Optimal Infrastructure for Acquiring and Processing of Data related to Anthropic Computer Incidents

Cornel RESTEANU[1], Electra MITAN[1], Marin ANDREICA[2], Gheorghe PACURAR[2]

[1] National Institute for Research and Development in Informatics,
Bucharest 011455, Romania,
resteanu@ici.ro, electra.mitan@gmail.com

[2] Economic Studies Academy,
Bucharest 010374, Romania,
marianadreica@gmail.com, gicu_p@yahoo.com

Abstract: The paper presents a good practice in choosing optimal software - hardware infrastructure couple engaged with acquiring and processing of data related to computer security anthropic incidents. The authors work in the Multi-Attribute Decision Making paradigm, mono-decision maker and mono-state of nature sub-paradigm. An assessment model is built. The model's objects are the software – hardware couples. Concerning the first couple's component, i.e. the software infrastructure, the selection process is done starting with Data Base Management Systems' complex taxonomy. Concerning hardware infrastructure, the choice of elements is based on considered availabilities. The model's attributes are couples' characteristics evaluated by an expert which gives grades according to his / her expertise. Based on this model, the TOPSIS method computes the merit of each couple. The couple with the greatest merit is considered as optimal.

Keywords: Computer Security, Cyber Attacks, Data Base Management Systems, Computer Systems, Clustering, Multi-Attribute Decision Making.

1. Introduction

Anthropic incidents in the field of computer science are caused by cyber weapons [1] or non-cyber weapons. In this paper only cyber weapons are taken into account. The main incidents produced by cyber weapons are:

- Infection with computer associated bacteria (with its varieties: regular bomb, timer bomb, logic bomb etc.), viruses (with the varieties: boot sector, appending, companion, crypto, critical, Trojan horse, binary, multi-partite, link, file jumper, stealth, morphy, runtime, parasitic, polymorphic, resident, spy etc.), worms (with the varieties: computer, network, host etc.);
- Sabotaging the firewalls installed for the protection of web applications;
- Attacks of email systems by spreading word viruses and flooding messages;
- The alteration of the functioning of the search engines by over posting ads or information that the user has requested in previous sessions;
- The violation of the access data of the users` (names, passwords, accounts etc.);
- Not respecting the access rights to some data;
- Decrypting data that are supposed to be secret;
- Destruction of the integrity of the systems' digital content (files / data bases) etc.

There have been created in almost every country several multi-level structured entities with a view to tracking the current status and the evolution of the anthropic cyber incidents produced by informatics means in order to build strategies against this phenomenon. Basic organisms are assigned to one area to work in by collecting data on incidents: how the attack happened, place of occurrence, date of occurrence, the damage and how was solved the problem locally. A summary report shall be submitted to a higher level national or international authority. It systematizes the information received, processes it statistically / graphically and transfers it to the integrating international authorities. The minimal purpose of these bodies is to warn the users in specific areas on imminent hazards as well as to define the tools designed to ensure security. Obviously, these entities must have adequate tools to perform the functions listed above. The infrastructure of these tools, both software and hardware, must be optimal for the system functionality to satisfy the needs of IT efficiency, which in this case are quite high. Therefore, paper's goal is to propose a way to find optimal software – hardware infrastructure for cyber security centres.

2. Security Organisms and the Optimization Modality of their Working Infrastructure

A classic example of an organization based monitoring system, also launching security

policies, is the NCNMER / RoEduNet agency (National Computer Network Management for Education and Research) that „administrates and develops the RoEduNet network and provides data communication services for research and academic institutions at all levels in Romania". An example of an intermediate body is TERENA (Trans-European Research and Education Networking Association) meaning "association for education and research networks in Europe, which provides the framework in which they innovate and develop technologies for changing knowledge, infrastructure and Internet access services used by the research and education communities safely". An example of an international integrator body is the agency ENISA (European Network and Information Security Agency) which is "the association responsible for IT security activities within the European Union, defining and promoting IT security standards in Europe". The most prominent organisms of Information Security are structured by NATO and they have correspondence on national level, for example ORNISS and CERT, due to the integration in Euro-Atlantic structures.

