

Una Propuesta para la Minería de Patrones Temporales Borrosos

Francisco Guil^{1*}, Alfonso Bosch¹, and Roque Marín²

¹ Departamento de Lenguajes y Computación
Universidad de Almería
04120 Almería

{fguil, abosch}@ual.es

² Dept. Ingeniería de la Información y las Comunicaciones
Universidad de Murcia
30071 Espinardo (Murcia)
roque@dif.um.es

Resumen La incorporación de la semántica temporal en las técnicas actuales de minería de datos viene motivada por la necesidad de obtener conocimiento útil en nuevos dominios de aplicación, como los dominios de aplicación dinámicos cuya naturaleza es variable en el tiempo. En este trabajo se proponen las bases de un nuevo algoritmo para la extracción de patrones temporales borrosos a partir de bases de datos relacionales basado en el esquema intertransaccional.

1. Introducción

La minería de datos es un paso esencial en el proceso global de extracción de conocimiento potencialmente útil a partir de un gran volumen de datos, proceso conocido como KDD (Knowledge Discovery in Databases). Consiste en la aplicación de técnicas de análisis de datos y algoritmos que dan como resultado una enumeración particular de estructuras sobre los datos [6]. Dependiendo del tipo de estructura vamos a poder hablar de técnicas de minería de datos globales, que dan como resultado un modelo sobre el conjunto de datos, y técnicas de minería de datos locales, que dan como resultado una enumeración de patrones existentes en los datos [13]. En el caso de las técnicas locales, el ejemplo más importante es la minería de reglas de asociación [1]. El objetivo principal es el de encontrar un mecanismo útil y sencillo de representación del conocimiento para caracterizar las regularidades existentes en grandes volúmenes de datos y usarlas como base para la toma de decisiones dentro de las organizaciones.

Desde la aparición de los primeros algoritmos para resolver el problema de la extracción de reglas de asociación, entre los cuales destacamos el algoritmo *Apriori* presentado por *Agrawal et al.* en [2], se ha desarrollado una gran cantidad de trabajo en varias direcciones, entre las cuales destacamos el desarrollo

* Este trabajo ha sido parcialmente financiado por el MCyT con el proyecto TIC2001-4936-E.

de algoritmos que mejoran en eficiencia al algoritmo *Apriori*, descubrimiento de reglas de asociación generalizadas, multinivel o cuantitativas, algoritmos para la extracción de reglas de asociación borrosas, algoritmos para la extracción de reglas de asociación con restricciones, mecanismos para el mantenimiento de las reglas de asociación extraídas, etc. En [12], *Hongjun Lu* y sus coautores presentan las reglas de asociación intertransaccionales multidimensionales y proponen técnicas para la extracción de este tipo de reglas, técnicas que desde el punto de vista conceptual y algorítmico suponen una generalización de las mencionadas anteriormente.

Muchas de las extensiones propuestas han estado motivadas por la necesidad de extraer conocimiento útil que refleje la naturaleza de distintos dominios de aplicación. Tal es el caso de los dominios de aplicación dinámicos, que poseen la característica de tener una naturaleza variable en el tiempo. Es en este tipo de dominios donde se hace necesaria la incorporación de la semántica temporal en las técnicas de minería de datos, dando lugar al área conocida como minería de datos temporales. Las técnicas de minería de datos temporales son una importante extensión debido a que tienen la capacidad de extraer actividad en vez de estados y, de esta manera, inferir relaciones de proximidad temporal, algunas de las cuales pueden indicar asociaciones causa-efecto [14].

Nuestro objetivo es encontrar un modelo temporal de minería de datos válido en el dominio de aplicación denominado Control Integrado, que consiste en la selección de un tratamiento adecuado para controlar las plagas o enfermedades en cultivos agrícolas. Tras realizar un estudio comparativo de los modelos de minería de datos existentes y ver la inaplicabilidad de dichos modelos en este dominio de aplicación, se propone el desarrollo de un algoritmo de extracción de patrones temporales borrosos a partir de bases de datos relacionales, el cual tiene como base conceptual al modelo intertransaccional. En este trabajo se presentan las bases de desarrollo de este nuevo algoritmo de minería de datos temporales.

