Slovak University of Technology in Bratislava
Institute of Information Engineering, Automation, and Mathematics

PROCEEDINGS

17$^{\mathrm{th}}$ International Conference on Process Control 2009

Hotel Baník, Štrbské Pleso, Slovakia, June 9 – 12, 2009

ISBN 978-80-227-3081-5

http://www.kirp.chtf.stuba.sk/pc09

Editors: M. Fikar and M. Kvasnica

Bíla, J., Jura, J.: Qualitative Modeling and Data Mining In the Monitoring of the Selected Ecosystem, Editors: Fikar, M., Kvasnica, M., In *Proceedings of the 17th International Conference on Process Control '09*, Štrbské Pleso, Slovakia, 586–594, 2009.

Full paper online: http://www.kirp.chtf.stuba.sk/pc09/data/abstracts/084.html

QUALITATIVE MODELING AND DATA MINING IN THE MONITORING OF THE SELECTED ECOSYSTEM

J. Bila, J. Jura

Department of Instrumentation and Control Engineering
Faculty of Mechanical Engineering, CTU Prague
Technická 4, 166 07 Praha 6, Czech Republic
Jiri.Bila@fs.cvut.cz, Jakub.Jura@fs.cvut.cz

Abstract: The paper proposes two data mining methods used for modeling and monitoring of the selected ecosystems. The first method consists in the following stages: 1. Formation of monitoring states – the design of a state model - the formation of monitoring super-classes – the design of a rule-based system on the set of monitoring super-classes – the design of an algorithm for the detection of unexpected situations. The second method has these stages: the construction of the qualitative matrix of state dependencies - the construction of the matroid on the set of states - the extraction of the matroid bases – special operations with matroid bases. The methodology is developed with data support of a real ecosystem in the South Bohemia. The frame structure of the measurement centers network and the associated structure of the system database is introduced. The demonstration of the methods is presented in a database segment acquired from one of the measurement centers.

Keywords: Qualitative Modeling, Database system, monitoring states, state model, data mining, knowledge discoveries on database, Concept lattice, Hasse diagram, matroid, bases.

1 INTRODUCTION

The Trebon region in the South Bohemia is known with its picturesque landscape and with its ponds (including Christmas surprises). Hardly anybody would expected that the Trebon ecosystem is violated by continuous parasitic dehumidifying and dehydrating which attacks the Small Water Cycle. Small Water Cycle (SWC) denotes the behavior of a local ecosystem (e.g., the Trebon region) in which the volume of water that comes in the ecosystem is evaporated and falls back in this system. (SWC is described e.g., in (Kravčik *et al.* 2008)).

However in our mentioned ecosystem is evaporated water quickly brought up in the zone in which does not condense yet and in this height zone is transported outside the ecosystem to distanced mountains where spontaneously condenses in rising air streams. (With regard to enormous volumes of brought vapor, the condensation is very dynamic and the rain is sometimes of downpour form.)

As a basic reason of these facts are usually introduced two phenomena:

- The overheating of the air above landscape surface as a result of the solar radiation phenomena on the extensive reflex areas (corn fields, wasted turf fields, concrete parking places, roofs of towns, etc.).
- The increased transport velocity (with regard to stable SWC - roughly before 100 years) in non condensation zone (with one prevailing direction).

Seemed to be, that the extension of vegetation fields (that makes evaporation slower) and the decrease of reflex areas surface (that will be rather difficult in towns), could be sufficient for improvement the situation.

The proposed paper that belongs to the project (Pokorný *et al.* 2008), deals with another possible ways yet.

The goals of the project (Pokorný *et al*. 2008) are concentrated on the description of the energy flows and of the development of biodiversity in the ecosystem. The ecosystem seems to be stable but interventions in the landscape proceeds for a long time and steadily. The research is oriented especially to micro-meteorology, to the thermodynamics of atmosphere in a close distance to surfaces of different localities (green vegetation, exhausted turf fields, paved surfaces, etc.) and to the SWC, (Kravčik et al. 2008). Data for the monitoring system are obtained from 11 measurement centers which are located in the selected places. Each of the centers measures 13 variables (usually at intervals of 6 minutes). Data are stored, tested and modified in the database system and they are used for a variety of calculations (for example for the calculation of the complete solar radiation and the radiation reflected from Earth surface and the radiation retained in the Earth surface) and for the discovery of novel knowledge.

2 STRUCTURE OF THE DATABASE SYSTEM

The database system (see Fig. 1) has the following components:

- *Central Module*: Allows to entry into the system and it provides the coordination function.
- *Class User Interface* of a central module (TUR_CM): It represents the user's screen with the capability of launching the basic functions of individual blocks.
- *Class of Measuring Stations*: It is a superior class for coordination of the functions of measurement stations.
- *Class Measuring station* (TMS): Includes basic handling operations with data from measurement stations.
- Class *User Interface* of the measuring station (TUR_MS): Represents the user's screen for control the handling operations with data from measurement stations.
- Class *Management* (TControl): The class that provides co-ordination of functions in the whole system.