According to the methodology of mathematical modeling, when is needed an optimum over a multitude of homogeneous entities, independent of each other, carrying more observable, measurable or computable characteristics, one proceeds as follows:

- One examines the associated entities and their characteristics, specifying the purpose of the mathematical model;
- Considering, on the entities set, the possible features as generators of equivalence relations, provides a comprehensive taxonomy of it [2, 3];
- Using the set of classifications offered by taxonomy, it associates for each entity a class to which it belongs;
- All entities, from all classes, subject to the optimization but inconvenient to the purpose of the proposed mathematical model are removed;
- A model of decision from the theory of Multi- Attribute Decision Making (MADM) is built [4, 5] highlighting the entities and their characteristics and other information necessary for the optimization;
- Over the decision model, there are generated problems which are multiple solved by applying various optimal choice methods [6];
- If the optimum is multiple, for obtaining the global optimal one applies, to the subset of the optimal entities, an optimal procedure based on an extended set of attributes.

For the acquisition and processing of information related to anthropic computer incidents, the optimum software – hardware infrastructure is strictly built by following the methodological instructions above.

3. Data Base Management Systems' Taxonomy

The database is „a comprehensive collection of information, but non-redundant, in relation to a specific purpose, the information being structured and interconnected to ensure the independence of applications that use them". Databases are structured on three levels: logical, conceptual and physical. Working on the three levels is achieved through so-called Data Base Management Systems (DBMSs). Currently, when talking about a data base its DBMS is considered as well. Therefore, the taxonomy of the DBMSs, necessary to become aware of their characteristics, will be presented in this section. There was no need to obtain a "general taxonomy" which would study the DBMSs making use of concepts such as classes, objects, attributes, relations, functions, restrictions, rules, axioms, events etc, in other words it was not necessary to make an ontological study [7].

Taxonomy is a scientific tool for establishing laws of classification and systematization for areas of reality which present a complex structure. One field in question involves separating its elements into groups and these into sub-groups, the passing 'taxon' → 'taxa' being an operation which can be iterated as many times as desired and possible. Switching from one group to subgroups must ensure that the subgroups are unambiguous, mutually exclusive, and their reunion will result in the group from which it started. In the following paragraphs the main clustering of databases will be highlighted.

3.1 Clustering by data distribution

One of the most important points of view in databases' clustering is the distribution of data and applications. Thus there exist:

- *Centralized databases* are those databases that are hosted by a single computer and all users get information for their applications exclusively on this basis, either in local or remote access;
- *Distributed databases* are those databases that are hosted by multiple nodes (with well-defined geographical addresses) of a computer network, and there is at least one user application that uses data from at least two nodes of the network. These databases are the most efficient but also require the use of synchronization tools and transactions protection under conditions of data logical integrity;
- *Federated databases* are those distributed databases that have a certain structure and hierarchy of the distribution, the simplest hierarchical structure is "star". They consist in a set of independent databases, considered at basic level, working for a higher-level database. Databases from the basic level may cooperate or not with each other. The master database does not reply information from basic level databases, but it processes that complex information for decision making. Federated databases can be managed with the same DBMS and they are called homogeneous, or can be managed with different DBMS and then they are called heterogeneous. Usually they are used when the amount of data to manage and use is high and related to different geographic regions, widely separated from each other;
- *Mobile databases* are actually a special case of distributed databases, indicating that nodes do not have a well defined geographical address. They made their presence felt with the advent of laptops, notebooks and mobile phones.

3.2 Clustering by data structure

In terms of data structuring, databases fall into two categories:

Structured databases, which divide into:

- *Hierarchical databases*, first databases that have emerged after the stage where the files were independent and were managed through File Management System (FMS) of the operating system. These databases are based on models such as the tree, the sequential multi-graph without circuits and lately the n-dimensional hypercube;
- *Relational databases*, inspired by relational algebra, thus benefiting all possible operations with sets (complementary union, intersection, difference, etc.);
- *Object oriented databases*, implements the object-class model that applies to both data structure and the building of data manipulation functions;

Unstructured databases, which can be:

- *Audio / video databases,* containing audio-visual recordings;
- *Topological databases,* built for Geographic Information Systems (GISs), with specific data structures;
- *Big data databases*, which emerged with the development of the Internet. Although in some regions may contain elements of structure, because of their wide diversity of structure, they are considered unstructured. Instead of the existing data manipulation language, like in the case of databases described above, a technique called "data mining" was developed, which provides the user with the interface to the database and data processing finding. Also, for the data handling, search engines were developed to provide access to digital content even to people without computer knowledge.