La organización del artículo es la siguiente. En la Sección 2 se realiza una breve descripción del problema. En la Sección 3 presentamos el algoritmo de extracción de secuencias a partir de bases de datos relacionales. En la Sección 4 se muestra el proceso de evaluación del algoritmo y se presentan los resultados obtenidos. Y, por último, en la Sección 5 se recogen las conclusiones y se plantean los trabajos futuros.

2. Descripción del Problema

Nuestro objetivo es encontrar un modelo temporal de minería de datos válido en el dominio de aplicación denominado Control Integrado. En este dominio, uno de los problemas principales con el que se encuentran los expertos es la selección del mejor producto (biológico o químico) para controlar una determinada plaga o enfermedad que está afectando a un cultivo. La selección del producto ha de hacerse de acuerdo con los requerimientos de la sociedad, de protección del medioambiente y económicos. Muchos son los factores que han de tenerse en cuenta en la selección de un producto, como por ejemplo, los datos climáticos,

estados fenológicos del cultivo, estado y situación de la plaga o enfermedad, etc. Otro aspecto clave en la selección es que hay que tener en cuenta la incompatibilidad que existe entre ambos tipos de productos. En otras palabras, los productos químicos pueden dañar a los productos biológicos si se aplican dentro de un intervalo de tiempo no permitido. Además, entre algunos productos químicos también existen incompatibilidades, y su aplicación simultánea puede producir efectos indeseables en el cultivo. En definitiva, la elección de los productos se hará teniendo en cuenta las posibles incompatibilidades que existen entre ellos, que pueden verse como restricciones temporales que deben satisfacerse.

Con la idea de asesorar al técnico agrícola en su decisión, se está trabajando en el desarrollo de un sistema de ayuda a la decisión basado en agentes inteligentes, cuyo núcleo es un sistema basado en conocimiento denominado SAEPI [5]. Este sistema experto toma la decisión de si es necesaria o no la aplicación de un determinado tratamiento bajo ciertas circunstancias. El siguiente paso es la selección del mejor plan de tratamiento, tarea de la que se encarga un modelo de decisión multicriterio. Usualmente, este tratamiento consta de múltiples acciones entre las cuales existen restricciones temporales. Además, se ha de tener en cuenta el histórico del cultivo y el plan de cultivo general. El siguiente paso es la generación de un plan de trabajo y para ello se hace uso de un modelo de redes de restricciones temporales borrosas disyuntivas [4]. Mediante este modelo se va a asegurar la aplicabilidad y consistencia del plan de tratamiento, dando como resultado un *schedule*.

Actualmente se está trabajando en el desarrollo de un modelo de minería de datos basado en restricciones temporales para su posterior integración en el sistema de ayuda a la decisión. Los patrones temporales extraídos a partir de la base de datos histórica de tratamientos tienen una doble finalidad. Por una parte, el conocimiento extraído va a permitir validar y complementar la base de conocimiento del sistema experto. Por otra parte, a partir de los patrones temporales extraídos se va a poder obtener un conjunto de restricciones temporales válidas para los planes de tratamiento. La extracción de las restricciones temporales a partir de bases de datos es una tarea muy interesante debido a que este tipo de conocimiento es muy difícil de manejar por parte de los expertos del dominio.

La extracción de patrones basados en restricciones temporales es un problema poco tratado en el área de la minería de datos temporales, área activa en la cual podemos encontrar una extensa y variada bibliografía. En [14], los autores realizan un estudio sobre los distintos métodos y paradigmas de descubrimiento de conocimiento temporal encontrados en la literatura. En este trabajo podemos ver una taxonomía de conceptos de minería temporal y una organización de los métodos en base a la taxonomía propuesta. Centrándonos en el paradigma de descubrimiento de conocimiento temporal basado en *Apriori*, básicamente podemos hablar de tres grupos pertenecientes al área de minería de datos temporales, como son las series temporales, minería de secuencias y minería de reglas de asociación temporales. En el primer caso, el objeto temporal es un valor escalar, comparable con otros objetos temporales del mismo tipo. En el segundo y ter-

