Paper

Objective classification of empirical probability distributions and the issue of event detection

Janusz Granat and Andrzej P. Wierzbicki

Abstract—The paper concentrates on the issue of classification of empirical probability distributions (histograms), which is useful both in management situations and in event detection or event mining. While existing approaches to event detection concentrate on the use of selected moments or other characteristics of empirical probability distributions, we postulate that full empirical distribution preserves more of needed information then selected moments of this distribution, thus multiple criteria classification of distributions can be most effective in event detection.

Keywords— objective classification, event detection, multicriteria analysis.

1. Introduction

While practically all multiple criteria approaches to decision analysis and support concentrate on rationally supporting *subjective* decisions, depending on some form of an elicitation of preferences of the decision maker, there are diverse decision situations where we should suggest decisions that are made *as objectively as possible*; the full objectivity is not attainable for many practical and philosophical reasons, but objectivity can be seen as an useful ideal or goal. Examples of such situations are, on the one hand, managerial decisions influencing many stakeholders, when an aggregation of preferences of stakeholders is impossible. On the other hand, such situations occur also in event detection, e.g., when automatically detecting a case of fire, we should not make decisions based on subjective, personal preferences.

We shall call the problem of supporting decisions in such a case the problem of *objective classification* (treating problem of *ranking* as a special case with singleton classes and the problem of *decision selection* and *detection* as special cases with classes *selected – not selected* or *detected – not detected*). We can define objective classification as dependent only on a given set of data, relevant for the decision situation, and independent from any more detailed specification of personal preferences than that given by defining criteria and the partial order in criterion space. Already in this definition, we see the limits to objectivity, because naturally the definition of criteria and their partial order, or of the relevant set of data, can be treated as subjective; however, they are often much more obvious and easy to agree upon than the detailed preferences defined, e.g., by a utility function or a set of weighting coefficients.

Most of classical approaches to multiple criteria decision analysis and support, e.g., based on weighted sum aggregation, are not easily adaptable to the case of objective classification. From known approaches, either the goal programming or the reference point approaches are easily adaptable, because goals or reference points can be defined reasonably objectively from statistics in a given set of data. We concentrate here on reference point approaches, because they have the property of producing always Pareto optimal options (which is not the case in goal programming).

In this paper we are focusing on the issue of classification of empirical probability distributions (histograms), which is useful both in management situations and in event detection or event mining. While existing approaches to event detection concentrate on the use of selected moments or other characteristics of empirical probability distributions, we postulate that full empirical distribution preserves more of needed information then selected moments of this distribution, thus multiple criteria classification of distributions can be most effective in event detection. One of advantages of reference point approaches is that they easily deal with so-called *multiobjective trajectory* analysis and optimization; this can be applied to issues of stochastic dominance and their generalizations needed for multiple criteria event detection based on classification of empirical probability distributions.

2. An algorithm for event detection

An outline of an algorithm for event detection based on histograms is as follows:

1. Calculate a typical histogram (based, e.g., on historical data):

   H_{ok} – the value of a typical histogram for an interval k, $k = 1,\ldots,N$.

2. Identify a priori the set of events E and(or) corresponding anomalies A:

   e_i – an event i, $i = 1,\ldots,I$,

   A_i – an anomaly corresponding to event i, $i = 1,\ldots,I$.

3. Identify the set of characteristic histograms for the set of events (using, if necessary, simulation of events and anomalies):

   H_{ik} – a histogram characteristic for event or anomaly i,

   while $i = 0$ corresponds to lack of anomaly (the typical histogram).

4. Define partial orders corresponding to each event (whether an event typically results in a decrease or increase or lack of change of typical histogram values, see below):

   PO_i – a partial order of histograms.

5. Calculate the histogram of the observed current measurements from the real system:

   $H_k, \quad k = 1,\ldots,N$ – the histogram of observed measurements.

6. Calculate achievement values comparing the histogram of observed measurements with histograms of the typical character and of character characteristic for specific events, treated as $I+1$ $(i = 0,\ldots,I)$ different reference trajectories.

