Crisis-Eval: A crisis information analysis and evaluation system

Petros Kofakis, Stelios Aleifantis

Keywords: *crisis, early warning, information retrieval, open source, event data*

Introduction

The main issue in the development of a methodological approach for implementation of an information system for crisis early warning is the definition of the limits of the phenomenon of crises in the level of political science and more specifically that of international relations. The academic study of this phenomenon has already offered a lot in the comprehension of crises. Our research work of the last four years attempted to propose and define for implementation a methodological approach, which will cover two distinct objectives:

- the sufficient management of an enormous and dynamic volume of information that is relevant with the problem of crises in which the special analysts can have easy, rapid, convenient and precise access, and
- the sufficient parameterisation of relevant information so as to create the necessary conditions for the definition of an information evaluation model that will enable the special analysts to systematically monitor the life cycle of a crisis, with a particular interest in the prediction of inter-state relations' crises.

In our work, research on the level of international relations and policy science have been matched with research in the field of information technology in order to ensure the development of a system that could contribute in the solution of the crises problem.

Setting the limits of the crises phenomenon

As Michael Brecher pointed out, crisis phenomena are understood in terms of a "unified crisis model", namely as a real world "international crisis" and as "foreign policy crisis" based on actors perception of the evolving situation. Therefore, the "unified crisis model" approach of the crisis phenomenon is based on the simultaneous existence of two general conditions in a conflicting situation:

A. a change in type and/or an increase in intensity of disruptive interaction between two or more states (actors), with the heightened probability of military hostilities, that in turn destabilizes their relationship and challenges the structure of an international system; and
B. the hostile act, disruptive event or environmental change generate three perceptions at actor level, namely:
    a. threat to one or more basic values
    b. finite time for response; and
    c. heightened probability of military hostilities before the challenge is overcome

The "unified crisis model" approach composes a crisis situation which enables the formation of specific variables, which in turn are subject to parameterisation, a cornerstone for the development of an *integrated information system* for situation monitoring and crisis early warning systems.

The central working hypothesis, tested and proved in our pilot application, for the development of an *integrated information system* is based on the following key-elements:

- The crisis:
  - concerns *specific issues* in the context of a conflicting situation between two sides,
  - is revealed with *concrete events* (actions or statements) producing changing tensions, and;
  - the appearance of specific perceptions of the involved sides is *expressed by actions or statements* in order the sides to cope with the evolving challenges.

- The *changing level of tension* in a conflicting situation between two sides constitutes a secure indication for monitoring deteriorating process that is the transformation of this situation from conflict to crisis and beyond.

- The crisis is a dynamic phenomenon that evolves in various *successive phases*; each phase has specific characteristics and is delimited by the existence specific «marginal situations» *thresholds* (onset, escalation, de-escalation phases)

- The systematic monitoring of "action-reaction" process between the two sides forming an *"event data" operational environment* as the core of an *integrated information system* offers an effective tool for information gathering, information analysis and crisis detection process in defined conflict situations.

As already mentioned, a major part of our general conceptual and methodological approach was based on the theoretical "unified crisis model", but our focus of inquiry is the dynamics of crisis situation as detected by the monitoring of "action-reaction" process generating tensions. The real time and automated collection, classification, retrieval, correlation and analysis of related information constituted the heart of an *integrated information system.* The scope of this tool is an *operational environment of information dominance* based on 'open sources' which provides the analysts with all necessary data for picture building and the understanding of crisis drives and situation trends. Our pilot tool, "*Crisis Eval"*, practically tested our working hypothesis with success providing optimistic encouragement for radical developments in the context of a concrete research project.

Although "*Crisis Eval*" is concentrating on the crisis phenomenon, building on the relevant concepts and methodologies it is feasible to aim at a wider monitoring area covering both conflict and crisis situations viewed as a continuum in the spectrum of relations between contending sides. In the definition of the crisis differentiation from the phenomena of conflict and war we have used the Michael Brecher distinctions. The relations between the crisis, the conflict, and the war, according to Michael Brecher, are presented in the figure below and cover an essential distinction development of an information system.

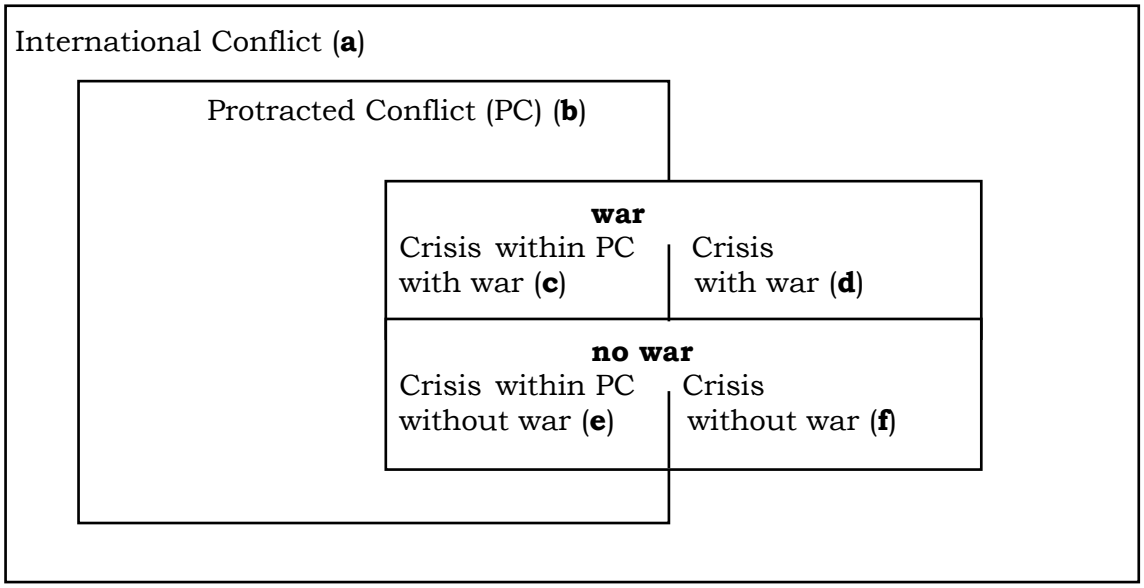


Figure: Conflict, Crisis, War
(Source: M Brecher, Crises in the World Politics:
Theory & Reality, Pergamon Press, 1993)

In any case, it is important to appreciate that both conflicts and crises constitute dynamic phenomena that evolve in various successive phases, each one having its own specific characteristics. This approach allows us to formulate specific "indices" for the follow-up of the situation, to fix «thresholds» for the transition from conflict to crisis and from one crisis phase to the other and from crisis to war or to crisis resolution.

Activities

The respective activities for follow-up of crisis are:

- Situation monitoring: its objective is the systematic collection of information and data, their processing and the creation of a picture of situation in order to emphasize anything that would need special follow-up
- Crisis detection: It includes the use of methodologies that examine the picture of a situation and tries to recognize events or trends that would imply that something extraordinary happens.

The volume of information that is produced daily in printed or electronic form is enormous thus the effort needed from analysts for the retrieval, classification and analysis of information is increasing exponentially.
Our objective was the development of methodologies and tools for the automated collection and classification of information as well as the assistance of analysts, so as to provide a systematic and objective evaluation of events and the analysis of a specific situation. These processes are realised with:

- precision (reduction of information overload)
- speed (reduction of time, hence quick response)
- user friendliness (automation of search, classification and cross-correlation of texts)
- topicality (in time)

The basic tasks that are executed in daily base by an analyst are:

- the collection and the classification of information (current) in electronic form from different sources
- the search and retrieval of this information
- the analysis and evaluation of events in order to estimate of the situation

The basic work scenario can be analysed as follows**:**

- the analysts monitor a list of subjects which can lead to crisis

Daily texts are collected from various sources. From one side we therefore have subjects and on the other hand information (texts). After the analysis of the parameters of each text, depending on subject area it is related, we calculate indicators for the recognition of forthcoming crisis.

