Database Mining

Leo Võhandu, *Rein Kuusik*, *Peeter Roosmann*

Tallinn Technical University, Estonia

We all are already used to the fact that databases are the living cycle histories of real life processes. Every user has his view for a given database. Complex data handling systems are synthesis of many differing views. The usage of databases is usually built on SQL interfaces and query languages. The active side is always the end user himself.
Massive build-up of databases after 80-ies has already created situations where hand or queries driven information systems are too much to handle for a usual stakeholder of the system.

The third wave of information processing in 90-ies is bringing in new ideas for that kind of situations. It consists of "database mining". That means "nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (Frawley et al. 1992). It encompasses a number of different technical approaches, such as clustering, data summarization, learning classification rules, deriving dependency (and other) from data, analyzing changes and derivations, and detecting anomalies (Mathews e a 1993).
Discovery of hidden data patterns is important in many fields, but its future prosperity lies in our understanding, especially in GIS field.

Our talk gives an overview about the methods developed at the Institute of Informatics for database mining. The formal basis for those methods is the theory of monotone systems created in our Institute (Mullat, Võhandu, Kuusik).

The main idea consists of covering database layers and data with IF<statement> THEN <statement> type sentences. The truth-value of such sentences can be controlled by a precision parameter. All needed formalisms will be presented in the talk.
Our practical experience shows that the working speed of covering and mining procedure is a very good one.