The model is developed as an application of the Object Modeling Technique (OMT) and with use of the Unified Modeling Language (UML). From the model generated by CASE Rational Rose 98 were used only the basic scripts. Two programming environments are used for database system implementation: MS SQL – for the proper implementation and for the maintenance of the system and the Visual Basic for the planned use of the software system STATISTICS and for the pre-modeling in MS ACCESS.

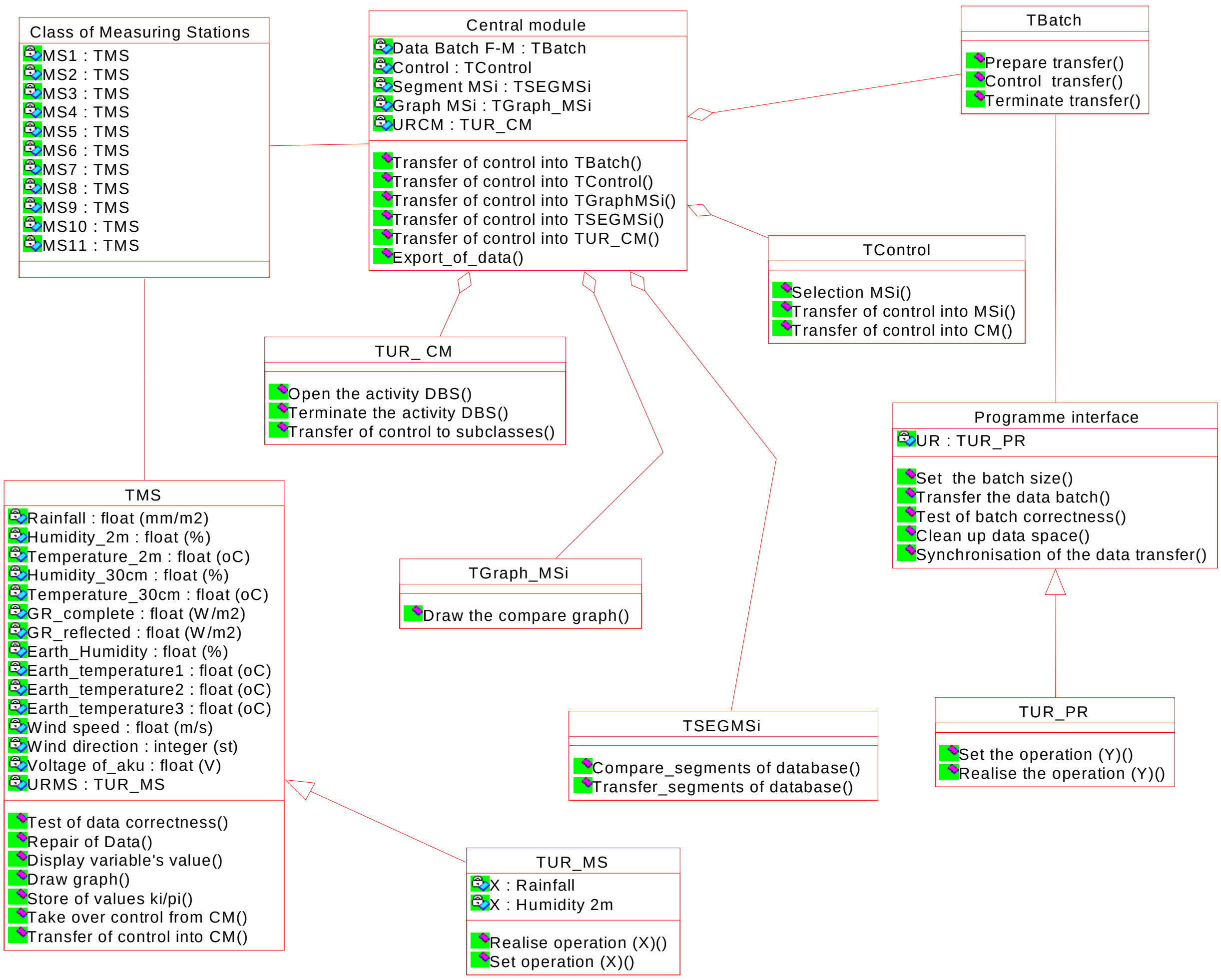

Figure1 – The schema of database system.

3 CONCEPT LATTICES AND MATROIDS

3.1 *Concept Lattices*

Definition 3.1: Triplet $C = (O, I, R)$ (where **O** is a set of objects, **I** is a set of items (attributes) and **R** is a binary relation $R \subseteq O \times I$) is called the Data Mining Context, (Jianliang *et al*. 2001).