3.3 Clustering by the users number

Taking into account the number of users, databases can be:

- *Single-User Database,* the single-user feature is given by the fact that at some point in the operation of the database, there is a single user. Although outdated, the single-user feature is essential in certain cases where updates and processing operations must be staggered in time so that a user must wait for the session of another user's work*;*
- *Multi-user database*, the multi-user feature is given by the fact that at the same moment, in the operation of the database, there are more users.

3.4 Clustering by size

This clustering has undergone significant changes over time. Both software and hardware resources have been developed at a sustained pace and users have developed and used applications increasingly significant in terms of size and complexity. Usually this classification

takes more into account the digital content of databases, but the DBMS capabilities are also considered sometimes so that it may manage large databases. The new classification outlines four categories of databases:

- *Small and medium databases,* are usually the single-user and multi-user databases developed for and implemented on micro or mini-computers;
- *Big and very Big databases* are necessarily multi-user databases implemented on computer systems / network with technical performance, large internal / external memory and special computational speed.

3.5 Clustering by memory support

By this criterion, databases are divided into:

- *Databases stored on disk,* representing the classic way in which databases have been developed;
- *Databases stored in memory,* they were identified as necessary databases with continuous operation (24 hours databases). Although they seem slightly volatile, there are many areas of their use.

3.6 Clustering by data dynamics

By this criterion, databases are divided into:

- *Operational databases* that are updated in real time and accurately reflect, at any time, system status data;
- *Analytical databases* containing static data stored periodically, at certain moments in time, to conduct statistics and decision support at a tactical or strategic level.

3.7 Clustering by digital content

In terms of digital content, databases are divided into traditional database, audio, video, multimedia, space research, bibliographic and security.

- *Traditional databases* themselves are divided into *scientific databases, technical, economic, managerial, administrative, socio-humanistic* etc. These databases use logical representations of data in formats such as numeric, alphabetic, alphanumeric, dates, memo and blob, i.e. formats which first appeared in the use of electronic computers*;*
- *Audio and video databases* allow management of audio and video in various formats and standards having a main compression technique working with large volumes of data;
- *Multimedia databases*, as indicated by their name, manage data with very different types and formats and use equipment specially designed and constructed for the purchase or play. The most expressive example for this type is databases for electronic games;
- *Spatial databases*, working with spatial-temporal objects, are specific computer-aided design mapping tools and, in the same time, geographic information systems (GIS). It should be noted that, working with traditional data applications hosted by this type of database, the systems developed can be very complex;
- *Documentary databases* manage various documents into editable format or image format. They have developed special tools for keyword search phrases, templates, etc.
- *Bibliographic databases* manage corrections of references as well as libraries with very large digital content. If the first category offers the author's name, title, possibly the summary, publication date and publisher, for the second category, besides from this type of information, it can give even the full content of incunabula, books, articles and images eloquent for them*;*
- *Security databases* manage all the incidents that occur in an informatics system / Intranet / a well-defined area of the internet.

Briefly, the databases' taxonomy is given in Table 1.

4. Available DBMSs' Clustering

In this section are given taxonomic characterization of the available DBMS sites of agencies involved in computer security:

- Access 2013 – generates the following types of databases: *centralized, structured (relational), multi-user, small, stored on disk, analytical, traditional (managerial)*.
- Paradox – generates the following types of databases: *centralized, structured (relational), multi-user, small, stored on disk, analytical, traditional (managerial)*.

