detection. Lecture Notes in Computer Science 2776, 2003, pp. 453-456.

Tarakanov A., Dasgupta D. An immunochip architecture and its emulation. //NASA/DoD Conf. on Evolvable Hardware (EH-2002). Alexandria, Virginia, July 15-18, 2002, pp. 261-265.