FAULT TREE CONSTRUCTION ON THE BASIS OF MULTIVARIATE TIME SERIES ANALYSIS.

At the analysis of potentially dangerous events (technological catastrophes, extreme phenomena of nature etc) there are often used models which have the form of fault trees (also called error trees). Fault tree (**FT**) allows to present in the graphical form a hierarchy of consecutively occurring stochastic events which cause certain undesirable event. On the basis of FT analysis, the forecast of the undesirable event can be made, as well as measures for its prevention elaborated.

Usually FT is designed by an expert on the basis of his experience and other a prior

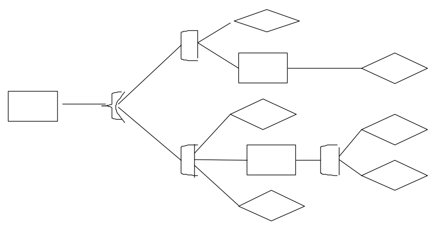

Fig. 1 An example of fault tree. – events; – leaves; – AND gate; – OR gate

information on the given event and on similar events. Often the expert may use ***statistical information*** describing the object under investigation and similar objects, and this information can also be used for the construction of fault tree.

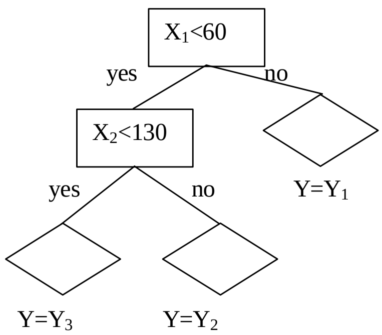

Fig.2. An example of decision tree

We assume that the problem in question has the following specific features:

– the set of features describing objects may simultaneously contain binary, nominal, and quantitative features. Thus, we have a multidimensional series composed of a set of binary, symbolic, and numerical sequences:

– the corresponding distributions of the multidimensional random process are unknown. The decision functions for the prediction are based on a limited statistical material. Since the sample size is small and the space of states has a high dimension, the derivation of statistically stable solutions becomes a problem;

– some features may not be measured at some instants of time. Then their values are missing in the time series.

In the given work we suggest a method for fault tree design (or construction of some of its parts, for finding of which it is necessary to analyze statistical data) by means of the analysis of multidimensional time series describing the dynamic properties of object.

The rest of the paper is organized as follows. In paragraph 2, we give short description of fault trees and decision trees. In paragraph 3, we briefly describe a method of decision tree construction for time series analysis. Paragraph 4 holds an algorithm for fault tree formation from decision tree. In the conclusion, we discuss the suggested method and indicate possible ways of future development.

Fault tree and decision tree

The root of FT (Fig.1) corresponds to so called ***top undesirable event*** (TUE). The nodes of the first layer of FT correspond to some events, preceded TUE and able to cause it. These events can be connected by ***gates*** AND, OR. Besides that, the probabilities that specified events cause TUE are evaluated. Similarly, each node (event) of the first layer of the tree can be matched to nodes (events) of the second layer etc. The leaves of the tree denote the events that are impossible to decompose due to the absence of information or other reasons.

Let it is required to find a model, where feature Y depends on features X. The example (Fig.2) illustrates such a model, which can be seen as a ***decision tree***. The decision tree (**DT**) consists of nodes stand for rectangular, the branches stand for pieces connecting the nodes. Each node corresponds with a certain feature and the branches correspond with a range of values. These ranges of values can be joint in the set of values of the given feature.

When precisely two branches grow out from an internal node (the tree of such type is called a dichotomous tree), each of these branches can give a true of false statement concerning the given feature as is shown on Fig. 2.

The value Y is ascribed for each terminal node of a tree (named "leaf"). In case of pattern recognition problem the given value is a certain class.

For any observation of x, using DT, we can find the predicted value Y. For this purpose we start with a root of a tree, we consider the feature, corresponding to a root and we define, to which branch the observed value of the given feature corresponds. Then we consider the node in which the given branch comes. We repeat the same operations for this node etc., until we reach a leaf. The value Y_S ascribed to *S-th* leaf will be the forecast for x.

Multidimensional time series and decision trees

Let a set of random features $X(t)=(X_1(t),..., X_n(t))$, whose values vary with time, be used to describe a certain object. We denote the set of possible values of X_j by D_j. The following types of features can be distinguished:

– binary feature: $D_j = \{0,1\}$;

–nominal feature: $D_j=\{u_j^1,...,u_j^{l_j}\}$ where $\{u_j^1,...,u_j^{l_j}\}$a is a set of symbols (names);

–quantitative feature: $D_j \subseteq R$ where R is a set of real numbers.