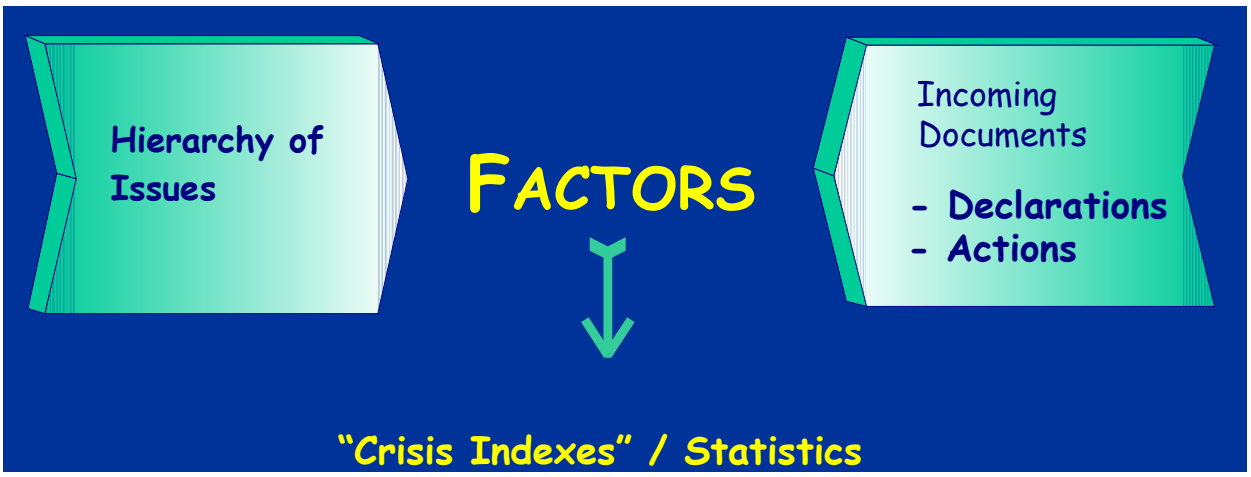


Figure: Scenario of work

In the following figure the data flow of the analysis and evaluation chain is given.

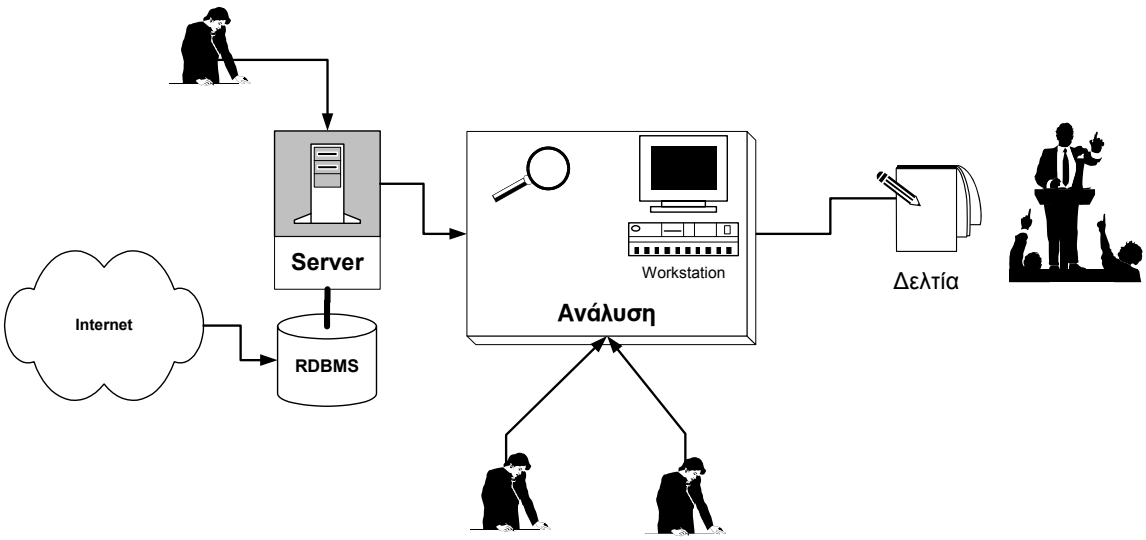


Figure: data flow of work.

Collection of information

Here we will point out once again that all the information and data are collected using open sources that is to say, sources which are accessible by everyone free of charge, or with subscription. These sources do not include information that is collected by other services and that is not directly accessible by everyone (classified, confidential or military content information).

Today, the main source of information from open sources is the internet. The internet and mainly the web constitute an enormous and very rapidly expanding source of information, practically for any subject, in electronic form and with minimal cost. The largest news agencies of the world and most newspapers have systematic internet presence providing news and information in real time (real-time newswire services as Reuters, AP, and online database services). Governments, but also political and other organisations, use systematically the internet for information dissemination, propaganda etc. In internet includes alternative sources of information that supplement the gaps of traditional means and sources. Information that is not present elsewhere is propagated via internet with the use e-mail, internet mailing lists, newsgroups etc. For these reasons the internet constitutes a valuable source of information (intelligence source) that it can provide timely warning for the prediction of crises.

Functionalities of the Crisis-Eval system

The development of the Crisis-Eval system was based on the integration of various heterogeneous systems and technologies at the communication but also at the functional level. The information technologies that were applied include: computers networks, intranet/internet technologies, systems for the classification and search of texts (full text retrieval systems), data base systems, document filtering) and simulation - modelling methodologies and systems. These technologies are mature and most of them available commercially. The solution that was selected was the use of distributed systems and intranet. That is to say, we used a Web Server as central point of access of all the system functionalities. In the user level the application is any web browser.

The basic units - functionalities of the Crisis-Eval system are:

- Management of sources of information
- Collection of information
- Categorisation and thematic organisation of information
- Management of knowledge for information search and retrieval
- Search and retrieval of information
- Evaluation of information - prediction of crises

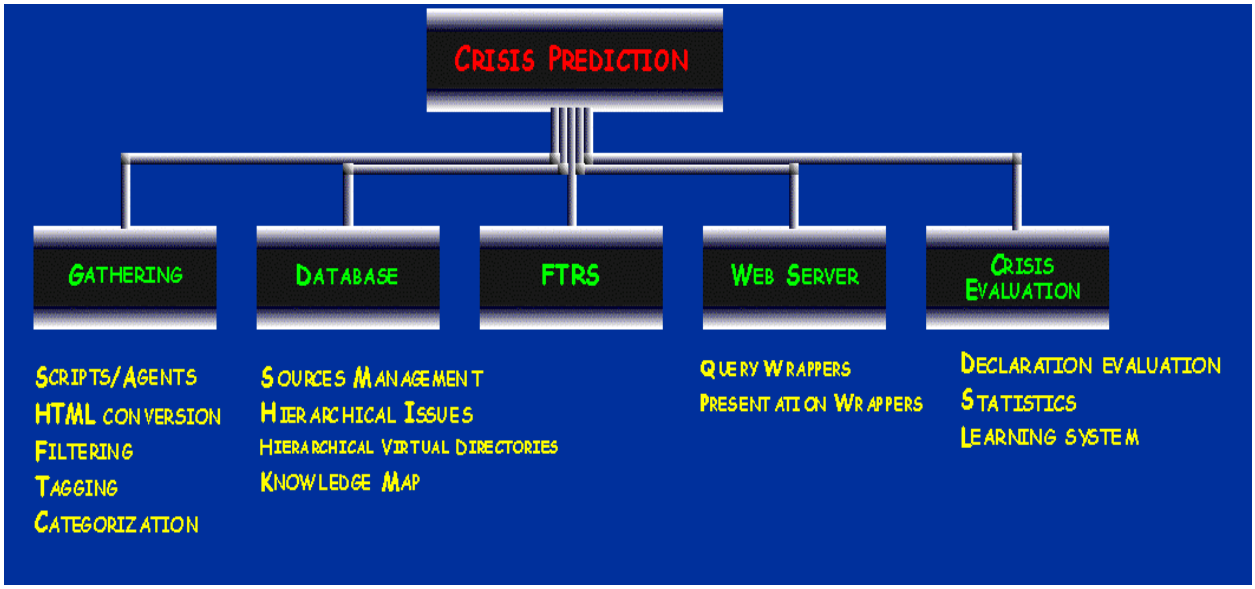


Figure: functionalities of the system

Collection - organisation of information

One of the more important operations of system is the collection of information, which is mostly created dynamically, changes continuously and should be collected in time. Extensive analysis of commercially available systems showed that the existing systems of information collection from the internet (agents, spiders, robots, crawlers, etc) could not be used immediately in our application because of the following reasons:

