Crisis Risk Assessment and Early Warning

Thomas Delavallade
160 bv de Valmy BP 82 92704 Colombes Cedex
thomas.delavallade@fr.thalesgroup.com

Christophe Marsala
Université Pierre et Marie Curie-Paris6, UMR 7606, LIP6, 8 rue du Capitaine Scott, 75015 Paris
christophe.marsala@lip6.fr

Philippe Capet
160 bv de Valmy BP 82 92704 Colombes Cedex
philippe.capet@fr.thalesgroup.com

Claude Michel
160 bv de Valmy BP 82 92704 Colombes Cedex
claude.michel@fr.thalesgroup.com

Abstract

Within a collaboration between THALES Land & Joint Systems and the LIP6 of Pierre et Marie Curie Paris VI University, we have developed a man-machine interface, which aims at helping experts in charge of strategic surveillance. Given a specific domain of interest, the designed tool processes long-term risk assessment based on macro-structural indicators and on-line crisis detection through the monitoring of news reports. Both automatic techniques have been built so that they can provide the end user with understandable results. It has first been applied to intra-state crises.

Keywords: Early Warning, Risk Assessment, Ontology, Fuzzy Logic.

1 Context and Goals

After the recent South East Asia Tsunami, crisis prevention has emerged as a critical issue in natural hazard sciences. This trend is not an isolated case, but can be observed as well in many other fields such as nuclear proliferation or civil war surveillance.

In order to set up prevention policies, there is a strong need to forecast the potential crises and also to understand the seminal and immediate causes. In this perspective, we have developed a software designed for strategic analysts. It aims at bringing into relief both long-term and short-term crisis signs, so that analysts can monitor the evolution of a critical situation and identify more easily the levers on which decision makers can act before a crisis happens.

For the longer term, structural risk assessment is carried out, while for the shorter term, an early warning system enables to monitor massive flows of events and launches alerts whenever critical event sequences are detected.

These systems allow analysts to work on a huge quantity of data, either structural indicators or structured events. Furthermore they provide guidelines to analyze their predictions through graphical tools.

We first give details of the two modules of our prototype in the next two sections. In section 4 we describe the implemented interface. Then a specific application centered on intra-state conflicts is presented, while the last section is dedicated to concluding remarks.

2 Risk Assessment Module

Regarding strategic surveillance, it is essential to anticipate as soon as possible the potential emergence of critical situations. This implies being able to forecast such situations and to understand why they are about to occur. This is the core task of the Risk Assessment Module.

Given structural information, which changes rather slowly, a risk level has to be assessed in order to reflect long-term trends that make up the context of the situation. Such information varies depending on the domain of interest and the analysis level required. For instance financial data at the company level are used for bankrupt detection [1], while mortgage loan data are the basis of credit risk assessment [5]. Concerning humanitarian crises, demographic indicators at the country level can be taken into account [7].

From these data various methodologies can be used to perform risk estimation. Qualitative assessments, through the aggregation of a group of experts' judgments, are the common practice, because they are considered more reliable and more accurate as well. A good illustration can be found in the study of country risk rating systems made by Gautrieaud [6]. Yet quantitative methods, partially or fully automatic, are also used and can yield assessments that are seen more objective as long as the internal processing is clear for the end user. These techniques are less expensive and time-consuming since they do not require any expert judgments. This is why we decided to opt for automatic techniques that output transparent results so that they can be accepted or rejected and understood by human experts: the end users.

Most quantitative models are based on purely statistical tools like regression analysis. The models developed by Collier at the World Bank are typical of what is usually done in political science [3]. However these models require non-trivial statistical know-how in order to interpret the results. Another minor but growing trend promotes the use of machine learning tools, which enable to deal with huge amount of input data [5]. Among these techniques, decision tree induction is quite popular, mostly because the learnt tree corresponds to a set of rules, easily understood by the user. Henceforth it fits quite well our requirements. In order to improve the robustness of our model and to get rules closer to human reasoning, we chose to integrate fuzzy reasoning. This led us to use Salammbô, the software developed by C. Marsala to build fuzzy decision trees [9].

From past data, for which we know if the situation is risky or not (class label), a set rules, under the form of a decision tree, is learnt. These rules map the source data to the class label and are then used on current data to predict the risk level of the current situation. Moreover the rules give some hints to understand the situation, from a structural point of view. See [4] for more details on this specific module.