The binary and nominal features are also referred to as qualitative features.

Let features be measured at the consecutive moments of time $t^1,..., t^\mu,...$. For definiteness we will assume that measurements will be carried out through equal intervals of time. We will designate through $x_j(t^\mu)=X_j(t^\mu)$ the value of feature X_j at the moment of time t^μ. Thus, we have n-dimensional heterogeneous time series $x_j(t^\mu)$, j=1,...,n, μ=1,2,....

Let us choose one predicted feature X_{j_0}, $1 \le j_0 \le n$. We designate, for convenience, this feature through Y. In this paper, we consider feature Y of qualitative type.

Let us consider the moment of time t^μ, and also a set of the previous moments of time $t^{\mu-1}$, $t^{\mu-2}$,…,$t^{\mu-l}$, where l is a given size ("deep of history"), $1 \le l < \mu$.

We suppose that conditional distribution $Y(t^\mu)$, when all previous values $X(t)$ are given, depends only on values of series in l previous moments of time.

Besides, we suppose, that this dependence is the same for any value μ. The given assumption means, that the statistical properties of series determining dependence are stationary.

For any moment of time t^{μ} , it is possible to form a set $v^{\mu}=(X_j(t^{\mu-i}))$, $i=1,...,l$, $j=1,...,n$, representing the time series in l previous moments of time. We will call a set v^{μ} background of length l for the moment t^{μ}.

It is required to construct a model of dependence of feature Y from its background for any moment of time. The model allows to predict the value of feature Y at the future moment of time on values of features for l last moments. In other words, the given model, using background, represents decision function for forecasting. In analogy to a usual pattern recognition problem, we will call a problem of the given type a problem of recognition of dynamic object. The analyzed object can change its class in the run of time.

We will represent a decision function for forecasting time series on its background as a decision tree. This decision tree differs from the described in paragraph 2 trees in statements concerning a features X_j in some i-th moment of time back are checked. For convenience, we

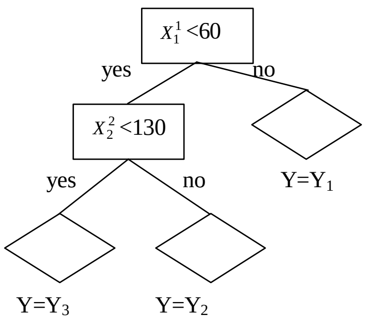

Fig.3. An example of decision tree for time series prediction

will designate these features, with a glance to background, through X_j^i (Fig. 3). Thus, X_j^i means feature X_j in i-th previous moment of time (concerning a present situation).

Let there be a set of measurements of features $X=(X_1,...,X_n)$ at the moment of time $t^1,...,t^N$ and value l is also given. Thus, we have a multivariate heterogeneous time series of length N.

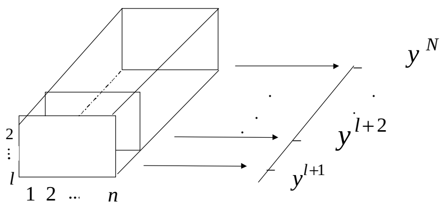

Fig. 4. 3-D data table for time series analysis

We generate set of all histories of length l for the moments of time $t^{l+1},...,t^N$: $A= v^{\mu}$, $\mu=l+1,...,N$.

For any given decision tree for forecasting by background (**DTFB**), it is possible to define its quality: we will designate through $\hat{Y}(t^{\mu})$ predicted value Y received with the help of a tree by background v^{μ}. The criterion of quality will be $Q=\frac{1}{N-l}\sum_{\mu=l+1}^{N}h(\mu)$,where $h(\mu)=L(\hat{Y}(t^{\mu}),Y(t^{\mu}))$, $L(a,b)$ is loss function, determining losses if class a is predicted then b is true class. Thus, Q is the estimate of risk function.

The initial problem of construction of DTFB is divided into some more simple pattern recognition problems.