- they did not allow the changes or modifications that we wanted (they were not open systems)
- they did not allow the communication with custom data bases or other systems
- they could not work in batch mode (parallel / distributed processing)
- they did not allow the import and export of rules for the collection and knowledge in the way that we wanted
- they did not allow the storage of data in the suitable way

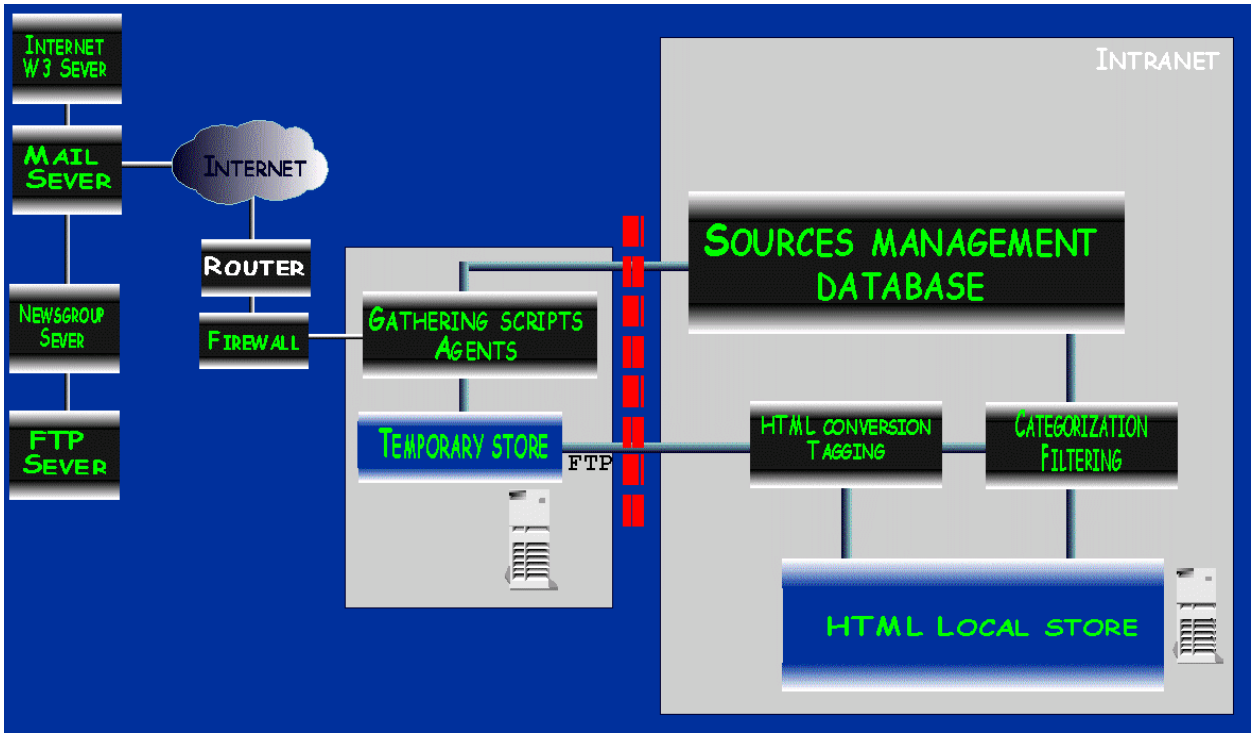


Figure: Information gathering and organization

The solution was to develop a data base for the management of all sources of information (internet URLs) called the "Sources Management database". Indicatively, this database contains the information like:

- Description of source
- Priority
- Way of collection
- Frequency of updates
- Quality/reliability of source
- Way of storage

Category and thematic (knowledge management for retrieval) The sources include sites with news, newspapers, governments, organisms, magazines, universities, inquiring centres, companies, institutions etc. Today, the database contains roughly about 700 sources of direct interest. About 70 sources are updated daily.
The application "Sources Management database" creates daily scripts (Perl) which execute automatically:

- the daily collection of specific information (web pages) from web servers, newsgroups, e-mail, FTP etc
- Cleaning of data and transformation in clean text (without pictures, hyper-links, scripts)
- Automatic appendix of structured information that characterizes each text (date, reliability of source, thematic group etc), as it is determined by the "Sources Management database"
- Storage in the data hierarchy of the local system (hierarchical organisation: source, date, thematic group).

The application allows the use of commercially available agent systems at the same time and also complementary. The storage of all the information as HTML pages allows:

- the access of the stored data in anyway: browsing, FTRS (Full Text Retrieval System) etc
- the independence between the data and the application that manages this information
- free and incremental appendix of structured information on each document text

Information retrieval

The addition of structured information in each text document) is essential for the retrieval and the effective detection of information that has been stored locally (content management). It also allows the search, retrieval and presentation depending on the special requirements of each user or team of users. All this structured knowledge originates:

- from the "Sources Management database"
- from the automatic categorisation systems that process off-line and progressively the documents texts that are stored in system
- from thematic lists and hierarchies that have been defined by the user and stored in the in the “Hierarchical Virtual Directories Database”.

During the search the "virtual" and real sources of information are presented to the user in a hierarchical structure (tree) from where the user can select or search texts.
For the search and retrieval of information that is stored as HTML web pages a FTRS system (Full Text Retrieval System) is used which communicates with the web server in order to receive search commands (queries) and return the search results.

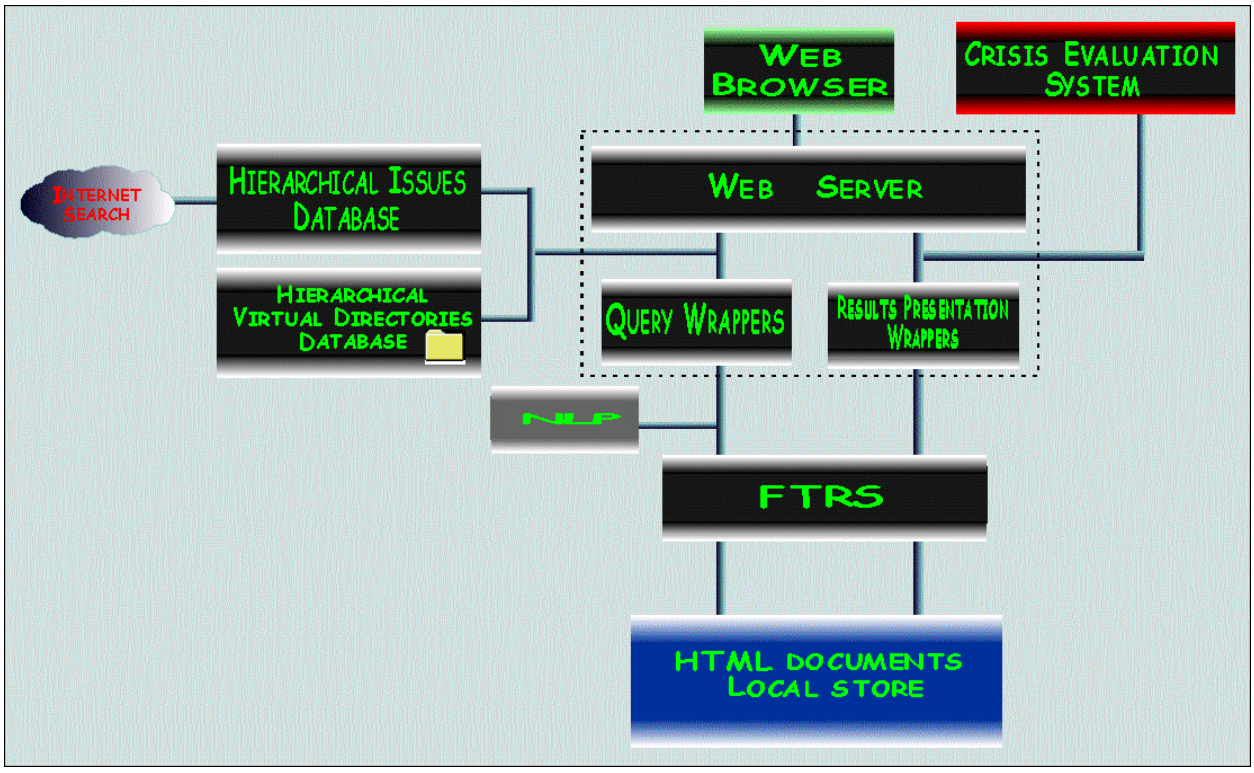


Figure: Search and retrieval of information

In our days the availability of cheap and effective means of data storage has allowed the creation and development of enormous data bases. But, while the collection and storage have been simplified, the problem is located in the retrieval of useful - each time - information especially in very large data bases. In order to exploit the information that has been stored effective ways of detection are required. A texts data base requires from the user a very good knowledge of its structure and subjects so that he can use it efficiently. The effort consequently is focused in the development of methods that will allow the effective detection of information that has been stored.