cer caso, el objeto temporal está asociado con el concepto de evento. Aunque los eventos están ordenados en el tiempo, su contenido no va a tener un orden asociado y, por lo tanto, no van a poder ser comparados para poder identificar tendencias. Esto hace que el espacio de búsqueda sea mayor y que la minería de secuencias o de reglas de asociación temporales sea un problema más complejo. La diferencia básica entre la minería de secuencias y la minería de reglas de asociación temporales va a estar en el concepto de orden. En el caso de la minería de reglas de asociación temporales, el objetivo es el de encontrar un conjunto de asociaciones junto con el intervalo temporal de validez a partir de un conjunto de datos. En el caso de la minería de secuencias, el problema consiste en obtener una enumeración de secuencias frecuentes (patrones secuenciales) a partir de un conjunto ordenado de secuencias. Una secuencia es un conjunto ordenado de eventos. Fue en [3] donde se introdujo por primera vez el problema de la minería de patrones secuenciales y, a partir de este trabajo, se ha desarrollado una gran cantidad de trabajo en varias direcciones, entre los cuales destacamos el desarrollo de algoritmos eficientes para la obtención de patrones secuenciales y el algoritmo diseñado para la obtención de episodios a partir de secuencias de eventos. En [9], los autores presentan una formulación universal de patrón secuencial que generaliza a los trabajos anteriores.

El problema que se plantea es diferente en naturaleza a los anteriormente citados. En un principio podríamos haber planteado la minería de secuencias como una alternativa viable, aunque la naturaleza variable del dominio de aplicación hace que descartemos este modelo. La variabilidad viene determinada por la inexistencia de un orden predeterminado de aparición de los hechos o eventos. Así pues, es necesaria la incorporación de un mecanismo de representación y razonamiento temporal basada en restricciones temporales, como es el modelo de redes de restricciones temporales borrosas, modelo en el cual se van a basar los patrones temporales que se pretenden obtener.

Debido a que la extracción de patrones temporales borrosos es un problema complejo, el algoritmo propuesto utiliza la técnica Divide y Vencerás, dividiendo el problema en dos subproblemas más tratables, tal y como se indica a continuación:

Algoritmo de Extracción de Patrones Temporales Borrosos

```
algoritmo ExtraccionPTB (D, sop_min, w)
  \\ D: conjunto de datos de entrada
  \\ sop_min: soporte mínimo
  \\ w: longitud de la ventana temporal
comienzo
  Paso1: Sec = Extraer_Secuencias(D, sop_min, w);
  Paso2: PTB = Obtener_Patrones(Sec);
fin
```

En un primer paso, el algoritmo extrae un conjunto de secuencias (*Sec*) a partir de un conjunto de datos, como puede ser una tabla o una consulta realizada sobre varias tablas relacionadas pertenecientes a una base de datos relacional. En

un segundo paso, el algoritmo obtendrá una enumeración de patrones temporales borrosos (*PTB*) a partir del conjunto de secuencias obtenidas en el primer paso. En este trabajo nos vamos a centrar en el algoritmo de extracción de secuencias, el cual se detalla en la siguiente sección.

3. Algoritmo de Extracción de Secuencias

Fuera del ámbito de la minería de datos temporales en [12] se presenta un nuevo algoritmo para la obtención de reglas de asociación multidimensionales intertransaccionales. Aunque el planteamiento es multidimensional, tanto la implementación como el estudio de casos de prueba se hizo para la versión unidimensional, donde la dimensión era el tiempo. El objetivo seguido por los autores es el de encontrar reglas de asociación que reflejen asociaciones frecuentes entre elementos que pertenezcan a distintas transacciones y no asociaciones entre elementos pertenecientes a la misma transacción, como ocurre con la minería de datos clásica o intratransaccional. La idea principal es la de obtener un conjunto de reglas que muestren de forma explícita el contexto en el que aparecen las asociaciones. Los algoritmos de minería de datos tradicionales ignoran la información contextual centrándose en la extracción de asociaciones que suceden dentro de la misma transacción, como pueden ser los productos que han sido comprados por el mismo cliente, el conjunto de eventos que han sucedido en el mismo instante de tiempo, etc. Es por esto por lo que reciben el nombre de algoritmos de minería de datos intratransaccionales. El tipo de conocimiento que se puede expresar con las reglas de asociación clásicas es del tipo "el 80% de los clientes que compran pañales también compran cerveza *durante la misma visita* al supermercado", o "existe una probabilidad alta de que en el *mismo día* en el que se sube el precio del producto A, el producto B se ve sometido a una bajada de precios". Sin embargo, las reglas de asociación intertransaccionales no solo muestran las asociaciones que ocurren a nivel de transacción, sino que también muestran las asociaciones entre elementos pertenecientes a distintas transacciones a lo largo de ciertas dimensiones. Un ejemplo de regla de asociación intertransaccional puede ser "en el 90% de los casos donde se ha tenido que aplicar el producto A para combatir la plaga P, *dos meses después* se ha tenido que aplicar el producto B". En este ejemplo, la dimensión temporal es la que constituye el contexto de la regla extraída. Veamos la diferencia entre

