# Discovering Temporal Relation Rules Mining from Interval Data

Jun Wook Lee[1], Yong Joon Lee[2], Hey Kyu Kim[2], Bu Hun Hwang[3], and Keun Ho Ryu[1]

[1] Database Laboratory, School of Electrical and Computer Engineering, Chungbuk National University, Cheongju 361-763, Republic of Korea
{junux,khryu}@dblab.chungbuk.ac.kr
[2] Postal Technology Development Department, Electronics and Telecommunications Research Institute, Korea
{yjl,hkkim}@etri.re.kr
[3] Department of Computer Science, Chonnam National University, Korea, 300, Yongbong-dong,Gwangju, Chonnam, Korea
bhhwang@chonnam.chonnam.ac.kr

**Abstract.** In this paper, we propose a new data mining technique that can address the temporal relation rules of temporal interval data by using Allen's theory. We present two new algorithms for discovering temporal relationships: one is to preprocess an algorithm for the generalization of temporal interval data and to transform timestamp data into temporal interval data; and the other is to use a temporal relation algorithm for mining temporal relation rules and to discover the rules from temporal interval data. This technique can provide more useful knowledge in comparison with other conventional data mining techniques.

## 1 Introduction

Due to the explosion of data volumes in various application areas, development of new types of data mining algorithms which can discover some useful and unknown knowledge has been focused [21]. Emerging data type of recent applications is from such as e-commerce, bioinformatics, GIS, medical care, and so on. Many works [6,7,8] on data mining have focused on the issues of temporal data mining for discovering temporal knowledge from temporal data, such as sequential patterns[5,9,10,11], similar time sequences [13,14], and temporal association rules [1].

An interesting type of rule is temporal rule from data which encompass temporal elements. Temporal rule is a mining technique that extends conventional data mining techniques. Through its extension, it can discover rules of temporal and causal relationships. This technique includes cyclic associations [15] that discover cyclically repeating association rules, that is, calendric associations [16,17] which mean association rules satisfying the temporal patterns expressed in the form of a calendar. In order to find some meaningful relationships from the temporal data, it is necessary to consider the relationships between data. Allen [2] introduces temporal relationships between intervals and operators for reasoning about relations between intervals. Moreover, the applications of Allen's interval

M.H. Shafazand and A M. Tjoa (Eds.): EurAsia-ICT 2002, LNCS 2510, pp. 57–66, 2002.

operators in geographical information system have been attempted in many works [3,4].

Previous studies have developed their research designs only with the data that were stamped with time points rather than time intervals. Basic studies on mining of useful patterns from interval data have been even partially attempted [18,19] recently. However, a lack of attention on that issue may be attributable to the complicated research process.

In this paper, we present a new data mining technique based on Allen's theory. Despite the scalability problem of the theory, the relationships and operators defined can help effectively discovering temporal relation rules from interval data. In order to reduce the processing time of time point data, algorithm uses the preprocessing method. The core of algorithm is discovering temporal relation rules from temporal interval data using temporal relationship operators introduced by Allen. This algorithm extends the AprioriAll[5], a typical sequential pattern algorithm, for discovering temporal relation rules.

The outline of the paper is as follows. Mining problems of temporal relation rules will be explored in section 2 and algorithms for discovering temporal relation rules will be convincingly explicated through a step by step process in section 3. In section 4, we will prove the effectiveness of algorithm through a series of experiments. Finally, further researches will be discussed.

## 2 Mining Problem Definition

### 2.1 Temporal Relation

**Definitions.** Let $e=(E,t)$ be an event with time point by. $E$ stands for an event type and $t$ means the occurrence time of event $e$. A sequence is an ordered list of event according to time. We denote an event sequence S by $<e_1, e_2, ..., e_n>$, where $e_i=(E_i,t_i)$ and $t_i \leq t_{i+1,}$ for each $i=1,..., n-1$. $[t_1, t_n]$ means the period between the first event $e_1$ and the last $e_n$ of $S$.

Also, Let $e'=(E, vs, ve)$ be an event with temporal interval. We call $vs$ the starting point of the temporal interval and $ve$ its end point. They are denoted as $e'.vs$, $e'.ve$. An sequence of this events is expressed as $S'=<e'_1, e'_2, ..., e'_n>$, where $e'_j=(E_j, vs_j, ve_j)$ and $ve_j \leq vs_{j+1}$ for each $j=1, ..., n-1$. Given time granularity U and the criterion time point V, a sequence S can be converted into a sequence S'. The period of S'is defined as [vs1,ven] and can be converted into [1,m], where m≥1.

Let $IE=\{e'_1, e'_2, ..., e'_n\}$ an event set with some temporal interval, where $e'_j=(E_j, vs_j, ve_j)$ for each $j=1, ..., n$. $e'_j.vs$ means the start point of event $e'_j$ with interval and $e'_j.ve$ is the end point. Some event pair $(x, y)$, which is included in event relation $\Omega=\{(x,y) \mid x, y \in IE, x \neq y\}$, has a binary temporal relationship $R(x, y)$. However, $x$ and $y$ are two separate events. A temporal relation is defined as $R(x,y)=\{P(x,y) \mid P \in IO\}$, where $\forall (x,y) \in \Omega$. The set of temporal interval operators is $IO=\{$*before, equal, meets, overlaps,* during$\}$ and P(x, y) is a binary predicate which expresses temporal relationship between x and y. For Example, event x occurs prior to the event of y in case of before(x,y). If $e'_1, e'_2, e'_3$ is the