7. Detect an event based on achievement values (by selecting the event with corresponding highest achievement value).

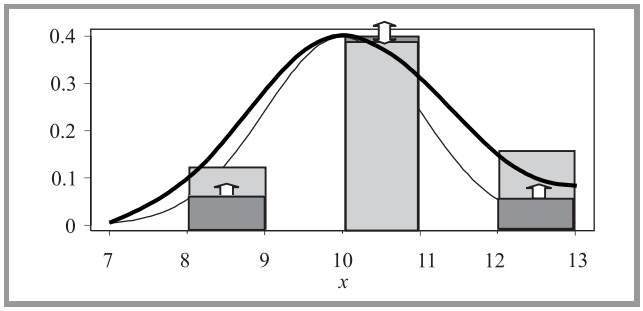

Fig. 1. Partial order of histograms.

A partial order of histograms (Fig. 1) is defined as follows. Two histograms H_α, H_β (with values $H_{\alpha,k}$, $H_{\beta,k}$ satisfy a given partial order $PO_i(H_\alpha, H_\beta)$, if:

$$PO_i(H_\alpha, H_\beta)$$

$$H_{\alpha,l} \le H_{\beta,l}, \qquad l = 1,\ldots,L$$

$$H_{\alpha,k} \ge H_{\beta,k}, \qquad k = 1,\ldots,K$$

$$H_{\alpha,m} \approx H_{\beta,m}, \qquad m = 1,\ldots,M$$

$$K + L + M = N,$$

where:

N – the number of intervals (common for both histograms).

Achievement values can be defined as follows (for the case $K = N$, $L = M = 0$; for other cases see, e.g., [8]):

$$\delta_i(H_k, H_{ik}) \,|\, PO_i = \min_k (H_k - H_{ik}) + \varepsilon \sum_k (H_k - H_{ik}),$$

$$i = 0,\ldots,I$$

Let us consider a very simple example (see Fig. 2). In the top left side of Fig. 2 we have two histograms H_{O1}, H_{O2} that represent histograms for normal operation of the system measured at two different outputs. However, in the bottom left side we have actually observed histograms H_1, H_2 that differ substantially from normal operation. In the right hand side of the picture we have defined histograms for two events e_1, e_2, with arrows indicating partial orders. Do the actually observed histograms correspond to normal operation, event e_1 or e_2?

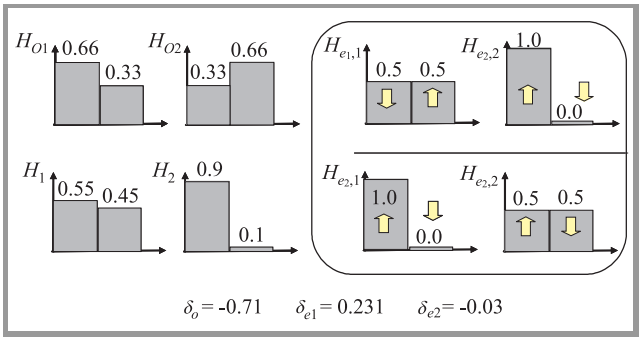

Fig. 2. An event detection.

If we calculate values of the achievement functions for the given histograms H_1, H_2 that result from the actual measurements, using normal operation histogram and the histograms characteristic for the two events as three different reference trajectories, we obtain:

$$\delta_o = -0.71, \quad \delta_{e1} = 0.231, \quad \delta_{e2} = -0.03.$$

Therefore, we can identify event 1, having the maximal achievement value, as the one best representing the measurements and the assumed partial order.

3. Fault detection

In this section we will discuss how the presented approach can be applied to detection of faults in a computer network.