Cuadro 1. Ejemplo de base de datos transaccional

T_{id}	Fecha	Productos
t_1	$date_1$	a, b
t_2	$date_2$	c, d
t_1	$date_3$	a, b
t_2	$date_4$	c, d

ambas técnicas a nivel de patrones (conjunto de elementos frecuentes o secuen-

cias, según el caso) extraídos. Consideremos la base de datos transaccional típica de un supermercado que se muestra en el Cuadro 1. El conjunto de atributos está formado por el número de transacción, el día en el que se ha producido cada transacción y los productos que se han comprado en ese día. Si aplicamos un algoritmo de minería de datos (por ejemplo, el algoritmo *Apriori* [2]) con un valor del 50% para el parámetro soporte mínimo, obtendremos el siguiente conjunto de patrones frecuentes:

$$\{a\}, \{b\}, \{c\}, \{d\} \{a, b\}, \{c, d\}$$

Sin embargo, si observamos detenidamente la base de datos y tenemos en cuenta la dimensión temporal, vemos que el producto *c* se compra siempre un día después de la compra del producto *a*. Lo mismo ocurre con el producto *d* y así con el resto de productos. Esta es la idea intuitiva del esquema intertransaccional: buscar asociaciones a lo largo de la dimensión (en este caso, temporal). Si aplicamos un algoritmo de minería de datos intertransaccional con un valor del 50% para el parámetro soporte mínimo, obtendremos el siguiente conjunto de patrones frecuentes:

$$\{a_0\}, \{b_0\}, \{c_0\}, \{d_0\}, \{a_0, b_0\}, \{c_0, d_0\}, \{a_0, c_1\}, \{a_0, d_1\}, \{b_0, c_1\}, \{b_0, d_1\}, \\ \{a_0, b_0, c_1\}, \{a_0, b_0, d_1\}, \{a_0, c_1, d_1\}, \{b_0, c_1, d_1\}, \{a_0, b_0, c_1, d_1\}$$

En este caso, un patrón está formado por un conjunto de pares (item, unidad temporal) denotados como *item_{ut}*. La unidad temporal refleja la distancia con respecto a la unidad temporal base. Como podemos observar, el conjunto de patrones intratransaccionales es un subconjunto del obtenido aplicando el algoritmo intertransaccional. De hecho, desde un punto de vista conceptual y algorítmico, el problema de extraer asociaciones intratransaccionales puede ser tratado como un caso particular del esquema intertransaccional. Las asociaciones intertransaccionales son más generales tanto semántica como formalmente, son más expresivas y poseen mayor capacidad de predicción. Por otra parte, este esquema plantea mayores desafíos en términos de eficiencia computacional.

El primer algoritmo, denominado *E-Apriori* (Extended Apriori), fue diseñado como una extensión del conocido algoritmo Apriori y presentado en [12,11] junto a una versión mejorada denominada *EH-Apriori* (Extended Hash-based Apriori extendido). En [15,16] se presenta un algoritmo eficiente diseñado específicamente para la extracción de este tipo de reglas de asociación denominado FITI (First Intra, Then Inter). Todos estos algoritmos tienen en común la utilización de varias estructuras intermedias para la obtención de asociaciones. Debido a que el problema de la extracción de patrones temporales es un problema de computación intensiva, es necesaria la implementación de algoritmos que permitan reducir al máximo los recursos computacionales necesarios, tanto el tiempo de ejecución como el espacio de almacenamiento consumido. Siguiendo la filosofía intertransaccional y teniendo en cuenta los requerimientos de tiempo y espacio, hemos desarrollado un algoritmo que utiliza una única estructura de datos, conocida con el nombre de *Árbol de enumeración de subconjuntos extendido* o *Árbol lexicográfico extendido*. Esta estructura de datos fue utilizada por primera