Table 1. The taxonomy of the databases

Distribution	Structure	Users	Size	Memory	Dynamics	Content
Centralised	Structured	Single-user	Small	Disk	Operational	Traditional
Distributed	- Hierarchical	Multi-user	Medium	Memory	Analytical	Audio
Federative	- Relational		Big			Video
Mobile	- Object oriented		Very big			Multimedia
	Unstructured					Spatial
	- Audio – video					Documentary
	- Topological					Bibliographic
	- Big data					Security

- MySQL – generates the following types of databases: *distributed / federative, structured (relational), multi-user, big, stored on disk, operational, traditional (technical and economic)*.
- Informix – generates the following types of databases: *distributed, structured (object oriented), multi-user, big, stored on disk, operational, multimedia*.
- IMS – generates the following types of databases: *centralized, structured (hierarchical), multi-user, big, stored on disk, operational, traditional (technical and economic)*.
- IDMS – generates the following types of databases: *centralized, structured, (hierarchical), multi-user, big, stored on disk, operational, traditional (technical and economic)*.
- DB2 – generates the following types of databases: *distributed / federative, structured (relational), multi-user, big, stored on disk, operational, traditional (technical and economic)*.
- Oracle – generates the following types of databases: *distributed / federative, structured (relational), multi-user, big, stored on disk, operational, traditional (technical and economic)*.
- GRASS GIS – generates the following types of databases: *centralized, unstructured, multi-user, big, stored on disk, operational, spatial.*

5. Possible Software Infrastructure

The databases taxonomy, contained in Table 1, is used for the reduction operation on the set of available DBMSs. It offers the complex clustering, implicitly providing all the characteristics of the databases. From the study referred to DBMS sites available, it appears that they can generate almost all classes of databases.

In this section, the interest is only in databases for the acquisition of information concerning the incidents that can occur in software - hardware systems from a very large geographic area. One will try to determine the necessary and sufficient characteristics of databases for the acquisition of data concerning the incidents and thus be able to drastically reduce the DBMS set that generates them.

First it should be noted that computer users not connected to the Internet are not the subject of this paper, they are assumed to be careful not to bring to their own computer / computers sources of incidents.

Users of computers connected to Internet must have security service providers that take care of the area including their computers. Accordingly, a first databases level for the purchase of information concerning the computer incidents may consist of databases belonging to Internet security providers. As it was asserted at the introduction, if these databases are running independently, each must be coupled with a superior database, hosted by a national / international organism with

attributions in cyber security that can do complex and comprehensive analysis in the meeting areas. This two level construction may be reiterated vertically any number of times.

After the analysis in terms of distribution, it can be concluded that the chosen DBMS must be able to generate a *federated* database. See Figure 1.

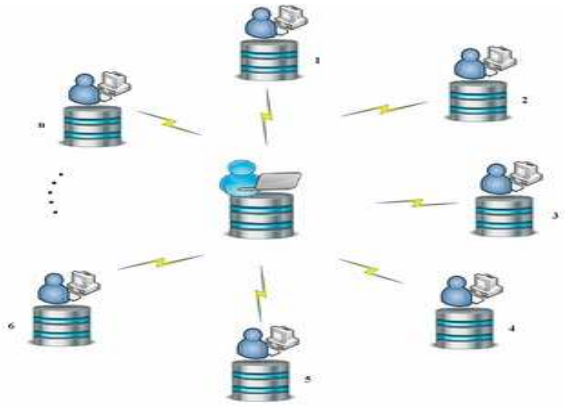

Figure 1. The federate databases

In terms of structure, it must bear in mind that any network administrator takes care of several computers linked together and that topology will clearly influence the propagation of cyber incidents in the network. On the other hand, the incoming server of the network will create a precedence relationship with the rest of the computers. Also one takes into account the fact that the upper level contains analysis indicators strongly interconnected.

Therefore, it will conclude that the database considered for recording and processing incidents must be a database that allows structuring, and it must be *relational*.

Compared to the number of users, it is obvious that the database must be *multi-user*. In terms of size, there may be several hundred databases on the first floor and a single database on a superior level. More effectively, a national system would be more rational for the system to have three levels, i.e. a basic level corresponding to the level of the Internet providers, the intermediate level corresponding with the county and the national level with maximum synthesis function, providing decision support for establishing computer security policy and instruments to sustain this policy. Do not forget that the system must be interconnected with other national systems that share information on demand or at a predetermined time interval. In conclusion, the database must be very *big*.

The support of the database will be the *disc*. This is enforced by the fact that it is working with very big volumes of current data, and must also be preserved the statistical data order in to determine the evolution of phenomena pertaining to cyber security.

At the core, the database must be *operational* and at the upper level, it must be *analytical*. This is obvious. In the first case, any delay in recording incidents will lead to erroneous decision making task. In the second case, making tactical or strategic processing may be required for the baseline to update data before launching their processing.