The Conceptual Lattice on the Data Mining Context (hereinafter called L) is formed by a natural grouping and associating objects with items and their values.

Each element of lattice L derived from the context C is a pair $\langle X, Y \rangle$, where X is a set of objects $X \subseteq O$ and Y is a set of items $Y \subseteq I$. The following conditions induced by the relation R holds for each pair $\langle X, Y \rangle$:

$$X = \{x \in O \mid \forall y \in Y, x\ R\ y\}, \tag{1}$$

$$Y = \{y \in I \mid \forall x \in X, x\ R\ y\}. \tag{2}$$

X is the largest set of objects and this set is described by the attributes from the Y and Y is the largest set of items common for objects of X. The Hasse diagram can be constructed by the use the tree ordering "<" (which is expressed by the lattice):

- Edge from H1 to H2 exist if H1 < H2 and for none element H3 holds condition H1 < H3 < H2.
- H1 is an antecedent of element H2 (H2 is the descendant of the element H1).
- A pair of $\langle X, Y \rangle$ is a node in Hasse diagram.

Hasse diagram expresses the relationship of *"generalisation / specifications"* between conceptual nodes. Hasse diagram is an acyclic graph with one additional condition: each pair of nodes has one common nearest antecedent and one descendant.

Example 1: Let us consider the context of data mining $C = (\{A_0, A_1, A_2, A_3, A_4\}, \{1, 2, 3, 4, 5, 6\}, R)$, where R is the relation represented in the table $\mathbf{M_G}$ in Figure 2. Hasse diagram for context *C* is Figure 3.

M_G	**1**	**2**	**3**	**4**	**5**	**6**
$\mathbf{A_0}$	**1**	**1**	**1**	**1**	**1**	**1**
$\mathbf{A_1}$	**1**		**1**		**1**	
$\mathbf{A_2}$		**1**				**1**
$\mathbf{A_3}$	**1**	**1**	**1**	**1**		
$\mathbf{A_4}$	**1**			**1**	**1**	**1**

Figure 2 – Relation R from *Example 1.*

The extraction of rules from Hasse diagram is very simple and is described, e.g., in (Jianliang et al. 2001). However - the extracted rules do not cover the semantic content of the selected fragment of the database in needed details. From this reason the table with rules is completed by quantities of Support (**Supp**) and Confidence (**Conf**), computed for each rule. One of computation methods is introduced, e.g., in (Bila and Jura 2007), and here are described only basic steps:

There is considered the set of situations S and the set of monitoring classes A (both sets are finite sets). Association rule is the expression $A_i \Rightarrow A_j$, where $A_i, A_j \in A$, $A_i, A_j \neq \emptyset$, $A_i \cap A_j = \emptyset$.

The rule $A_i \Rightarrow A_j$ means "each situation s, which is indicated (**Ict**) by monitoring class Ai (**Ict**(s, A_i)), is also indicated by class A_j, (**Ict**(s, A_j))".

Support and Confidence of the rule $A_i \Rightarrow A_j$ are expressed as follows:

$$\mathrm{supp}(A_i, S) = (\#(s \in S \mid \mathbf{Ict}(s, A_i)) / (\#(\mathbf{S})), \tag{3}$$

$$\mathbf{Supp}\ (A_i \Rightarrow A_j, \mathbf{S}) = \mathrm{supp}(A_i \cup A_j, \mathbf{S}), \tag{4}$$

$$\mathbf{Conf}\ (A_i \Rightarrow A_j, \mathbf{S}) = \mathbf{Supp}\ (A_i \Rightarrow A_j, \mathbf{S}) / \mathrm{supp}(A_i). \tag{5}$$

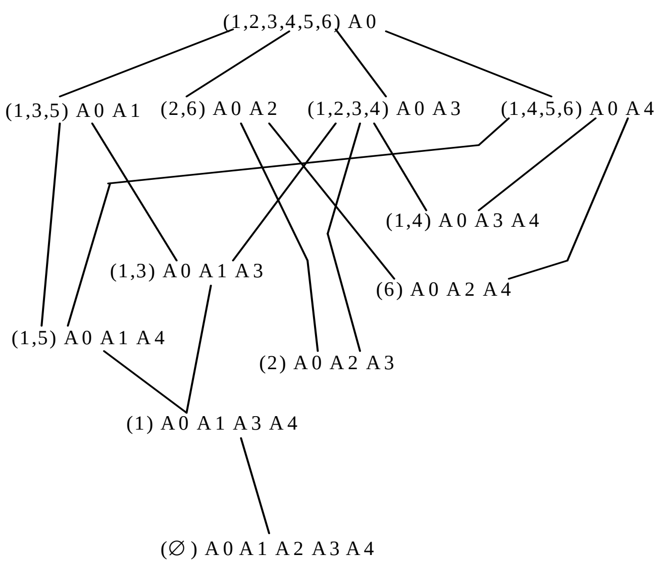

Figure3 – Hasse diagram constructed from matrix $\mathbf{M_G}$.