vez en el ámbito de la minería de datos no temporal por *R. J. Bayardo* [10]. En la figura 1 podemos ver el árbol de enumeración de subconjuntos extendido correspondiente a la base de datos que aparece en la Tabla 1.

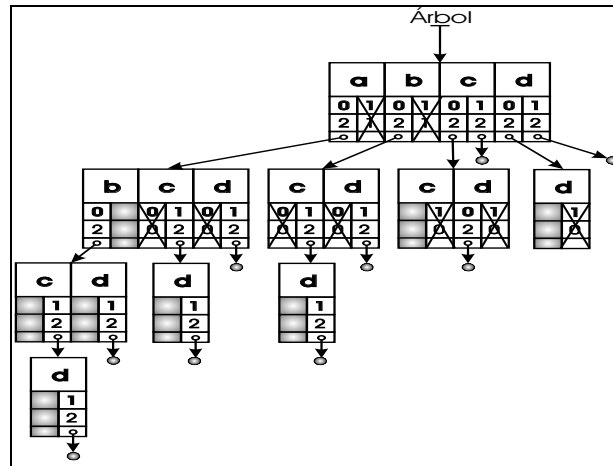


Figura 1. Ejemplo de árbol de enumeración de subconjuntos extendido

En [7,8] se extiende el concepto del problema y se presenta un esquema para la extracción de reglas de asociación intertransaccionales generalizadas. La extensión está basada en el concepto de expansión de contexto y se propone como técnica para dotar a las reglas de asociación de una mayor flexibilidad y expresividad, dando lugar a un esquema de extracción de conocimiento más completo y general del tipo "es muy probable que *entre dos y tres días después* de la aplicación del producto A, se de lugar la aplicación del producto B". Esta expansión de contexto está basada en el concepto de intervalo determinado por dos valores cuya distancia está delimitada por un parámetro. Pero esta serie de extensiones no aborda el problema del tratamiento impreciso de información, necesario en algunos dominios de aplicación para extraer conocimiento expresado en un lenguaje similar al lenguaje natural, el cual es impreciso por naturaleza. Para ello, en este trabajo proponemos la incorporación de la teoría conjuntos borrosos de *Zadeh* [17] gracias a lo cual vamos a poder extraer conocimiento del tipo "es muy probable que, *aproximadamente una semana después* de la aplicación del producto A, se de lugar a la aplicación del producto B", o "es muy probable que el producto A se aplique un *poco después* de la aplicación del producto B". Tal y como se refleja en los ejemplos, la utilización de etiquetas lingüísticas nos va a permitir que en las reglas de asociación extraídas se expresen relaciones temporales tanto cualitativas como cuantitativas.

Tradicionalmente, los algoritmos propuestos en la bibliografía se basan en obtener asociaciones a partir de bases de datos transaccionales. Aunque existen mecanismos para traducir cualquier tabla de una base de datos relacional a una base de datos transaccional (véase Tabla 2), esto supone un nuevo proceso de

preminería de datos muy costoso. Además, cada vez es menos habitual encontrarse con bases de datos transaccionales fuera del ámbito comercial, siendo los sistemas relacionales los que abarcan gran parte del mercado. Hay que destacar que, por claridad en la exposición, se ha escogido una base de datos transaccional como ejemplo, aunque el algoritmo está diseñado para tratar directamente con bases de datos relacionales.