In terms of content, the database is a *security* database. In this class it is also a *managerial* database because it is made considering the management of cyber security assurance activities at an operational, tactical and strategic level.

Consequently, the DBMS pattern characterization vector for purchase and processing of data related to anthropic computer incidents is: *Federative, Relational, Multi-user, Very large, Disk, Analytical+ Operational, Security*.

For some presented DBMSs, in number of eleven, the reason to be removed from the set of DBMSs, which with it originally started, is that the vector of characteristics does not coincide with this pattern, pattern specific to a DBMS that are considered capable of generating databases for recording and processing cyber incidents. Through this process of elimination remains a set of DBMSs with a cardinality significant lower in comparison with the initial set.

Basically, using the Table 1 as a screen through which the available DBMSs are selected, it stands out as being eligible the following DBMSs: ***MySQL*** [8, 9], ***DB2*** [10, 11, 12] and ***ORACLE*** [13, 14].

In conclusion, the mathematical model's entities will have as software component these three DBMSs.

6. Possible Hardware Infrastructure

The agencies are able to use and choose, as possible, the following hardware infrastructure for generating and running of those databases:

An ad-hoc network **-**A server, average configured, plus 10 CORE2 EXTREME QUADCORE computers.

A Grid site **-** RO-01-ICI, integrated in the National Profile infrastructure and also in the European Grid infrastructure.

A Cloud computing site **-** SoftLayer, equipped for the use of high-performance computing, free for one month of use.

7. DBMSs - Hardware Couples

It is recalled that in the previous sections were chosen as convenient three DMBS and three hardware supports considered as acceptable for building a system for cyber incidents acquisition and processing. There will therefore be nine database - hardware support pairs:

o(1) = MySQL – Ad-hoc network,
o(2) = MySQL – Grid,
o(3) = MySQL – Cloud Computing,
o(4) = DB2 – Ad-hoc network,
o(5) = DB2 – Grid,
o(6) = DB2 – Cloud Computing,
o(7) = ORACLE – Ad-hoc network,
o(8) = ORACLE – Grid,
o(9) = ORACLE – Cloud Computing.

In the language of MADM theory [15], this couples set is called the set of $o[i]_{i=\overline{1,9}}$ objects, the model analyst's task being to determine the optimal object.

This is possible by highlighting, for the set of objects, a lot of independent features, observed, measured or computed, in relation to which it is possible to determine the optimum object. In the MADM language, these characteristics are called attributes or discriminators.

It is noted that columns in Table 1 are common to all objects and therefore can be taken directly as attributes $a[j]_{j\equiv\overline{1,7}}$ in model. It is also known that object discrimination is intended and in terms of hardware infrastructure and so naturally another attribute $a[j]_{j=8}$ can be added.

Pattern analyst must think how to properly assess these eight attributes for each item. First attributes takes turns and differentiate for each DBMS, thus 7 × 3 = 21 sub-attributes to be taken into account in the assessment and then at the attributes of model itself. Therefore, sub-attributes will be:

sa(1) = Distribution – MySQL,
sa(2) = Distribution – DB2,
sa(3) = Distribution – Oracle,
sa(4) = Structuring – MySQL,
sa(5) = Structuring – DB2,
sa(6) = Structuring – Oracle,
sa(7) = Number of Users – MySQL,
sa(8) = Number of Users – DB2,
sa(9) = Number of Users – Oracle,
sa(10) = Size of Digital Content – MySQL,
sa(11) = Size of Digital Content – DB2,
sa(12) = Size of Digital Content – Oracle,
sa(13) = Storage mode – MySQL,
sa(14) = Storage mode – DB2,
sa(15) = Storage mode – Oracle,
sa(16) = Dynamics – MySQL,
sa(17) = Dynamics – DB2,
sa(18) = Dynamics – Oracle,
sa(19) = Digital Content – MySQL,
sa(20) = Digital Content – DB2,
sa(21) = Digital Content - Oracle.

One can observe that there have been determined the names of the lines and columns of a matrix called the decisional matrix $oa[i,j]_{i=\overline{1,9},\ j=\overline{1,8}}$ and the computing base was established for the first 7 attributes, the 8th attribute being directly evaluated.