The system of rules constructed from Hasse diagram (Figure 3), with the calculated values of **Supp** and **Conf**, is introduced in Table 1. (From a few possible rules (e.g., $A_1 \Rightarrow A_2$ or $A_2 \Rightarrow A_1$) are introduced rules with higher quantities of **Conf**.)

Table 1

Rule	Rule r_i	Supp(r_i)	Conf (r_i)
1	$A_1 \Rightarrow A_4$	0.33	0.66
2	$A_1 \Rightarrow A_3$	0.33	0.66
3	$A_2 \Rightarrow A_4$	0.165	0.5
4	$A_2 \Rightarrow A_3$	0.165	0.5
5	$A_3 \Rightarrow A_4$	0.33	0.5
6	$A_1\, A_4 \Rightarrow A_3$	0.165	1
7	$A_1\, A_3 \Rightarrow A_4$	0.165	1

3.2 *Matroids*

Matroid, e.g., in (Oxley 2001) is introduced in this paper as the structure:

$$\boldsymbol{M} = \langle \mathbf{X}, \mathbf{IND}, \{N_1, N_2,..., Nn\}\rangle = \langle \mathbf{X}, \mathbf{IND}, \boldsymbol{\Gamma}\rangle, \quad (6)$$

where **IND** is the relation of Independence and N1, N2, ... , Nn are independent sets. Relation **IND** is a binary relation. The independent set contains elements that each two are in the relation **IND**. Relation **IND** is usually constructed with help of binary dependence relation (**DNT**).

The following definition of the relation **DNT** were used for the construction of matroid:

Definition 3.2.1: Knowledge elements x_1, x_2 are Dependent (**DNT**(x_1, x_2)) if one (or two, three or four) of the following conditions holds (hold):

(i) Knowledge elements x_1, x_2 contribute to the solution of the *given problem* in the same or similar ways, means and principles. (Expert criterion.)

(ii) There are changes (variations) of x_1 which are associated (in an actual time-space of solution synthesis for a given problem) with changes (variations) of x_2 (or vv.).

(iii) The application of x_1 in a problem solution implies the application of x_2 (in this problem solution) (or vv).

(iv) Application of x_1 in the problem solution excludes the application of x_2 (in the problem solution) (and vv).

Note: 3.2.1: The solved problem (in (i)) is in our case, e.g., ”the partition of energy in states of the ecosystem” or ”the stabilization of sufficient humidity in the interaction zone (0-200 m) above landscape surface of the ecosystem”, etc.

Respecting Definition 3.2.1 the following theorem holds.

Theorem 3.2.1:

$$(\text{Not } ((\mathbf{DNT}(x_1, x_2)) \Leftrightarrow (\mathbf{IND}(x_1, x_2)). \quad (7)$$

The following three conditions holds universally, e.g. in (Oxley 2001).

(I1) $\varnothing \in \Gamma$;

(I2) If $N_1 \in \Gamma$ and $N_2 \subseteq N_1$, then $N_2 \in \Gamma$.

(I3) N_p, $N_q \in \Gamma$ and holds $\#(N_p) < \#(N_q)$. Then exists $x_e \in N_q - N_p$ and holds $(N_p \cup x_e) \in \Gamma$. ♦

Note 3.2.2: Condition (I3) is called the *independence augmentation* axiom.

A maximal independent set is called the **Basis** ($\aleph$) of the matroid.

Theorem 3.2.2.: (on the number of elements of Bases): If a matroid has more than one Basis, all the bases have the same number of elements. ♦

Theorem 3.2.3.: Let β(X) be a system of subsets of set X. Then the following conditions are equivalent:

(i) For every non-negative function h: X --->R_+, *the greedy algorithm* (Oxley 2001) finds set B $\in \beta$(X), for which H(B) = Σ h(x), (for all x $\in$ B) has its maximal quantity.

(ii) There exists such a matroid ***M*** on X that the maximal elements from β(X) (i.e. the elements as B) represent the system of all bases of matroid ***M***. ♦

Theorem 3.2.4: Set B discovered by *the greedy algorithm* on matroid ***M*** with regard to the real function h: X --->R_+ and with regard to the maximum of function H(B) = Σ h(x), (for x $\in$ B) represents one of the bases of matroid ***M***. ♦

4 STATE DESCRIPTION OF THE SELECTED ECOSYSTEM, MONITORING SUPER-CLASSES AND THE DETECTION OF THE UNEXPECTED SITUATIONS

The qualitative description, that is now introduced, is based on the assumption that we are able to concentrate effectively the information about the ecosystem behavior into states and into the transition between these states. The state is understood from one side as a language denotation of some time-space fragment form the life of the ecosystem, and from the other side it is understood as some typical result of measurement in centers M1 - M11. (In other words – each ”language” state corresponds to a set of sequences of 13 components from 11 measurement centers. Illustration segments of vectors from database are not introduced.)