We will present set v^{μ} as the table $v^{\mu} =(X_j(t^{\mu-i}), i=1,...,l, j=1,...,n$ containing l rows and n columns. Then, the initial information for forecasting is the set of tables v^{μ}, together with the values of predicted feature Y specified for each table $Y(t^{\mu})$, $\mu=l+1,...,N$. It is possible to present set $A= v^{l+1},\ldots, v^N$ as the three-dimensional table of dimension $l{\times}n{\times}(N-l)$ to which the vector $(y^{l+1},...,y^N)$ corresponds (Fig. 4). However, available methods of recognition with using of DT use bi-dimensional tables as input information. Below we suggest one of the possible ways to use given methods for the analysis of three-dimensional data tables.

Consider l tables $X_j(t^{\mu-i}),Y(t^{\mu})$, where $j\in\{1,2,...,n\}, \mu\in\{l+1,l+2,...,N\}, i\in\{1,2,...,l\}$. Thus,

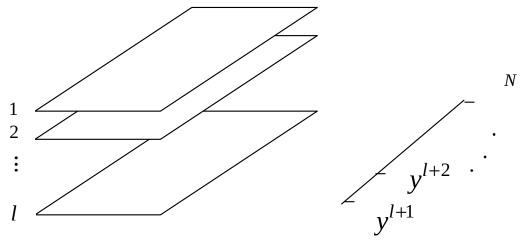

Fig. 5. Horizontal cuts of 3-D data table

we have l horizontal cuts of 3-D data table (Fig. 5). For each cut (two-dimensional table), the decision tree minimizing Q is constructed. In a result we receive a set of trees $T_1,T_2,..., T_l$. We will denote the best of these trees (with minimal Q) as T^*.

From decision tree to fault tree.

Let feature Y corresponds to the coming of some undesirable event: Y=1, if this event comes, Y=0 otherwise. Thus, the number of predicted classes equals two. Consider loss function with the following properties: $L(0,0)=L(1,1)=0$; $L(1,0)<<L(0,1)$ (the error of classification of undesirable event costs much more than the error of classification of

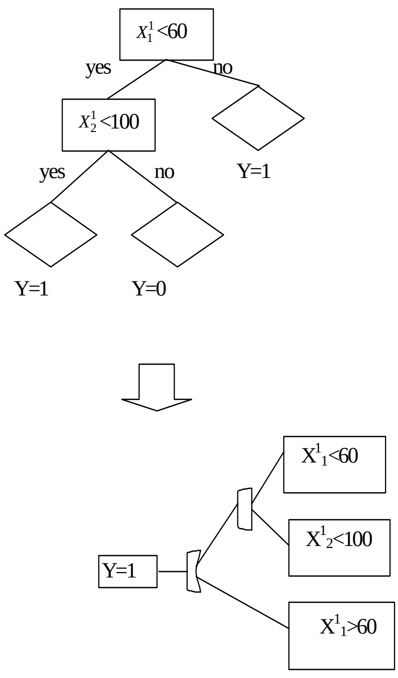

Fig.6. Formation of FT first layer from decision tree

ordinary event).

In T^*, all features are measured for the same moment of time in the past. This tree is used then for the formation of first layer of FT. (Note that in other variant of the method all trees $T_1, T_2, ..., T_l$ can be used for FT formation).

Consider the following algorithm:

1. Find all statements in paths from root node of T^* to the leaves that correspond to undesirable events;
2. Use these statements as nodes (events) of current level of FT (Fig. 6). The conjunctions of statements are connected with TUE by means of AND gates, and different paths are connected with TUE by means of OR gates. Choose non-terminal nodes of FT;
3. For each non-terminal event E_m of current level of FT form new feature Z_m: Z_m=1, if this event takes place, Z_m=0 otherwise, where m=1,…,M, M is number of non-terminal nodes;
4. For each m construct sub-tree for prediction Z_m on the basis of its background;
5. Use statements of obtained DTFB as nodes (events) for next level of FT (Fig.7);

6. Repeat steps 3-5 until no non-terminal nodes left.

The constructed FT may be corrected and complemented by an expert. The probabilities of events in FT are evaluated as frequencies of correspondent observations in time series.

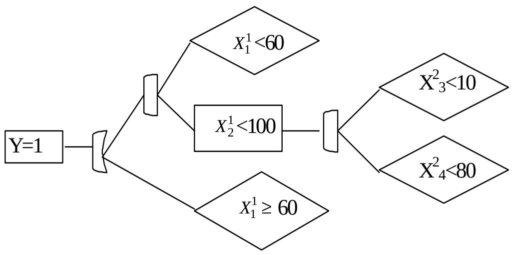

Fig.7 FT next layer formation