Next, the values for each entity of this decisional matrix will be given. This will be done in three steps:

a. It will be appreciated the valence note to produce and manage a database with features given conform with taxonomy,
b. It will be appreciated the valence note to produce and manage a database in terms of exploitation using the 3 chosen to support hardware,
c. The decision matrix notes are filled in with other data related to the so-called decision-making context for the attributes: variation limits, weights and the criterion of the optimal (i.e. min or max).

Thus, for the chosen DBMSs, the valence notes, expressing the characteristics listed in conformity with the taxonomy, are:

- Compared to the distribution, in order to build a federated database, one can say that the valence to produce and manage a database thus can be noted as MySQL = 8, DB2 = 10, Oracle = 9;
- Compared to the structure, in order to build a relational database, it can be appreciated that the valence to produce and manage a database thus can be noted as MySQL = 9, DB2 = 10, Oracle = 10;
- Compared to the number of users, in order to build a multi-user database, it can be appreciated that the valence to produce and manage a database thus can be noted as MySQL = 9, DB2 = 9, Oracle = 10;
- Compared to the size of digital content, in order to build a very large database, it can be appreciated that the valence to produce and manage a database thus can be noted as MySQL = 9, DB2 = 9, Oracle = 10;
- Compared to the storage on magnetic media, in order to build a database on disk, one can say that the valence to produce and manage such a database in addition the fast access to high capacity disks can be noted as MySQL = 9, DB2 = 10, Oracle = 9;
- Compared to dynamic data management, in order to build a database that is both operational and analytical, one can say that valence to produce and manage a database thus can be noted as MySQL = 9, DB2 = 10, Oracle = 10;
- Compared to digital content, in order to build a database for security management, one can say that valence to produce and manage a database thus can be noted as MySQL = 9, DB2 = 10, Oracle = 10.

In order to produce and manage a database performing in terms of good exploitation, the use of the 3 chosen supporting hardware is assessed like in following:

- Any database hosted by the ad-hoc network will be awarded 8 points;
- Any database hosted by Grid or Cloud computing will be awarded 10 points.

Taking into account the obtained marks for each attribute and renumbering, we get the decisional matrix $oa[i,j]_{i=\overline{1,9},\ j=\overline{1,8}}$, contained in Table 3. The lines and columns of this matrix are indexed by $a[j]_{j=\overline{1,8}}$ and respectively by $o[i]_{i=\overline{1,9}}$. In this table, beside the decisional matrix, for the attributes there are also highlighted: the variation limits *Lo* and *Up* (in this case, grades from 1 to 10), the absolute weights *W* (real numbers from 0 to 1) and the considered sense in optimization (*m* = minimum or *M* = maximum). In this context, the decisional matrix is complete and consistent [16]. Also, the matrix contains the vectors needed for solving using TOPSIS method i.e. o^+, o^-, d^+, d^- and $eval\_o[i]_{i=1,9}$.

8. Solving with TOPSIS Method

One chooses, as a method of solving the multi-attribute decision problem, the TOPSIS method (Technique for Order Preference by Similarity to Ideal Solution).

The method is based on computing of the distance between the model objects and two *ideal* points. The first point, called *ideal+* and marked o^+, is an ideal object characterized by attributes with values defined by $a^+[j]=\max_{i=\overline{1,i}} oa[i,j]$ for those $j\in\overline{1,j}$ that denotes a maximum attribute, or with values defined by $a^+[j]=\min_{i=\overline{1,i}} oa[i,j]$ for those $j\in\overline{1,j}$ that denotes a minimum attribute. The second point, called *ideal-* and marked o^-, is an ideal object characterized by attributes with values defined by $a^-[j]=\min_{i=\overline{1,i}} oa[i,j]$ for those $j\in\overline{1,j}$ that denotes a maximum attribute or with $a^-[j]=\max_{i=\overline{1,i}} oa[i,j]$ for those $j\in\overline{1,j}$ that denotes a minimum attribute. In the opinion of the method's authors, the distance taken into account must be the one derived from L2 norm, the Euclidean distance, which gives a better measure in calculating the distances. The optimal solution is given by the object which is at the lowest distance away from the positive ideal point and, at the same time, at the highest distance away from the negative ideal point. The indicator expressing, for every object, the mix of those distances is called *merit* and computed by a formula that ensures that it takes values between 0 and 1. All the i merits are recorded in the vector $(eval\_o[i])_{i=\overline{1,i}}$ that will highlight, taking into account the maximum of its elements, the optimal object / objects.