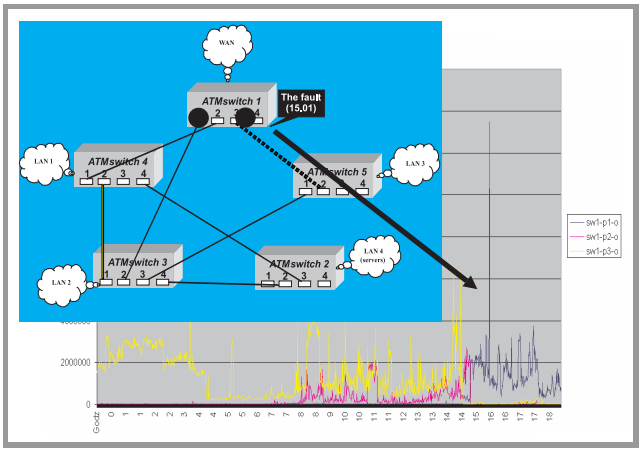

Fig. 3. A fault of the computer network.

Figure 3 shows a network with five switches. On 16.01 we can observe a fault of the connection between switch 1

and switch 5. The observations of the network throughput are presented on Figs. 4 and 5.

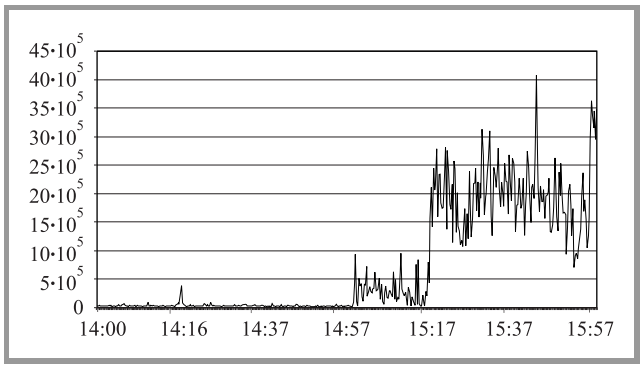

Fig. 4. Observation of the network throughput – switch 1, port 1.

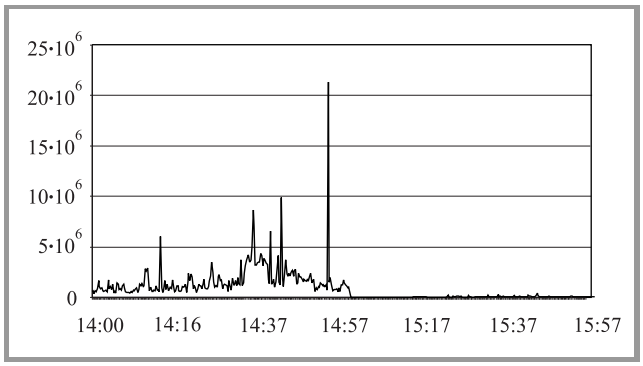

Fig. 5. Observation of the network throughput – switch 1, port 3.

We can transform this observations to the histograms shown of Figs. 6 and 7.

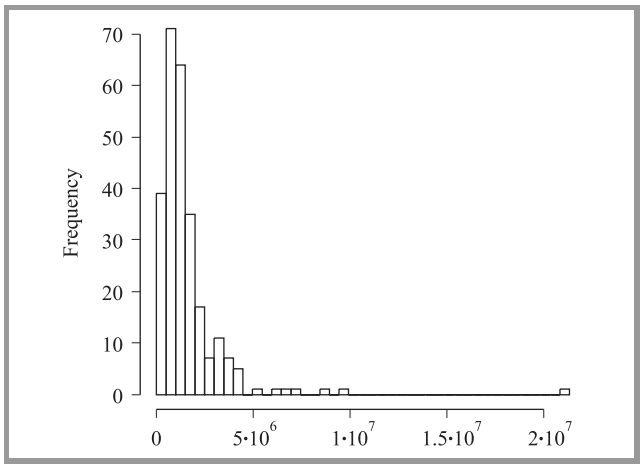

Fig. 6. A histogram of the observation on switch 1, output 3 – normal state of operation.

We see on Figs. 6 and 7 that the number of intervals that are significant from the point of event detection is small. Only these few intervals can be considered in the process of event detection.