4.1. *Monitoring states*

The following monitoring states were designed for the representation of the ecosystem behavior:

Air humidity and water

S1… *Low local humidity,*
S2… *Middle local humidity,*
S3… *High local humidity,*
S4… *Local fog,*
S5… *Region fog (covers the area larger than 20 km^2),*
S6… *High volume of water soaked in the soil,*
S7… *Local floods,*
S8… *Violence of the small water cycle (SWC),*

Weather:

S9… *Rain,*
S10… *Snow,*
S11… *Long time local dry atmosphere, (arid soil),*
S12… *Semi-clear weather,*
S13… *Very cloudy weather and overcast,*
S14… *Strong wind,*
S15… *Storm,*

Evaporation:

S16… *High evaporation (no water goes back down the surface),*
S17… *Middle evaporation – (a part of evaporated water goes back to places of evaporation),*
S18… *Low evaporation,*
S19… *Violated evapotranspiration.*

The semantics of symbols in the matrix of transitions (Table 2) is the following one:

- Symbol ”1”denotes a possible transition or a possible simultaneous existence of states (e.g., (S2→S4) OR (S2 AND S4)). The certainty (uncertainty) of the transition would be represented by a membership function and would cover not only the proper fact of the transition but also some external circumstances as ”where” and ”in what time interval”, etc. The values of certainties are not introduced in this paper.
- Symbol ”0” denotes non existence (high uncertainty) of the transition.
- The symbol ”*” denotes a non relevancy of the question about the transition between considered states.

The qualitative model of the ecosystem behavior would be constructed as a *state diagram* associated with the transition matrix. This is rather simple operation, however it leads to an incomprehensible result. This step is dropped in this paper.

Table 2 – Qualitative matrix of transitions

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19
S1	1	1	0	0	0	0	*	*	*	*	1	*	*	*	*	1	1	1	1
S2	1	1	1	1	1	0	*	0	1	1	0	1	1	*	0	1	1	1	0
S3	0	1	1	1	1	1	1	0	*	*	0	*	*	*	*	1	1	0	0
S4	0	0	1	1	1	*	*	*	1	1	0	1	1	*	0	0	1	1	0
S5	0	1	1	1	1	*	*	0	1	1	0	1	1	*	1	0	0	1	0
S6	0	1	1	1	0	1	1	0	*	*	0	1	1	*	*	0	1	1	0
S7	0	1	1	1	0	1	1	0	1	0	0	1	1	1	1	0	0	1	0
S8	1	0	0	0	0	0	1	1	0	0	1	0	0	1	0	1	0	0	1
S9	0	1	1	1	1	1	1	0	1	1	0	1	0	1	1	0	0	1	0
S10	0	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0
S11	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1
S12	*	*	*	0	0	1	*	*	1	1	0	1	1	0	0	0	1	1	0
S13	*	*	*	0	0	*	*	0	0	0		1	1		*	0	1	1	0
S14	*	*	*	0	0	*	*	*	*	*	0	1	1	1	1	1	0	0	0
S15	1	1	1	1	0	1	1	0	1	1	0	1	0	1	1	0	1	1	0
S16	1	1	0	0	0	0	0	1	0	0	1	1	0	*	1	1	1	0	1
S17	1	1	0	1	1	1	0	0	1	1	0	1	0	*	1	0	1	1	0
S18	1	0	0	0	0	0	*	1	0	0	1	1	0	*	1	0	0	1	0
S19	1	0	0	0	0	0	0	1	0	0	1	*	*	*	*	1	0	0	1

4.2 *Monitoring super-classes and the detection of unexpected situations*

Monitoring super-classes are introduced as sub-diagrams of transitions (extracted from the complete state diagram) that have a special semantic content in the ecosystem life. Analyzing the system database, we introduced the following supper-classes:

A1. Moist extreme of SWC. (It does not lead directly to violence of SWC but it necessitates stabilization actions (natural or artificial) into ecosystem.)