1st step: It consists in applying the weights of the attributes to the corresponding elements of the objects-attributes matrix:

$$oa[i,j]=oa[i,j]\cdot w[j] \quad (\forall) i=\overline{1,i},\ j=\overline{1,j}.$$

But in this case this step is not performed because the model was not given absolute weights (importance) of the attributes.

2nd step: Because the model contains only maximum attributes, the attributes vectors, characterizing the ideal points, are computed by the formulas:

$$a^{+}[j]=\max_{i=\overline{1,i}} oa[i,j] \text{ for } j=\overline{1,j},$$
$$a^{-}[j]=\min_{i=\overline{1,i}} oa[i,j] \text{ for } j=\overline{1,j}.$$

By computing, one obtains $a^{+}[j]_{j=\overline{1,8}}$= (10, 10, 10, 10, 10, 10, 10, 10) and $a^{-}[j]_{j=\overline{1,8}}$= (8, 9, 9, 9, 9, 9, 9, 8), which will characterize the imaginary objects o^{+} and o^{-}.

3rd step: For each $i\in[1,i]$, the Euclidean distance between o_i and o^{+} respectively o^{-} is determined in $\Re^{j}$ by the formulas:

$$d^{+}[i]=\left(\sum_{j=1}^{j}\left(oa[i,j]-a^{+}[j]\right)^{2}\right)^{1/2},$$
$$d^{-}[i]=\left(\sum_{j=1}^{j}\left(oa[i,j]-a^{-}[j]\right)^{2}\right)^{1/2}.$$

By computing, one obtains $d^{+}[i]_{i=\overline{1,9}}$= (3.74, 3.16, 3.16, 2.44, 1.41, 1.41, 2.44, 1.41, 1.41) and $d^{-}[i]_{i=\overline{1,9}}$= (0, 2.24, 2, 2.82, 3.46, 3.46, 2.44, 3.16, 3.16).

4th step: Perform evaluation of the function which gives the objects merits

$$eval_o[i]=\frac{d^{-}[i]}{d^{-}[i]+d^{+}[i]},\ (\forall)i=\overline{1,i}.$$

By computing, one obtains $(eval_o[i])_{i=\overline{1,9}}$ =(0, 0.41, 0.39, 0.54, 0.71, 0.71, 0.50, 0.69, 0.69).

5th step: The decision matrix is filled with the TOPSIS algorithm's computing results, see Table 3. The optimal objects are highlighted through the analysis of the vector of merits.

In this case one notices that the biggest merits are those of the *o*[5] and *o*[6] objects, therefore they are optimal. At the first trial, any of them can be chosen for implementation.

6th step: If the optimal decision-making context extends with new attributes, the unique optimum can appear.

In this case, the decisional context is extended with a new attribute, the operating costs of the federated databases:

Table 2. The final solution

Operating costs of the database (Lei/month)	
o_5=DB2-Grid	o_6=DB2-Cloud computing
1000	950

Consequently, the DB2 database which runs on Cloud Computing is the desired optimum.

9. Conclusions

The novelty and originality of the method for determining the optimal torque software infrastructure - hardware infrastructure for the acquisition and processing of data related to computer anthropic incidents produced by using cyber weapons are judged according to research in the border areas that it approaches. Its complexity is related to the complexity of

Table 3. Decisional matrix in standard work format

O					o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	o_9	o^{+}	o^{-}
A	W	Lo	Up	m/M											
a_1		1	10	M	8	8	8	10	10	10	9	9	9	**10**	**8**
a_2		1	10	M	9	9	9	10	10	10	10	10	10	**10**	**9**
a_3		1	10	M	9	9	9	9	9	9	10	10	10	**10**	**9**
a_4		1	10	M	9	9	9	9	9	9	10	10	10	**10**	**9**
a_5		1	10	M	9	9	9	10	10	10	9	9	9	**10**	**9**
a_6		1	10	M	9	9	9	10	10	10	10	10	10	**10**	**9**
a_7		1	10	M	9	9	9	10	10	10	10	10	10	**10**	**9**
a_8		1	10	M	8	10	10	8	10	10	8	10	10	**10**	**8**
d^{+}					**3.74**	**3.16**	**3.16**	**2.44**	**1.41**	**1.41**	**2.44**	**1.41**	**1.41**		
d^{-}					**0**	**2.24**	**2**	**2.82**	**3.46**	**3.46**	**2.44**	**3.16**	**3.16**		
eval					**0**	**0.41**	**0.39**	**0.54**	**0.71**	**0.71**	**0.50**	**0.69**	**0.69**		