Precision is the ability of the system to provide only the information that is relevant that is to say, how mush more should I search in order to I find what really I want

Recall is the ability to recover all the information that is relative that is to say, how I can be sure that I have found everything. Most FTRS systems support the management of knowledge and the creation of lists and thesaurus for the thematic organisation of data.

Our objective, when designing the Crisis-Eval system, was the implementation of a sub system for the storage of all the knowledge for the search (that is to say the way the user searches and formulates his queries) and its thematic organisation in way independent of the FTRS system that we use. It has been proved that roughly the 10% of query terms that are used by the user; do not exist in the searched texts. The use of concepts and structured queries can significantly the precision and the recall rate of the system. In the case of our system we have developed a "hierarchical issues database" which it contains a conceptual hierarchical list:

- outlines of subjects
- likely queries (search scenarios) or
- search scripts (pre-formulated complex queries)

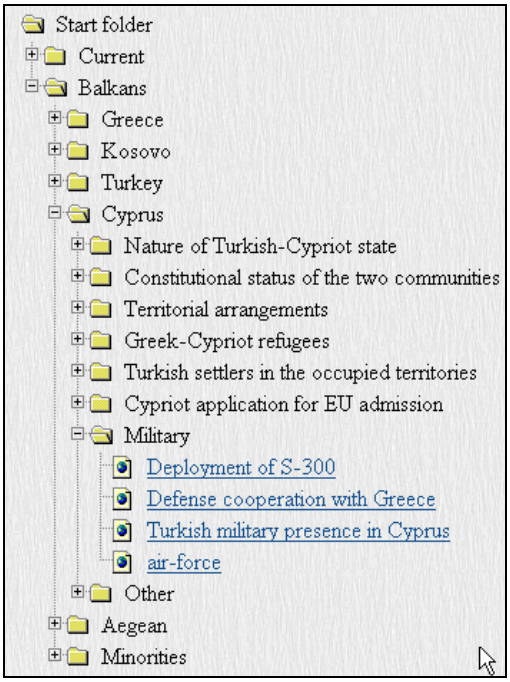
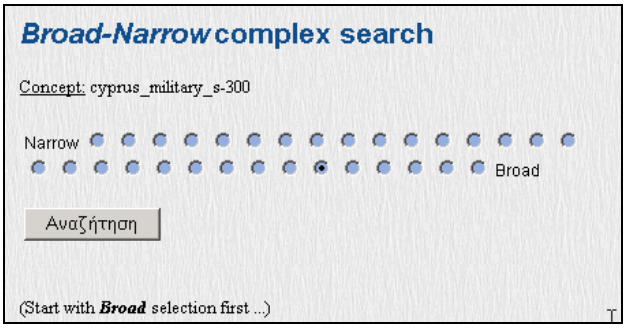


Figure: indicative conceptual search hierarchy

This hierarchical list has the form of tree whose leaves of correspond in concrete search scenarios and contain predefined templates for searches related to a specific subject. These templates are translated automatically in queries according to the syntactic rules (Boolean, proximity, fuzzy or vector) the specific search engine and type of search that we use. The system can also transfer searches to specific internet search engines.

Below we give an example of the concept selection and query that the system creates automatically combining specific terms (keywords) from the database. Example of query that is generated automatically by the system for the concept S-300, term that is perhaps not contained at all in most of the texts that will be searched by the system.	
	English "light" query term: S-300 AND (russian AND missiles) AND russian AND cyprus AND russia AND (S-300 OR missile OR greece OR turkey OR cypriot OR turkish OR (turkish AND cypriot) OR (air AND defence) OR anti-aircraft OR anti-missile OR (military AND action) OR (anti-missile AND defence AND system) OR (ground-to-air AND missiles) OR (surface-to-surface AND missiles) OR (missile AND installations) OR (medium-range AND missiles) OR (defense AND minister) OR mediterranean OR deployment OR island OR israel OR defense OR greek OR crisis OR military)

The user of the database can therefore using specific search scenarios for a particular search topic / subject, can discriminate from the crowd of information, these that he really needs, without the risk to be lost in a enormous volume of information (information overload).

Analysis and evaluation of information

The **stage that follows** the search and retrieval of information is the analysis and evaluation for the estimation of situation. We have therefore developed a tool/model for the evaluation of statements - actions - developments so that we can extract with increased safety and certainty conclusions regarding the possibility of a crisis. The procedures are executed in a systematic and objective manner, that is to say, different users should arrive at the same conclusions for specific events that are entered into the system database.

First Stage: characterization and parameterisation of events

In this stage we input information to the system database in the form of text that describes actions or statements of involved actors in the crisis and this information

is then evaluated manually or automated. This sub system, as the entire *Crisis-Eval* system, is web-based and runs through any browser. The system creates dynamically a hierarchical set of parameters (decision tree) which is used to quantify the parameters that characterize a text. For each text that contains statements or actions, depending on the thematic area that it is related, we calculate a «crisis indicator" which gives a measure of the crisis probability (registration of tendency).

In the following table we give an example of events (actions) for the issue: "*FYROM / Status of the Albanians in FYROM*" and respective calculated ECP (Estimated Crisis Possibility).