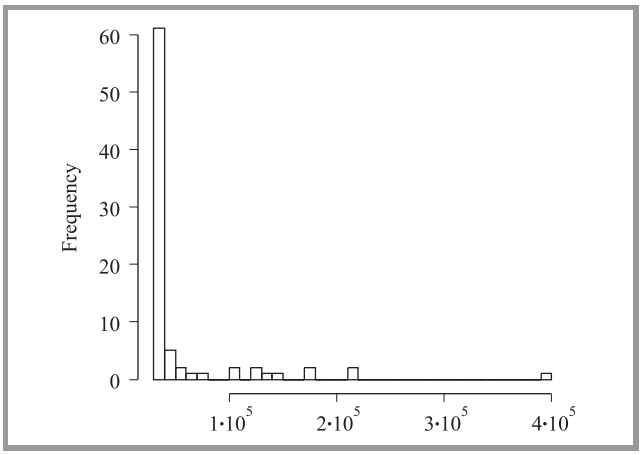

Fig. 7. A histogram of the observation on switch 1, output 3 – after fault of the network.

4. Conclusions

The paper presents a new concept of event detection that is based on histograms and multicriteria approach. This approach allows to consider only selected intervals of the histograms what is especially important in the problem of event detection where the significant changes can be detected only in selected intervals. It is also possible to consider histograms corresponding to observations of various points of the system. In Section 2 we have presented one of possible algorithms for event detections; other variants of the presented algorithm are also possible. The disadvantage of this approach is that we have to specify a priori the given sets of events and the histograms characteristic for given events. The advantage is that we use in a possibly most complete way the statistical information contained in the measurements.

References

[1] J. Granat, "Events driven approach for supporting real-time management of complex systems", in *Proc. First Word Congr. Int. Feder. Syst. Res. IFSR*, Kobe, Japan, 2005.

[2] J. Granat, "Event mining based on observations of the system", *J. Telecommun. Inform. Technol.*, no. 3, pp. 87–90, 2005.

[3] J. Granat and A. P. Wierzbicki, "Multicriteria analysis in telecommunications", in *Proc. 37th Ann. Hawaii Int. Conf. Syst. Sci. HICSS'04*, Big Island, Hawaii, USA, 2004.

[4] J. Granat, "Multicriteria analysis for behavioral segmentation", *J. Telecommun. Inform. Technol.*, no. 3, pp. 39–43, 2004.

[5] D. Luckham, *The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems*. Boston: Addison-Wesley, 2002.

[6] L. Perrochon, W. Mann, S. Kasriel, and D. C. Luckham, "Event mining with event processing networks", in *Pacific-Asia Conference on Knowledge Discovery and Data Mining, LNCS*. London: Springer, 1999, vol. 1574, pp. 474–478.

[7] A. P. Wierzbicki, "On the completeness and constructiveness of parametric characterizations to vector optimization problems", *OR Spektrum*, vol. 8, pp. 73–87, 1986.

[8] A. P. Wierzbicki, M. Makowski, and J. Wessels, *Model-Based Decision Support Methodology with Environmental Applications*. Dordrecht: Kluwer, 2000.

Janusz Granat received his M.Sc. in control engineering (1996) and his Ph.D. (1997) in computer science from the Warsaw University of Technology, Poland. He holds a position as an Assistant Professor at the Warsaw University of Technology, and is the leader of a research group on applications of decision support systems at the National Institute of Telecommunications in Warsaw. He lectured decision support systems and various subjects in computer science. His scientific interests include data mining, modeling and decision support systems, information systems for IT management. Since 1988 he has been cooperating with IIASA. He contributed to the development of decision support systems of DIDAS family and the ISAAP module for specifying user preferences. He has been involved in various projects related to data warehousing and data mining for telecommunication operators. He was also involved in EU MiningMart project.
e-mail: J.Granat@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Institute of Control and Computation Engineering
Warsaw University Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Andrzej P. Wierzbicki – for biography, see this issue, p. 22.