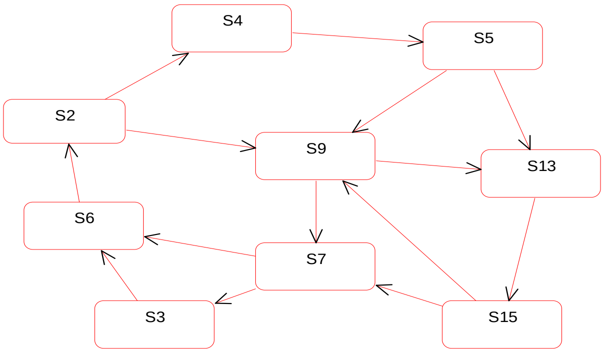

Figure 4 – Moist extreme of SWC.

A2. Dry extreme of SWC. (This situation leads directly to violence of SWC and evapotranspiration.)

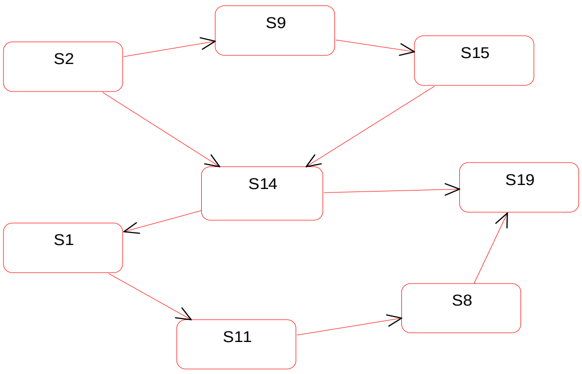

Figure 5 – Dry extreme of SWC.

A3. Stabilizing behavior of ecosystem (auto-regulation) directed to the recovery of SWC and standard behavior of ecosystem.

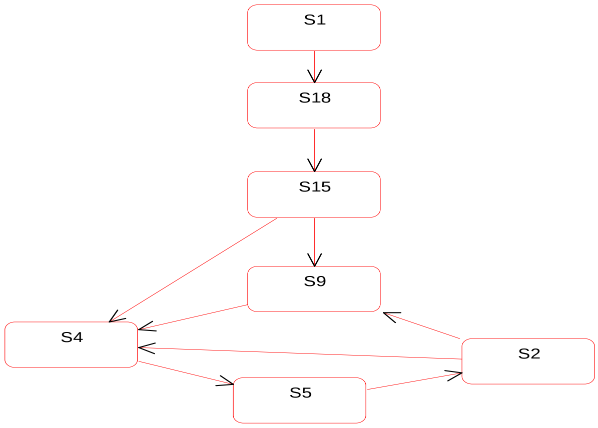

Figure 6 – Stabilizing behavior of the ecosystem.

As it seen from Figs. 4 - 7, some diagrams have common sequences of states. These circumstances are from one side natural and they come from the physical and phenomenological characteristics of ecosystem behavior. From the other side they are consequences of the description style and of the reduction of information about the ecosystem in the monitoring states and super-classes.

The fact that the sets of sequences of monitoring states in super-classes are non disjoint implies the difficulties in identifications of the actual behavior of ecosystem. From the other side it brings the additional information about the ecosystem, the information available by means of *data mining methods*.

A4. Standard behavior of ecosystem with SWC.

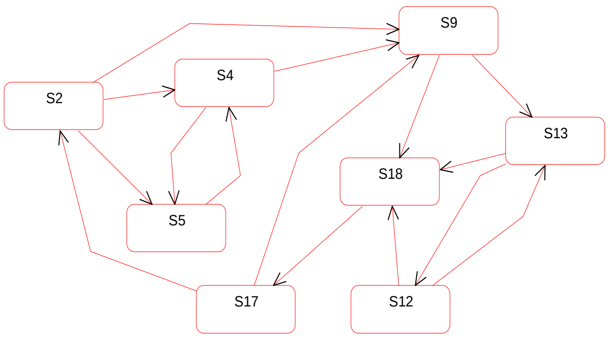

Figure 7 – Standard behavior of ecosystem with SWC.

One of such ways by data mining methods has partially been described in Section 3.1. The result of analysis of common state sequences in super-classes diagrams is seen in matrix $\mathbf{M}_G$ in Fig.2. This matrix is transformed into Hasse diagram (in Fig.3) and then into the system of rules – Table 1. These rules with their values **Supp** and **Conf** are the additional information about the ecosystem behavior. (Let us remind you the explanation of the rule content: The rule $A_i \Rightarrow A_j$ brings the content ”each situation s, that is indicated (**Ict**) by monitoring super-class A_i (**Ict**(s, A_i)), is also indicated by super-class A_j, (**Ict**(s, A_j))”.

Note 4.2.1: See, please, as an example, ”relatively high” values of **Supp** and **Conf** of rules r_1 a r_2. These rules say that after every Moist extreme of SWC follows the Stabilizing process or the Standard behavior of the ecosystem.