Date	Actor1	Actor2	Event	ECP
1/6/2001	m2	mk	Advance elements of the rebel Central Brigade have pushed to within 5km of Kumanovo, in spite of daily Macedonian air strikes, tank artillery and mortar barrages.	5788
9/6/2001	m2	mk	Ethnic Albanian militants seized control of a suburb within shooting distance of Macedonias capital Saturday, sending scores of frightened civilians fleeing amid fears that the insurgency could engulf Skopje... the rebels have threatened to fire on Skopjes airport, near Aracinovo, if attacked.	5788
10/6/2001	mk	m2	The Macedonian army on Monday resumed heavy shelling of villages held by ethnic Albanian rebels, ignoring threats by insurgents to attack downtown Skopje if the onslaught did not cease. The army began hitting the villages of Slupcane and Matejce shortly after first light at 6:15 a.m. 0415 GMT and the blasts were still continuing more than an hour later.	5788
4/6/2001	mk	m2	Macedonian artillery and helicopters attacked ethnic Albanian rebels in their villages on the day new talks to resolve the conflict were due to begin. Yesterday long-range artillery pounded the villages of Matejce, Otlja, Slupcane and Orizare on the lower slopes of Macedonias north-eastern mountains.	5788
6/6/2001	mk	m2	Detonations and shootings can still be heard from the villages of Gajre, Lisec, Sipkovica and Popova Sapka above Tetovo.	5788
8/6/2001	mk	m2	The army began artillery and helicopter attacks in the early morning against villages to the northeast of the capital Skopje that the rebels have controlled for over a month.	5788
9/6/2001	mk	m2	- In a new bid to halt a five-month guerrilla rebellion, Macedonian President Boris Trajkovski in a speech to parliament proposed a partial amnesty today for fighters who lay down their arms. At the same time, the government launched its fiercest artillery bombardment so far against the guerrilla strongholds.	5788
11/6/2001	mk	m2	Macedonian security forces on Monday resumed their bombardment of the positions of ethnic Albanian rebels, ignoring the threat of the rebels to attack Skopje, the capital of Macedonia, Albanias ATA news agency reported According to ATA, the bombardment began at 06:15 am 0415 GMT in the north villages of Sllupcanje and Matejce where government troops have been fighting violently the ethnic Albanian rebels recently. XINHUA NEWS AGENCY BULLETIN 11/06/2001	5788
27/6/2001	mk	m2	New fighting erupted in several parts of Macedonia today, wilting hopes that a NATO-supervised withdrawal Monday of ethnic Albanian rebels from a village near the capital would calm tensions and get peace talks going again. Macedonian forces used artillery north of the capital, Skopje, to strike some of the villages where the rebels were ferried in buses provided by U.S. Army troops. More fighting was reported west of the city of Tetovo in an area where rebels have held territory since February.	5788
3/6/2001	mk	m2	Using artillery and helicopters, the army targeted four villages held by guerrillas since the insurgency, which first erupted in February, flared up again on May 3.	5788
25/6/2001	mk	m2	But the government bombardment of the rebel-held village of Aracinovo in the Skopje suburbs intensified yesterday as the authorities ignored international appeals for restraint.	5788
22/6/2001	mk	m2	Detonations and automatic gunfire resounded on the slopes of Mount Sara above Tetovo. Shooting could also be heard around Slupcane, in the northwest of the country and near Aracinovo, an NLA-held town on the outskirts of the capital.	5788
28/6/2001	mk	m2	Macedonian forces have resumed shelling of positions held by ethnic Albanian rebels, as western diplomats try to shore up support for the government. The shelling of the village of Nikustak, which began yesterday, follows the evacuation of rebels from nearby Aracinovo in a Nato-led operation.	5788
23/6/2001	m2	mk	The guerrillas, who say they are fighting only to end discrimination against the minority, responded to the offensive with an attack on a police checkpoint in the village of Vorce, wounding five members of the security forces.	4736
22/6/2001	mk	m3	A number of undamaged houses belonging to people who had left were set ablaze.	4736

Date	Actor1	Actor2	Event	ECP
11/6/2001	m2	mk	Yet when darkness fell, guerrillas fired on a police vehicle near the city of Tetovo, wounding its occupants and endangering fragile hopes that the first bilateral truce in four months of conflict might be nursed into a serious peace bid.	4736
12/6/2001	m2	mk	Albanian terrorists, lying in ambush, opened fire at Macedonian Interior Ministrys car on the Tetovo-Jazince rood near the village of Odri and wounded 9 policemen in the car.	4736
25/6/2001	al	Lk	The police forces of Kukes seized on Tuesday three Kosovar citizens who were transporting in their vehicle weapons to Macedonia.	4736
25/6/2001	m2	mk	Macedonias MIA news agency reported at 2140 gmt on Monday night that there were new clashes tonight between Macedonian security forces and Albanian terrorist groups in the Kumanovo-Lipkovo area of northern Macedonia. Albanian terrorists fired at Macedonian security forces positions from the direction of the villages of Karadak, Vistica, Brest, Otla, and Matejce, the agency said.	4736
23/6/2001	m2	m1	Ethnic Albanian terrorists on Saturday morning launched an attack on a passenger train travelling between the Macedonian capital Skopje and the western town of Kicevo. said the attack had taken place in the vicinity of Tetovo and that the fighters had used light and artillery weapons. The train was hit by a number of shells, but none of the passengers was injured, the agency added.	4736

Below we describe the essential processes for the operation of the system as well as the work flow that is required for the input and the retrieval of data.

Description of system use

In the lower part of first page (web based application) there is a form for the input of data on any issue. First we select the issue from a list of existing subjects. If the issue is not found in the existing subjects then we can fill the field *other*. Then we select the actor (country or organisation) that has made the action or the statement. In the next list we can select the actor (country or organisation) which is related with the action or the statement (related actor). Then we complete the date of the event, and the type of the event (statement or action). We input the event text optionally the internet address (URL) of the source where we found the text.

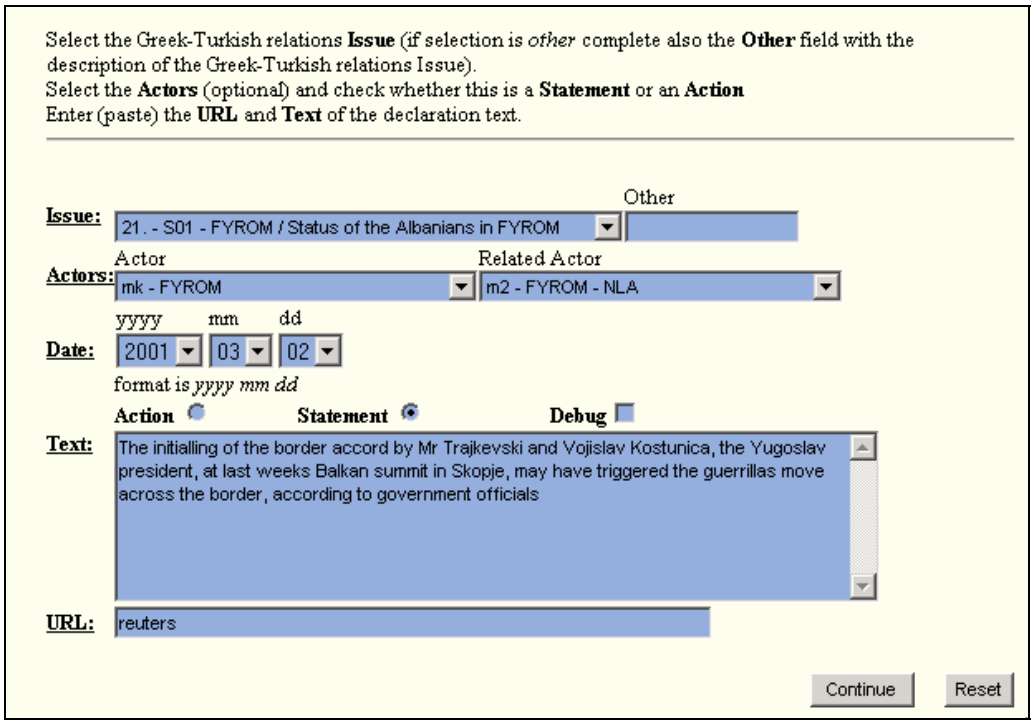
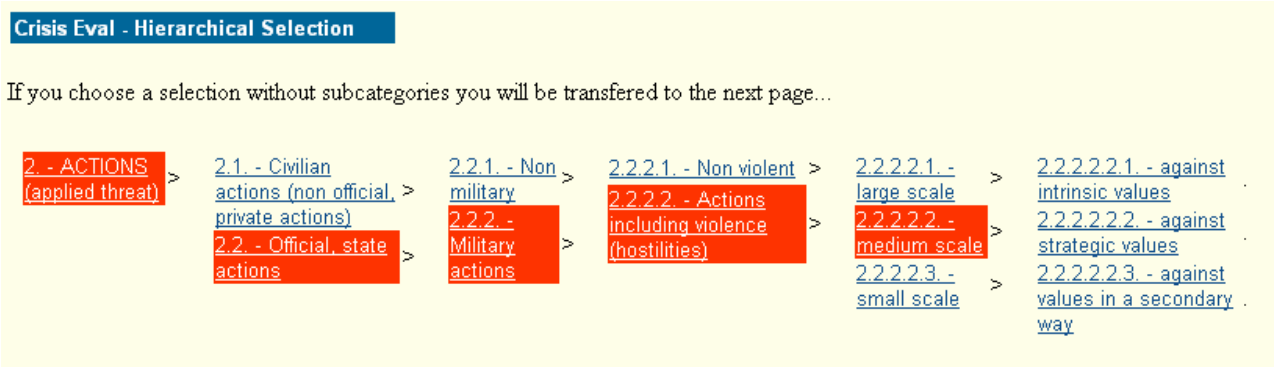


Figure: First page of the application

If the information refers to a statement, then a form for the selection of the country of origin of person that made the statement appears. There exist two lists for the choice of persons for a selected country / organisation. The first one is nominal and contains the names of the persons; the second is the status of the person without the name (e.g. prime minister etc).
In case where the person does not exist, then we have the ability to add in the database the information for the new person.
If the event is an action then we are transferred to the hierarchical decision tree. Then, in the next page, we are must categorise the event according to the hierarchical decision tree options, which reflect the "hierarchical system of parameters". The categorisation tree, in an intermediate state, is shown in the following picture.