3 Early-Warning Module

The risk assessment module is necessary to globally grasp a situation, but is not sufficient to follow the day-to-day evolution of a critical situation. We need an early warning module to monitor such an evolution and to process a refined and more accurate analysis. Such a task is usually done by experts. We believe however, that automatic processing of ongoing events can be very useful and even essential. Indeed it enables the experts to take into account much more data, that what would have been possible otherwise.

Our dynamic early warning module consists in watching a flow of data related to the phenomenon to be predicted, in order to spot event sequences that are likely to end up in crisis. The data under watch need to be structured before they enter the system. As data acquisition is not the central focus of our work, we assume that text-mining tools are used upstream, like KEDS for conflict related event data [11]. The system is also supplied with a library of scenarios, built by experts, which correspond to typical developments of crises, and an ontology describing the domain knowledge. It contains all the knowledge necessary to link input data to the scenarios.

Some studies in political science, like Schrodt's work [12], developed fully automatic systems; we preferred to include experts' knowledge to guide the system. This is a bit more time consuming, but for specific, well defined and bounded applications, we tend to believe that experts' guidance can provide accurate results. Furthermore, this is a way to make the system accepted by end users since they are more confident in such non-fully automatic systems.

A scenario describes typical developments of a specific type of crises, *i.e.* it depicts how a system moves from a normal, sound state to a critical, anarchical state. We use the template formalism developed at THALES Land & Joint Systems to express such a description. The general principle of a template is the breaking down of a complex phenomenon into a combination (conjunction, disjunction, etc.) of less complex phenomena until elementary phenomena are reached, namely the events directly observed from the input data. A complex template, like for example the one pictured on Figure 2, is in fact a tree, which corresponds to the aggregation of several templates. Fuzzy logic is used to build refined aggregation operators.

The core of our system consists in comparing input event data and known scenarios of crisis developments, through a constrained pattern matching processing. This task enables to detect

the early signs of what usually ends up in crisis (with regard to historical knowledge). It is done by our recognition engine, which assesses the degree of match between event sequences and scenarios, taking into account constraints specified by the experts during the scenario editing. Fuzzy logic can be used to specify these constraints, which eases the scenarios editing since linguistic variables are closer to human reasoning than real numbers. See [10] for detailed explanations of this module.

4 Man-Machine Interface

Our tool was developed not only in order to detect forthcoming crises but also and overall in order to help experts watch the evolution of a situation. That is why we designed a man-machine interface, which scrutinizes event data flows, as described earlier. Analysts who are supposed to be the final users must have tools to understand the historical context and to visualize the flow of events and the evolution of the situation as estimated by the recognition engine. This requires that the risk scores computed by the system are understandable. This is as important as the score computation itself. Indeed experts, who do not understand the results the machine presents them, may mistrust the system, so that it becomes useless. To ensure transparency of the early warning system partially recognized templates are displayed and the user can browse the tree structure, to see which parts of the scenarios are recognized and which are not. He can browse down to the leaves and thus discover which events triggered the recognition. Visualizing the early stage of a partially recognized scenario draws the attention of the user who is then supposed to watch out more carefully the event sequences highlighted by the tool. Concerning the long-term risk analysis, the interface displays the rules used by the system to estimate the risk level. Besides, the user can modify the parameters from which the computation is done, in order to see how the risk level is consequently modify. This interface, presented in Figure 1, and the functionalities offered by the tool to visualize the recognition results, to understand these results, can be considered as the real output of our system.

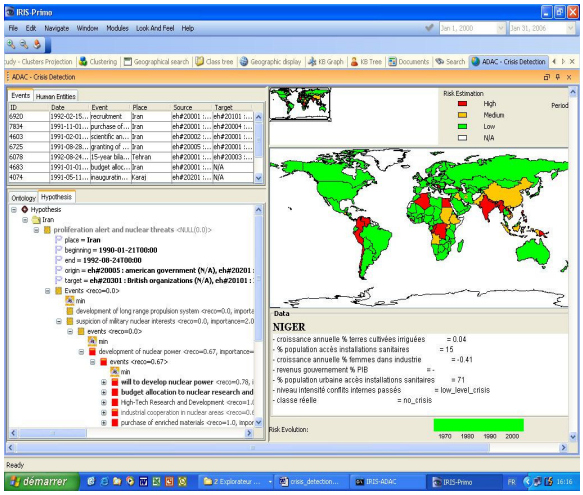

Figure 1: A screenshot of the crisis detection interface

5 Application: ethnic conflicts detection

So far we have depicted in a fairly general way the methodologies implemented in our crisis anticipation system. To make this description clearer we present in this section the application that gave birth to this prototype, namely the intra-state conflicts anticipation.

