Discovering Temporal Association Rules: Algorithms, Language and System

Xiaodong Chen
Department of Computing and Mathematics
Manchester Metropolitan University, U.K.
e-mail: x.chen@doc.mmu.ac.uk

Ilias Petrounias
Department of Computation
UMIST, U.K.
e-mail: ilias@co.umist.ac.uk

Abstract

Most previous work on association rule discovery [1] overlooks time components, which are usually attached to transactions in databases. This results in the loss of the chance to discover some meaningful time-related rules. This work addresses temporal issues of association rules with the corresponding algorithms, language and system for discovering temporal association rules.

1. Temporal association rule discovery

A temporal association rule [2] can be considered as a pair <AR, TF>, where AR is an implication of the form X $\Rightarrow$ Y and TF is a temporal feature (e.g., a valid period, a periodicity or a specific calendar) that AR possesses. In the task of temporal association rule discovery, both association rules and temporal features are expected to be extracted from the database. From a practical point of view, it is too expensive to find all possibly hidden temporal association rules from large databases without any limitation, due to the two-dimensional solution space. The methodology adopted in this paper is eventually discovering potential temporal rules by alternatively using special search techniques for various restricted problems in an interactive and iterative process.

Three forms of interesting mining tasks for temporal association rules with certain constraints are identified. These tasks are the discovery of valid time periods during which association rules hold, the discovery of possible periodicities that association rules have, and the discovery of association rules with temporal features. The search techniques and algorithms ([2], [3]) for those individual tasks have been developed and implemented in a temporal mining system.

The results of experiments with synthetic datasets show that many time-related association rules that would have been missed with traditional approaches can be discovered with the techniques and approaches presented in this paper.

2. Temporal mining language and system

The integrated query and mining system (IQMS) is a prototype system with an integrated query and mining interface (IQMI), supporting the discovery of temporal association rules. The kernel of the system is a temporal mining language (TML), integrated with Oracle SQL. By using this system, data miners can explore temporal association rules from databases in an iterative process, as shown in Figure 1. With the query function, the data in any Oracle database can firstly be analysed, according to the business requirements, in order to get some useful information (e.g., summary information about the data) for designing mining tasks. The mining task can then be designed and expressed in the TML language and fulfilled by the ad hoc mining function of the system. The mining results need to be further analysed to judge if the expected knowledge has been found or whether the mining task should be adjusted to make another mining effort.

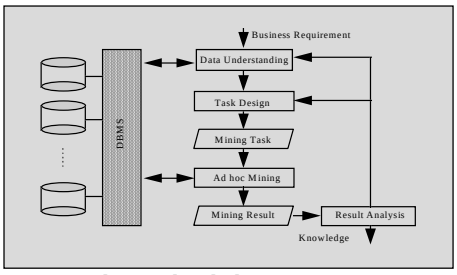

Figure 1: IQMI-based mining process

The significance of such an integrated system has been shown in three aspects in the dynamic data mining efforts. Firstly, *data selection and sampling* for different data mining tasks are easy to achieve with the query function that is integrated in the system. Secondly, *ad-hoc mining* for different application requirements is possible to fulfil with the mining language supported by the system. Finally, all data mining activities can be undertaken in a more flexible data mining process based on IQMI.

3. References

[1] R. Agrawal, T. Imielinski, and A. Swami, Mining Associations between Sets of Items in Massive Databases, *Proc. of SIGMOD'93*, Washington D.C., USA, pp.207-216.

[2] X. Chen, I. Petrounias, and H. Heathfield, Discovering Temporal Association Rules in Temporal Databases, *Proc. of IADT'98*, Berlin, Germany, pp.312-319.

[3] X. Chen and I. Petrounias, Mining Temporal Features in Association Rules, *Proc. of PKDD'99*, Prague, Czech Republic, pp.295-300.