In case that the rules (as in Table 1) are well founded and verified, it is possible to use it for the detection of unexpected situations:

The *unexpected situation* is detected as the declination from behavior of ecosystem determined by system of the rules. The slight declinations are included in Table 1 in values of quantity **Conf**. Let us discuss, e.g., the value **Conf** (r_1). The value 0.66 indicates that not all informations about the ecosystem behavior are known and that may come on other rule from Table 1. Such a situation is not dangerous (it is of the type UX^1, details, e.g., in (Bila and Jura 2007)) and we are able to be prepared for it. Similar reasoning holds for some other rules from Table 1.

However we are interested in unexpected situations of the type UX^3,(details, e.g., in (Bila and Jura 2007)), that are induced by violence of the relation between the carrier of information (in our case – the language of rules) and language expression (the rule).

Respecting Table 1 such dangerous rules are:

$$\{A_1 \Rightarrow A_2, A_2 \Rightarrow A_1, A_3 \Rightarrow A_2, A_3 \Rightarrow A_1, A_4 \Rightarrow A_3\}.$$

These rules indicate some abnormal situation (e.g., that after Moist extreme of SWC will come immediately Dry extreme), which is not considered in Table 1 (and which has never been observed yet). Such rules warn us against the possibility of situations, when we do not know what will be set in. (And even not only that we do not know which of super-classes (from Table 1) will come on but we do not know what will be set in at all.)

Such a statement is very unpleasant and in cases of exacting technological processes their operation is stopped till clarification of reasons and characteristics of the situation. In our case is not possible to set aside the ecosystem and is necessary to prepare measures providing minimization of eventual losses.

Note 4.2.2: For the discussion if six situations {1, 2, 3, 4, 5, 6} in matrix $\mathbf{M_G}$ are sufficient representatives of situations in database with millions vectors of numerical data:

The dimensions of matrix $\mathbf{M_G}$ depends on the number of monitoring super-classes. And the number of super-classes depends on the complexity of modeled ecosystem and on the ability of an expert to concentrate information about the ecosystem to the designed super-classes.

In our case we have 4 active super-classes (A0 is a default super-class) and hence there are 15 possible configurations. Four configurations are disjoint (assigned to only one super-class) and one configuration is wrong in principle (assigned to all super-classes). It is no reasonable to consider spontaneous behavior of A1 and A2. Three configurations from remaining nine are simply considered as unreal. Matrix $\mathbf{M_G}$ is in our case sufficient. (The assertion about the number of columns of matrix $\mathbf{M_G}$ holds universally.)

Table 3 – Qualitative matrix of **DNT** relation

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19
S1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1	1
S2	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1
S3	1	1	1	1	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1
S4	1	1	1	1	1	0	0	1	0	0	1	0	0	1	0	1	1	1	0
S5	1	1	1	1	1	0	0	0	0	0	1	0	1	1	0	1	1	1	0
S6	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	0
S7	0	0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
S8	1	0	0	1	0	1	1	1	1	1	1	0	0	1	0	1	0	1	1
S9	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1
S10	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	0	0	1	0
S11	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1
S12	0	1	1	0	1	0	0	0	0	0	0	1	1	0	0	1	1	1	0
S13	0	0	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1
S14	1	0	0	1	1	0	0	1	0	0	1	0	0	1	1	1	1	1	1
S15	0	0	0	1	0	0	1	1	1	0	1	1	1	1	1	0	1	1	1
S16	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0	1	1	1	1
S17	1	1	1	1	1	1	0	0	1	0	1	0	1	0	0	1	1	1	1
S18	1	1	1	1	1	0	0	0	1	0	1	0	1	0	0	1	1	1	1
S19	1	1	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1	1	1

5 CONSTRUCTION OF MATROID, DISCOVERY OF BASES AND IMPLICATION FROM BASES

Four bases with 5 elements have been discovered in the matroid:

B1 = {S2, S7, S12, S14, S15},

B2 = {S2, S7, S8, S12, S15},

B3 = {S4, S7, S12, S15, S19},

B4 = {S7, S10, S12, S14, S19}.

The operations with bases have now two continuations:

- The extension of the matroid by some new artificial monitoring state - for improvement of the monitoring system function. It means the extension of some of bases (respecting relation **DNT** (**IND**)). (The extension of the matroid is not a trivial operation and usually is realised by the construction of modular cut.)
- The use of consequences of theorems 3.2.3 and 3.2.4 from section 3.

Let us consider a nonnegative real function h: X --->R+ that assigns to states S1 – S19 real numbers. According to theorems 3.2.3 and 3.2.4 - at least one of the discovered bases Bj concentrates highest value:

$$H(B) = \sum_{S_i \in B_j} h(S_i), \text{ pro } j \in \{1,2,3,4\}. \qquad (8)$$

The further implications from consequences above depends on the interpretation of function ”h”, on her physical (event. non physical) essence, on her position in the monitored ecosystem. Partial results to these questions we find by deeper investigation and data mining in the monitoring system database.