Cuadro 2. Ejemplo de conversión de BD relacional a BD transaccional

$Fecha_i$	C_1	C_2		$Fecha_i$	Items
0	0	0	$C_1 = 0 \Rightarrow$ a	0	a, b
1	1	1	$C_2 = 0 \Rightarrow$ b	1	c, d
2	0	0	$C_1 = 1 \Rightarrow$ c	2	a, b
3	1	1	$C_2 = 1 \Rightarrow$ d	3	c, d

4. Evaluación Empírica y Resultados

Para la realización de los experimentos se ha contado con una base de datos real de tratamientos agrícolas formada por 20 tablas que contienen atributos numéricos, categóricos y temporales (*timestamp*). A partir de esta base de datos se ha seleccionado un conjunto de datos formado por 7.751 registros.

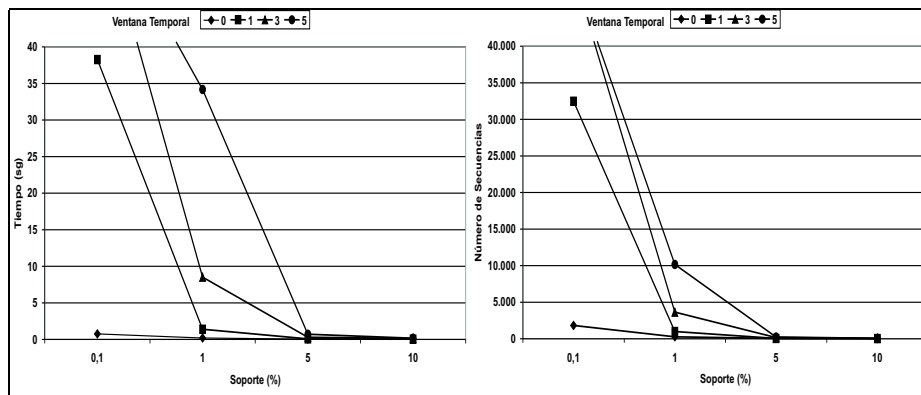


Figura 2. Resultados experimentales

Para estudiar el comportamiento del algoritmo de extracción de secuencias a partir de bases de datos relacionales se han realizado una serie de experimentos, cuyos resultados se muestran en la Figura 2. Los parámetros escogidos han sido soporte y longitud de la ventana temporal. Como podemos observar, tanto el tiempo de ejecución del algoritmo como el número de secuencias obtenidas crece

de forma exponencial conforme vamos disminuyendo el valor de soporte y vamos aumentando el valor de la longitud de la ventana temporal. Para un valor de ventana temporal igual a 0 (que equivale a un proceso de minería no temporal), el crecimiento es prácticamente lineal. Este resultado refuerza lo que ya se indicaba anteriormente. Por un lado tenemos que la incorporación de la semántica temporal en las técnicas de minería de datos es un problema complejo y, por otro lado, es necesario para la obtención de conocimiento útil en dominios de aplicación dinámicos. Podemos ver que, para un valor de soporte bajo, el número de secuencias obtenidas en el caso "no temporal" es prácticamente despreciable si la comparamos con el caso temporal (longitud de la ventana temporal ≥ 1).

5. Conclusiones y Trabajos Futuros

En este trabajo se presentan las bases para el desarrollo de un algoritmo de minería de patrones temporales borrosos a partir de bases de datos relacionales. Al igual que en el algoritmo *A priori*, se ha aplicado el método Divide y Vencerás para el diseño del algoritmo, dividiendo el proceso global en dos subprocesos. En el primer paso, el algoritmo extrae un conjunto de secuencias a partir de un conjunto de datos. En un segundo paso, el algoritmo obtiene un conjunto de patrones temporales borrosos a partir del conjunto de secuencias obtenidas previamente.

Actualmente, la primera parte del algoritmo se encuentra en fase de prueba y validación. En este trabajo se presentan los resultados preliminares que se han obtenido aplicando el algoritmo sobre una base de datos real. Para una correcta evaluación, estamos realizando experimentos con otras bases de datos, como son las bases de datos sintéticas que se utilizan en la mayoría de trabajos existentes en la literatura y bases de datos pertenecientes a otros dominios de aplicación similares.

Se propone como trabajo futuro la implementación de la segunda parte del algoritmo, la cual se encuentra en fase de análisis.