Picture: Hierarchical System of Parameters (abstract parameters)

With the red background we present the categories that have been selected. After we select a path in the decision, the selection path and the related category numbers of each node are displayed. This way of categorisation allows an objective categorisation independent from subjective factors related to the specific user of the system. The degree of parameterisation depends from the user requirements and more specifically from the degree of specialisation of the information analyst (researcher or official) and mainly from the aim of use (emphasis in the prediction, situation estimation etc).

During the development phase we have used different degrees of parameterisation, while the pilot system *Crisis-Eval* was based on a balance between the user requirements and in the aim of use of the tool.

At the end of the selection process the special weight of the selected category is presented as well as calculated indicator ECP (Estimated Crisis Possibility). If the information entered is not correct then the user has the possibility to correct the fields he wishes by going back to the corresponding page. If the information entered is correct then it can be imported in the system database. The *Crisis-Eval* system provides ways of easy navigation and search of the events that are stored in the database. The system also provides tools for the modification, deletion and addition of records as well as the addition new countries and persons. There exist three ways for the visualisation of information in the database: as list, in detail, or also with the two ways. We also have the possibility to define filters (for example records that are after a specific date), and define the sorting order for the projection of information that interests the user.

The system also provides ways for easy management. Specifically, we have the possibility to modify all the tables and records of the database:

- the table that contains countries
- the table that describes the degree with which a person can influence decisions
- the country of each person
- the table of subjects
- the table of persons

Graphs

For any issue that we select there exists the possibility to graphically visualise the indicator ECP, for specific actors or combination of actors. In the left side of page via menu we can select the issue, the actors (actor, related actor) and the date for the graph that will appear in right part of the page.

Second Stage: Calculation of metadata and graphics

The basic restrictions that resulted from the application of only the first stage were the following:

- the events were completely independent and distinguishable between them
- the time parameter was not incorporated in the results

This resulted in the creation of graphic representations of ECP indicator that had discontinuities and was difficult follow in the time. Also, it was difficult for the user to interpret situations where we had many events in the same day. Thus, we added an additional stage which can function in cascade with the first without any influence on the design and parameterisation of first stage. The new stage does not suppress, neither contradicts the first one. Functionally it is considered as an additional stage in the output of the old one (cascade) and gives new values (meta-data) for the plotting of graphs.

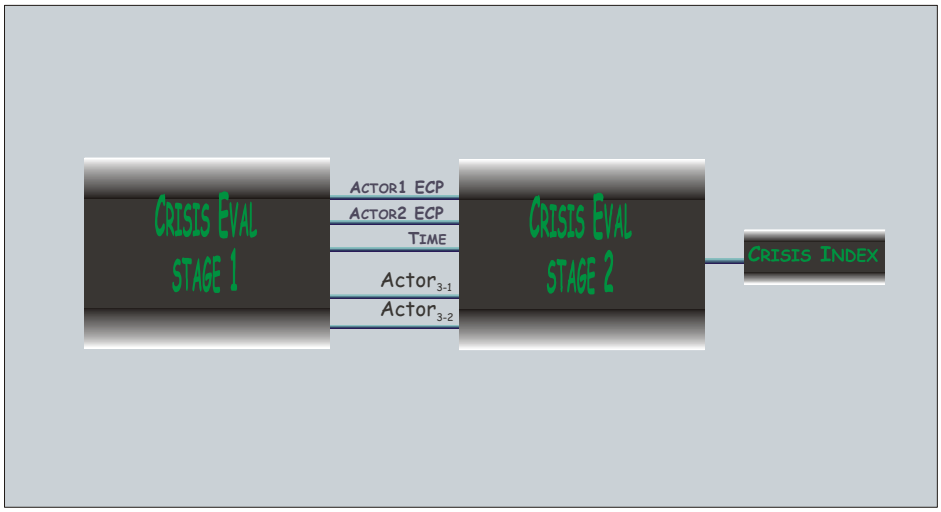


Figure: The two stages of Crisis-Eval

Addition of the 2nd stage, new concepts and assumptions

- *Concept of accumulation*: the first stage did not consider the number of events and the history (track). If in a window of time there do exist many events then these should create higher pressure – escalation than a unique event.
- *Concept of variable time window*: The width of a time window can be altered in order to compensate uncertainties in the data input. i.e., difference of the time the event happened and the time of input in the database, but also in the percentage of the coverage of the events. The predefined (default) time window is 24 hours.
- *Concept of detonation*: If time interval between two events is large then pressure – escalation that is created from a new event should be smaller than that if it happened after a short time interval.
- *Effect of third actors*: The first stage did not take into account the contribution of 3rd actors (coalitions) in the creation of pressure –escalation.

Physical equivalent of the new model

For the better comprehension of second stage we give one physical equivalent of its model.

- In the model of the second stage we consider two reservoirs, one for each actor (Actor1, Actor2).
- At time t=0 and the two reservoirs are empty
- Each new event of the first actor (Actor1) concerning the second actor (Actor2) adds in the reservoir of actor (Actor1) a specific quantity of liquid (CEV: Crisis Event Volume) that correspond in each one of the cases calculated at the 1st stage by the system of evaluation. The same process applies also for the second actor (Actor2).

- If in a time window (one day or 24 hours) there do exist many events then the total pressure of the reservoir CPt at the specific time moment t is:

  $\mathbf{CP_t = C_{Accu1}*CEV_1 + C_{Accu2}*CEV_2 + C_{Accu3}*CEV_3}$

  where C_{Accu} is the corresponding factor of accumulation for the event [1,2,3….]. In practice the maximum number of events during the same day (24 hours) that we use for the calculation of total pressure of the reservoir CP_t is fixed to 10 events which have been sorted in descending order of intensity. The corresponding coefficients are C_{Accu} *= (1, 0.8, 0.6, 0.4, 0.4, 0.3, 0.2, 0.2, 0.2, 0.1)* and the formula becomes

  $\mathbf{CP_t = CEV_1 + 0.8*CEV_2 + 0.6*CEV_3 + \ldots}$
- the detonation in time is simulated as a leakage in our reservoir. If during day there exist some events then we close the tap that is to say there does not exist any detonation. If no then we consider a detonation equal to $\mathbf{C_{leak}*CP_{t-1}}$ that is to say:

  $\mathbf{CP_t = CP_t - C_{leak}*CP_{t-1}}$

  where C_{leak} is the detonation coefficient. A typical value of C_{leak} = 0.2
- The effect of third actors is simulated as a leakage (negative sign) in the reservoir of a specific quantity of liquid (CEV: Crisis Event Volume). That is to say for each new event of third actor (Actor3) concerning a particular actor (Actor1 or Actor2) we remove from the reservoir of the respective actor a specific quantity of liquid (CEV: Crisis Event Volume) that corresponds in each one of

the "cases" as they are calculated by the evaluation system. If there are many events the concept of accumulation is also in effect (with negative sign).

- The new model allows the future addition in the list of cases of events (statements or actions) that des-escalate (negative sign).
- The total volume in each container – reservoir (sum of the volume that emanates from the actor and volume that emanates from third actors creates a pressure (CP: Crisis Pressure) for each actor.
- The combination of pressures from the two actors creates the combined pressure or escalation of crisis (CCE: Combined Crisis Pressure - Escalation).
- The combined pressure or escalation of crisis is detonated (damping) in an expansion tank with specific damping factor.