Regarding the long-term risk assessment module, input data are made of structural indicators, taken at the country level, in various fields such as economy, energy or demography, to name a few. The occurrence of an internal conflict stands for the target concept Salammbô has to learn. The first results are quite satisfying since we managed to reach an 80% recognition accuracy, when estimating the classification performance of our system in a 10-fold cross-validation.

For the early warning module, we focused on the armed conflicts that stroke in the nineties the Great Lakes region in Eastern Africa. The input data are event data coming from the PANDA project [2], which are structured news reports from Reuters. From these news, conflict related events are kept, as well as dates, locations, source and target agents.

Inspired by Gurr and Harff's model of ethnic conflict emergence [8], the scenario of figure has been created. It is of course much simpler than the original model. The goal was to experiment our methodology with a toy scenario, yet rather plausible though simplified, in order to show convincing demonstrations of the usefulness of our tool.

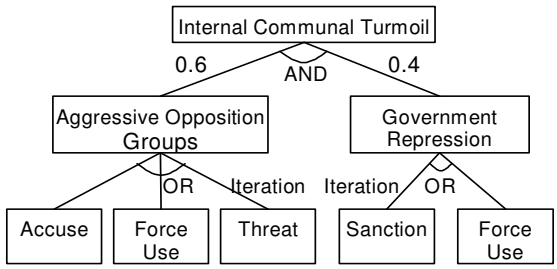

Figure 2: An ethnic conflict template

6 Future Work and Conclusion

In this paper we described the prototype we developed to anticipate crises and how we applied it in the particular context of ethnic conflict detection. Based on generic methodologies, structural risk is assessed to give a contextual overview of the situation, and event data flows are compared to known scenarios of crisis developments, so that on-line monitoring of the evolution of the situation is possible.

In order to show that our methodologies are generic we wish to work on applications related to different types of crises like financial crises.

Our two modules perform some automatic computations of risk scores, for both short and long term. But both intend to provide experts with useful ways of understanding a critical situation, its roots and immediate potential causes. Here is the real goal of our application, which definitely not aims at replacing the experts.

Acknowledgements

We would like to thank Laure Mouillet and Emmanuel Collain from THALES Land & Joint Systems, and to the whole fuzzy logic team: LOFTI of the LIP6 laboratory from Pierre et Marie Curie, Paris VI University.

References

[1] J. Batten et al (2005). Assessing the Time Horizon of Bankruptcy Using Financial Ratios and the Maturity Schedule of Long-Term Debt, *Third European Risk Management Conference*, Antwerpen, Belgium.

[2] J. Bond and D. Bond (1998). The Protocol for the Assessment of Nonviolent Direct Action (PANDA), Codebook for the P24 Data Set.

[3] P. Collier, A. Hoeffler (2001). Greed and Grievance in Civil War, *World Bank working paper.*

[4] T. Delavallade (2005). Country Risk Ratings: A new Methodology to Assess Internal Conflicts Risks, *Third European Risk Management Conference*, Antwerpen, Belgium.

[5] J. Galindo, P. Tamayo (2000). Credit Risk Assessment using Statistical and Machine Learning: Basic Methodology and Risk Modelling Applications, *Computational Economics*, volume 15, pages 107-143.

[6] S. Gautrieaud (2002). Le risque pays : approche conceptuelle et approche pratique, *working paper*, Centre for Development Economics Montesquieu University, Bordeaux IV, France.

[7] J. A. Goldstone et al (2000). State Failure Task Force Report: Phase III Findings, *working paper*, http://www.cidcm.umd.edu/inscr/stfail/.

[8] T. R. Gurr, B. Harff (1998). Systematic Early Warning of Humanitarian Emergencies, *journal of Peace Research*, volume 35, number 5, pages 551-579.

[9] C. Marsala (1998). Apprentissage inductif en présence de données imprécises : construction et utilisation d'arbres de décisions flous, *PhD Thesis*, Pierre et Marie Curie Paris VI University, France.

[10] L. Mouillet (2005). Modélisation, reconnaissance et apprentissage de scénarios de conflits ethno-politiques, P*hD Thesis*, Pierre et Marie Curie Paris VI University, France.

[11] P.A. Schrodt, S. G. Davis, J. L. Weddle (1994), Political Science: KEDS–A Program for the Machine Coding of Event Data, in *Social Science Computer Review*, volume 12 pages. 561–588.

[12] P. Schrodt (2000). Pattern Recognition of International Crises Using Hidden Markov Models (D Richards, Editor), *Political Complexity: Nonlinear Models of Politics*, *Ann Arbor: University of Michigan Press*, pages 296-328.