the areas addressed together as well as the use of the latest computer techniques. The method was based on multidisciplinary tools belonging to: Computer Science (Hardware, Databases, Computer Security), Taxonomical Analysis, Operation Research (Multi-Attribute Decision Making) designed to provide together the optimal solution to solve a problem with a high degree of difficulty.

Writing this paper, the authors open the way to development of very complex models in the field of software and hardware infrastructure acquisition for cyber security centres. Further developments can appears by enlarging the attributes set, introducing multiple states of nature, meaning that there are variant of specialized hardware, and considering more than one expert in the assessment process.

REFERENCES

1. VERT, G., R. DOURSAT, **A Fuzzy Taxonomic Approach for Classifying and Identifying System Attacks and Automating Attacks Response**. In: Proceedings of 4th WSEAS International Conference on Computational intelligence, man-machine systems and cybernetics, Miami, Florida, USA, November 17-19, 2005, pp 29-34.
2. XU, R., D. C. WUNSCH II, **Clustering**, Wiley-IEEE Press, 2008.
3. EVERITT, B. S., S. LANDAU, M. LEESE, D. STAHL, **Cluster Analysis**, Wiley Series in Probability and Statistics, 4th ed., 2009.
4. TZENG, G. H., J. J. HUANG, **Multiple Attribute Decision Making: Methods and Applications**, Chapman & Hall, CRC Press, 2011.
5. YOON, K. P., C. L. HWANG, **Multiple Attribute Decision Making: An Introduction**, SAGE Publications, 1995.
6. RESTEANU, C., M. ŞOMODI, M. ANDREICA, E. MITAN, **Distributed and Parallel Computing in MADM Domain using the OPT CHOICE Software**. Wisconsin, USA: In: Proceedings of the 7th WSEAS International Conference on Applied Computer Science (ACS'07), 2007, pp. 376-384.
7. STAAB, S., R. STUDER, **Handbook on Ontologies**, Springer eBooks, Series: International Handbooks on Information Systems, 2009.
8. SCHWARTZ, B., P. ZAITSEV, V. TKACHENKO, J. D. ZAWODNY, A. LENTZ, D. J. BALLING, **High Performance MySQL: Optimization, Backups, and Replication**, Sebastopol: O'Reilly Media, 2008.
9. THOMPSON, L., L. WELLING, **PHP and MySQL Web Development. The definitive guide to building database-drive Web applications with PHP and MySQL**, Boston: Addison-Wesley Professional, 2008.
10. VAN DER LANS, R. F., **SQL for DB2 Developers: The Complete Guide for Optimal Performance**, Indianapolis: IBM Press, 2007.
11. CHONG, R. F., C. LIU, **DB2 Essentials: Understanding DB2 in a Big Data World** (3rd Ed.), Indianapolis: IBM Press, 2013.
12. MULLINS, C. S., **DB2 Developer's Guide: A Solutions-Oriented Approach to Learning the Foundation and Capabilities of DB2 for z/OS** (6th Ed.) Indianapolis: IBM Press, 2012.
13. KYTE, T., **Expert Oracle Database Architecture: Oracle Database 9i, 10g, and 11g**. Programming Techniques and Solutions, New York: Apress, 2010.
14. MAFTEI, E., C. MAFTEI, **ORACLE from 9i to 11g for Application Developers** - Vol. 1 (part. 1+2), Cluj-Napoca: Publisher Albastra, 2010.
15. RESTEANU, C. **MADM - Theory and Practice**, ICI Publishing House, Bucharest, (in Romanian), 2006.
16. ERGU, D., G. KOU, **Data Inconsistency and Incompleteness Processing Model in Decision Matrix**, Studies in Informatics and Control, vol. 22 (4), 2013, pp. 359-366.