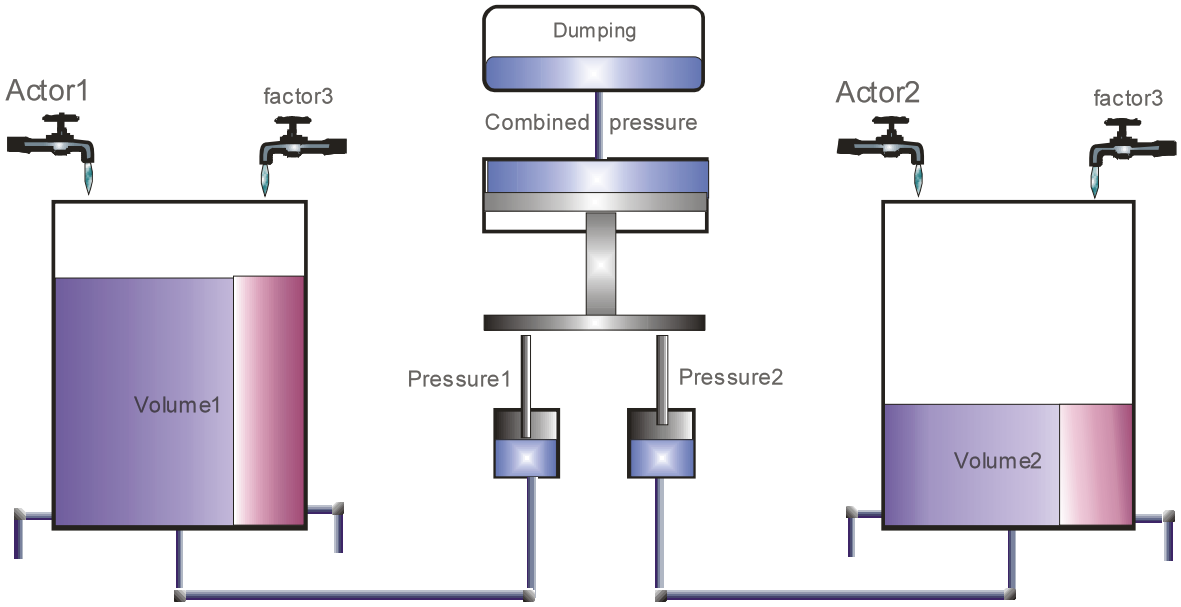


Figure: Physical equivalent of 2nd stage

The mathematical formula is:

$$CP_{Actor1(t_i)} = a_1 * \sum_{t_i} CEV_{Actor1} - a_2 * CP_{Actor1(t_{i-1})} + c_1 * \sum_{t_i} CEV_{Actor3-Actor1} - c2 * \sum_{t_i} CEV_{Actor3-Actor2}$$

Output - Graphs

The new values at the output of the additional stage of the new model that are used for the creation of graph plots are:

- **Actor1CP$_t$**: Pressure that creates actor 1 with respect to actor 2 the specific time (date) t.
  There exists accumulation from the previous days and also from the events of the same day. There also exists a detonation leak. Leakage coefficient $C_{Leak} = 0.2$
  Accumulation factor $C_{Accu} = (1, 0.8, 0.6, 0.4, 0.4, 0.3, 0.2, 0.2, 0.2, 0.1)$
  If there exists an event: **Actor1CP$_t$ = CEV$_1$ + 0.8*CEV$_2$ + 0.6*CEV$_3$ + ...**
  If no: **Actor1CP$_t$ = CP$_t$ - C$_{leak}$*CP$_{t-1}$**

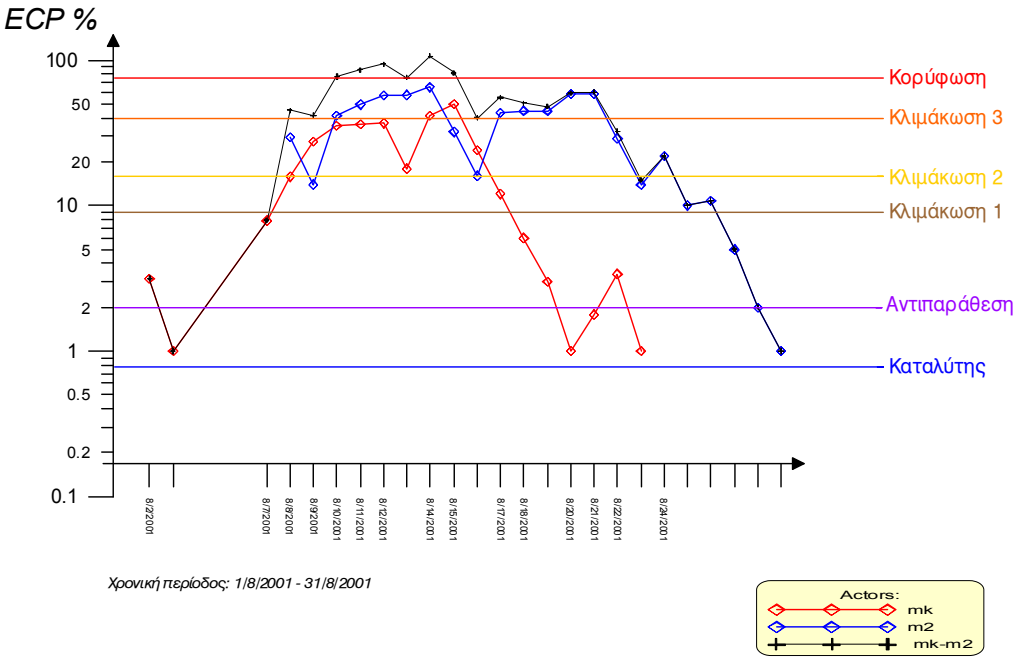
- **Actor2CP$_t$**: Pressure that creates actor 2 with respect to actor 1 the specific time (date) t.
  There exists accumulation from the previous days and also from the events of the same day. There also exists a detonation leak. Leakage coefficient C_{Leak} = *0.2*
  Accumulation factor C_{Accu} = *(1, 0.8, 0.6, 0.4, 0.4, 0.3, 0.2, 0.2, 0.2, 0.1)*
  If there exists an event: **Actor2CP$_t$ = CEV$_1$ + 0.8*CEV$_2$ + 0.6*CEV$_3$ + …**
  If no: **Actor2CP$_t$ = CP$_t$ - C$_{leak}$*CP$_{t-1}$**

- **Actor31CP$_t$:** Pressure that creates actor 1 with respect to actor 1 the specific time (date) t.
  There exists accumulation from the previous days and also from the events of the same day. There also exists a detonation leak. Leakage coefficient C_{Leak} = *0.2*
  Accumulation factor C_{Accu} = *(1, 0.8, 0.6, 0.4, 0.4, 0.3, 0.2, 0.2, 0.2, 0.1)*
  Actor31CP$_t$ = CEV$_1$ + 0.8*CEV$_2$ + 0.6*CEV$_3$ + …

- **Actor32CP$_t$**: Pressure that creates actor 3 with respect to actor 2 the specific time (date) t.
  There exists accumulation from the previous days and also from the events of the same day. There also exists a detonation leak. Leakage coefficient C_{Leak} = *0.2*
  Accumulation factor C_{Accu} = *(1, 0.8, 0.6, 0.4, 0.4, 0.3, 0.2, 0.2, 0.2, 0.1)*
  Actor32CP$_t$ = CEV$_1$ + 0.8*CEV$_2$ + 0.6*CEV$_3$ + …

- **Actor12CP$_t$:** Combined Crisis Pressure - Escalation (CCE)
  Actor12CP$_t$ = (Actor1CP$_t$ + Actor2CP$_t$)/2

- **Actor31-2CP$_t$:** Pressure difference that creates actor 3 with respect to actor 1 and actor 2.
  Actor31-2CP$_t$ = Actor31CP$_t$ - Actor32CP$_t$
  Positive values indicate support of actor 3 to actor 2, negative values indicate support of actor 3 to actor 1.