Referencias

1. R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In P. Buneman and S. Jajodia, editors, *Proc. of the ACM SIGMOD Int. Conf. on Management of Data, Washington, D.C.*, pages 207–216. ACM Press, 1993.
2. R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In J. B. Bocca M. Jarke and C. Zaniolo, editors, *Proc. of 20th Int. Conf. on Very Large Data Bases (VLDB'94), Santiago de Chile, Chile*, pages 487–499. Morgan Kaufmann, 1994.
3. R. Agrawal and R. Srikant. Mining sequential patterns. In P. S. Yu and A. L. P. Chen, editors, *Proc. of the 11th Int. Conf. on Data Engineering, Taipei, Taiwan*, pages 3–14. IEEE Computer Society, 1995.
4. A. Bosch, M. Torres, and R. Marín. Reasoning with disjunctive fuzzy temporal constraint networks. In *Proc. of the 9th Int. Symposium on Temporal Representation and Reasoning (TIME'02)*, pages 36–43. IEEE Computer Society, 2002.

5. J. J. Cañadas, I. M. del Águila, A. Bosch, and S. Túnez. An intelligent system for therapy control in a distributed organization. In H. Shafazand and A. M. Tjoa, editors, *Proc of the First EurAsian Conference on Information and Communication Technology (EurAsia-ICT'02), Shiraz, Iran*, volume 2510 of *Lecture Notes in Computer Science*, pages 19–26. Springer, 2002.
6. U. Fayyad, G. Piatetky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AIMagazine*, 17(3):37–54, 1996.
7. L. Feng, T. Dillon, and J. Liu. Inter-transactional association rules for multi-dimensional contexts for prediction and their application to studying meteorological data. *Data & Knowledge Engineering*, 37:85–115, 2001.
8. L. Feng, Q. Li, and A. Wong. Mining inter-transactional association rules: Generalization and empirical evaluation. In Y. Kambayashi, W. Winiwarter, and M. Arikawa, editors, *Proc. of the 3rd Int. Conf. on Data Warehousing and Knowledge Discovery (DaWaK'01), Munich, Germany*, volume 2114, pages 31–40. Springer, 2001.
9. M. V. Joshi, G. Karypis, and V. Kumar. A universal formulation of sequential patterns. In *Proc. of the KDD'2001 Workshop on Temporal Data Mining, 7th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining San Francisco*. ACM, 2001.
10. R. J. Bayardo Jr. Efficiently mining long patterns from databases. In L. M. Haas and A. Tiwary, editors, *Proc. of the ACM SIGMOD Int. Conf. on Management of Data (SIGMOD 1998), Seattle, Washington, USA*, pages 85–93. ACM Press, 1998.
11. H. Lu, L. Feng, and J. Han. Beyond intra-transaction association analysis: Mining multi-dimensional inter-transaction association rules. *ACM Transactions on Information Systems (TOIS)*, 18(4):423–454, 2000.
12. H. Lu, J. Han, and L. Feng. Stock movement and n-dimensional inter-transaction association rules. In *Proc. of the Workshop on Research Issues on Data Mining and Knowledge Discovery (DMKD'98), Seattle, Washington*, pages 12:1–12:7, 1998.
13. H. Mannila. Local and global methods in data mining: Basic techniques and open problems. In P. Widmayer, F. Triguero, R. Morales, M. Hennessey, S. Eidenbenz, and R. Conejo, editors, *In Proc. of the 29th Int. Colloquium on Automata, Languages and Programming (ICALP 2002), Malaga, Spain*, volume 2380 of *Lecture Notes in Computer Science*, pages 57–68. Springer, 2002.
14. J. F. Roddick and M. Spiliopoulou. A survey of temporal knowledge discovery paradigms and methods. *IEEE Transactions on Knowledge and Data Engineering*, 14(4):750–767, 2002.
15. A. K. H. Tung, H. Lu, J. Han, and L. Feng. Breaking the barrier of transactions: Mining inter-transaction association rules. In *Proc. of the 5th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, San Diego, CA, USA*, pages 297–301. ACM Press, 1999.
16. A. K. H. Tung, H. Lu, J. Han, and L. Feng. Efficient mining of intertransaction association rules. *IEEE Transactions on Knowledge and Data Engineering*, 15(1):43–56, 2003.
17. Lotfi A. Zadeh. Fuzzy logic. *IEEE Computer*, 21(4):83–93, 1989.