6 DISCUSSION OF RESULTS

As it has been said in the Abstract and Introduction, the paper introduces in the research of nature systems two new methods. The fact that these methods lead to qualitative results induces their further continuation in reasoning about consequences and in the design of monitoring strategies that contribute to the solution of the problem of dehumidifying and dehydrating ecosystem.

Results of Section 4. give the background for the design of the monitoring system and for the detection of unexpected situations. In case that the ecosystem will be in order, it will move in transitions of state diagram and in the transitions between monitoring super-classes. The further research has three directions:

- The accomplishment of the correspondence between data acquired from measurement stations and the monitoring states.
- The verification of concordance of ecosystem behavior and state model under reduction or extension of the set of monitoring states.
- The application of the algorithm for the detection of unexpected situations for both the previous cases.

Discussion of results of section 5 is even more interesting.

At first – what could represent function ”h” ?

According to definition (e.g., in (Oxley 2001)) function ”h” represents some additive quantity. Let us suggest three designs:

a) The values of function ”h” represent the values of physical energy concentrated in states of the ecosystem. Highest energies would be concentrated, in states of bases. Let us consider . e.g., the first and the fourth:

- ”S2 (Middle local humidity), S7 (Local floods), S12 (Semi-clear weather), S14 (Strong wind), S15 (Storm)” or
- ”S4 (Local fog), S10 (Snow), S12 (Semi-clear weather), S15 (Storm), S19 (Violated evapotranspiration)”.

This interpretation corresponds to our knowledge only partially, nevertheless it is a non trivial hypothesis. We should have to consider, if this discovery has at least a local importance or if it is correct only with regard to dehydrating landscape or to stabilization of humidity in the interaction zone. For the verification we have to go back to database of measurement data. However if this interpretation is true it will have general surprising consequences.

b) Function ”h” corresponds to some scalar variable that has influence on stabilization of humidity in the interaction zone above surface of landscape. The investigation of this issue opens the field for basic research.

c) Function ”h” indicates activities of an emergent intelligence, e.g., in (Hendriks-Jansen 1996), concentrated in the ecosystem. Emergent intelligence leads the ecosystem to states that are optimal for it and its environment, however for us are these activities unpredictable by our temporary theoretical means. This issue also belongs to a basic research.

7 CONCLUSIONS

The proposed paper illustrates the application of two data mining methods and their use in ecosystem monitoring. As a data support has been used the data acquired from measurement centers. The paper contributes to explanation of some local climatic phenomena, as it is the violence of small water cycle and sharp extremes of weather (and v.v.). The trends encoded in the proposed qualitative models describe indirectly energetic balances in the local ecosystem, their influence on SWC and the consequences usually assigned to so called ”green-house effect”. The proposed results and hypotheses support conceptual substances of New Water Paradigm, (Kravčik *et al*. 2008).

ACKNOWLEDGEMENT

This article belongs to a research project supported by grant MSMT No 2B06023.

8 REFERENCES

Kravčik, M., Pokorný, J., Kohutiar, J., Kováč M. and Tóth, E. (2008). Water for the Recovery of the Climate. A New Water Paradigm. Typopress-publishing house s.r.o., Košice.

Pokorný, J., Jirka, V., Pechar, L., Bíla, J., Jura, J., Hofreiter, M., Petrová, R., Zicha, J., Kobrzek F. and Mareček, J. (2008). *Partial Report to project 2B 06023*, MSMT, Czech Republic.

Jianliang, Z. Yi. S., Da Ruan and Pengfei, S. (2001). *Interesting Rough Lattice-based Implication Rules Discovery*. (Da Ruan, J.Kacprzyk, M.Fedrizzi (Eds..)), Soft Computing for Risk Evaluation and Management, Heidelberg, Springer-Verlag, pp. 155-169.

Oxley, J. G., (2001). *Matroid Theory*, Oxford Science Publications, Reprinted Edition.

Hendriks-Jansen, H. (1996). *Praise of Interactive Emergence or why Explanations Don’t Have to Wait for Implementations*. In: Proc of 4th Int. Worskshop on the Synthesis and Simulation of Living Systems- Artificial Life IV. (Brooks, R.A and Maes, P. Eds.), The MIT Press, Cambridge, Massachusetts, London, England, pp. 70-79.

Ripl, W. (1995). *Management of Water Cycle and Energy Flows for Ecosystems Control* – The Energy- Transport-Reaction (ETR) Model. Ecological Modelling, No. 78, pp. 61-76.

Bila J. and Jura, J. (2007). *Fuzzy Concepts in the Detection of Unexpected Situations*. Acta Polytechnica. Vol. 47, No.1. pp. 5-8.