ΔΙΑΓΡΑΜΜΑ 7 - FYROM CRISIS 2001/ Status of the Albanians in FYROM

Figure: example of graphics plot

Conclusions

Up to today the use of the Crisis-Eval system, which is under continuous development, allowed:

- the detection and registration of the analysts (experts) knowledge
- increase of precision (reduction of information overload)
- reduction search time of (hence quick response)
- automation of search, classification and cross-correlation of texts
- creation of a large database (historical database)
- standardisation and homogenisation of the parameters for the description of a situation and a crisis
- precise qualitative and quantitative follow-up of a crisis development

References

1. Balabanovic, M. and Shoham, Y. (1995) *Learning information retrieval agents: Experiments with automated Web browsing*. Proceedings of the 1995 AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments, Stanford, AAAI Press.
2. Balabanovic, M., Shoham, Y., and Yun, Y. (1997) *An adaptive agent for automated web browsing*. Stanford University Technical Report CS-TN-97-52.
3. Bartell, B.T., Cottrell, G.W., and Belew, R.K. (1994) *Learning the optimal parameters in a ranked retrieval system using multi-query relevance feedback*. Proceedings of the Symposium on Document Analysis and Information Retrieval Las Vegas.
4. Belkin, N. & Croft, B. (1992) *Information Filtering and Information Retrieval: Two Sides of the Same Coin.* Communications of the ACM, 35, 12, pp. 29 – 38.
5. C. Buckley, J. Allan, and G. Salton (1993) *Automatic retrieval with locality information using SMART*. Proceedings of the First Text REtrieval Conference TREC-1, pages 59-72.
6. C. J. van Rijsbergen, (1979) *Information Retrieval*. Butterworth & Co (Publishers) Ltd., London, second edition.
7. Chen, H. (1995) *Machine Learning for information retrieval: Neural networks, symbolic learning and genetic algorithms*. Journal of the American Society for Information Science (JASIS), 46(3), 194-216.
8. Clifford A. Lynch (1991) *The Z39-50 information retrieval protocol: An overview and status report.* ACM Computer Communication Review, 21(1):58- 70.
9. Cohen, W.W. (1995) *Text categorization and relational learning.* Proceedings of the International Machine Learning Conference (ICML-95), pp. 124-132, Lake Tahoe, CA, Morgan Kaufmann.
10. Cohen, W.W. and Singer, Y. (1996) *Context-sensitive learning methods for text categorization.* Proceedings of SIGIR-96.
11. Czeslaw Danilowicz (1994) *Modeling of user preferences and needs in Boolean retrieval systems*. Information Processing & Management, vol. 30, no. 3, pp. 363-378,
12. David D. Lewis (1992) *Representation and Learning in Information Retrieval*. PhD thesis, Department of Computer Science, University of Massachusetts.
13. David K Gifford, Robert W. Baldwin, Stephen T. Berlin, and John M. Lucassen (1985) *An architecture for large scale information systems.* Proceedings of the 10th ACM Symposium on Operating Systems Principles, 19(5):161-171, December 1985.
14. Donna Harman (1993) *Overview of the first Text REtrieval Conference.* In Proceedings of the 16th Annual International ACM/SIGIR Conference , pages 36-48, Pittsburgh, PA.
15. Douglass R. Cutting, Jan O. Pedersen, Per-Kristian Halvorsen, and Meg Withgott (1990) *Information theater versus information refinery.* In Paul S. Jacobs, editor, AAAI Spring Symposium on Text-based Intelligent Systems.
16. G. Salton (1968) *Automatic Information Organization and Retrieval*. McGraw-Hill Book Company.
17. G. Salton and M. J. McGill (1983) *Introduction to Modern Information Retrieval.* McGraw-Hill Book Company.
18. Gerard Salton (1989) *Automatic Text Processing - The Transformation, Analysis, and Retrieval of Information by Computer*. Addison-Wesley.

19. H. C. Arents and W. F. L. Bogaerts (1993) *Concept-based retrieval of hypermedia information - from term indexing to semantic hyperindexing*. Information Processing and Management, 29(3):373-386.
20. J.P. Callan, W.B. Croft, and S.M. Harding (1992) *The inquery retrieval system.* Proceedings of the Third International Conference on Database and Expert Systems Applications, pages 78-83.
21. Joe Carthy and Alan Smeaton (2000) *The design of a topic tracking system.* In Proc. 22nd Annual Colloquim on Information Retrieval Reseach, Cambridge, UK.
22. Johannes Fürnkranz, Johann Petrak, and Robert Trappl (1997) *Knowledge Discovery in International Conflict Databases.* Applied Artificial Intelligence, 11(2), pp. 91-118.
23. Lang, K. (1995) *NewsWeeder: Learning to filter NetNews.* Proceedings of the 1995 Machine Learning Conference, ICML-95, Lake Tahoe, CA, Morgan Kaufmann.
24. M. Perkowitz and O. Etzioni (1995) *Category translation: Learning to understand information on the Internet*. Proc. IJCAI-95, Montreal, Morgan Kaufmann.
25. Marcia J. Bates (1990) *Design for a subject search interface and online thesaurus for a very large records management database*. In Proceedings of the 53rd Annual Meeting of the American Society for Information Science, pages 20-28.
26. Michael R. Genesereth, Arthur M. Keller, and Oliver Duschka (1997) *Infomaster: An Information Integration System*. In proceedings of 1997 ACM SIGMOD Conference, May 1997.
27. P. B. Danzig, S.-H. Li, and K. Obraczka (1992) *Distributed indexing of autonomous Internet services*. Computing Systems, vol. 5, no. 4, pp. 433-459.
28. R. Armstrong, D. Freitag, T. Joachims, and T. Mitchell (1995) *WebWatcher: A learning apprentice for the World Wide Web*. Proc. 1995 AAAI Spring Symp. on Information Gathering from Heterogeneous, Distributed Environments, Stanford, March 1995, AAAI Press.
29. R. Cooley and J. Srivastava (1997) *Web Mining: Information and Pattern Discovery on the World Wide Web (A Survey Paper)*. Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'97), November 1997.
30. Raya Fidel. (1991) *Searcher's selection of search keys: I. the selection routine.* Journal of the American Society for Information Science, 42(7):490-500.
31. Robert Trappl (1992) *The Role of Artificial Intelligence in the Avoidance of War*. Cybernetics and Systems '92, R. Trappl (Ed.), pp.1665-1680. World Scientific Publishing: Singapore.
32. Robert Trappl (1986) *Reducing International Tension through Artificial Intelligence: A Proposal for 3 Projects*. Power. Autonomy, Utopia: New Approaches Towards Complex Systems pp.97-103, R. Trappl (Ed.). Plenum: New York.
33. Robert Trappl, Johannes Fürnkranz, and Johann Petrak (1996) *Digging for Peace: Using Machine Learning Methods for Assessing International Conflict Databases*. In Proceedings of the 12th European Conference on Artificial Intelligence (ECAI-96), pages 453-457, Budapest, Hungary, 1996. John Wiley & Sons.
34. Schrodt A.Philip, Event Data in Foreign Policy Analysis**,** in Laura Neack, Patrick J. Haney and Jeanne A.K. Hey (eds), *Foreign Policy Analysis: Continuity and Change in Its Second Generation,* New York: Prentice Hall, 1994
35. Shih-Hao Li and Peter B. Danzig (1995) *Vocabulary problem in Internet resource discovery.* Proceedings of the Second International Workshop on Next Generation Information Technologies and Systems, Naharia, Israel, June 1995, pp. 139-145.
36. Susanne Schmeidl (2001) *Early Warning and Integrated Response Development*. Inter-regional Forum on Coping with Crises and Conflicts, April 2 – 4, 2001, Bucharest, Romania.
37. Turney, P.D. (1997) *Extraction of Keyphrases from Text: Evaluation of Four Algorithms.* National Research Council, Institute for Information Technology, Technical Report ERB-1051.
38. Yiming Yang, Jaime Carbonell, Ralf Brown, Thomas Pierce, Brian T. Archibald, and Xin Liu (1999) *Learning approaches for detecting and tracking news events.* IEEE Intelligent Systems Special Issue on Applications of Intelligent Information Retrieval, 14(4):32 – 43.
39. Yiming Yang, Thomas Pierce, and Jaime Carbonell (1998) *A study on retrospective and on-line event detection.* In Proc. ACM SIGIR, pages 28-36, Melbourne.