

메타 규칙 생성에 의한 RFID 데이터 마이닝 기법

김영희^{0*}, 이창열^{**}, 강은영^{*}, 김응모^{*}*성균관대학교 정보통신학부 컴퓨터공학과, **동의대학교 컴퓨터공학과
pink77hee@skku.edu⁰, lcy@deu.ac.kr, eykang@skku.edu, umkim@yurim.skku.ac.kr

A Method Mining RFID Data Using Generation Meta-Rules

Younghee Kim⁰, Changyeol Lee^{**}, Enyoung Kang^{*}, Ungmo Kim^{*}⁰Dept. of Computer Engineering, SungKyunKwan University^{**}Dept. of Computer Engineering, DongEui University

요 약

RFID 데이터에 대한 연관 규칙을 효율적으로 생성하기 위해서는 단일 개념 레벨에서 연관 규칙을 찾는 방법과는 달리 다단계 개념 레벨에서 의미 있는 정보를 발견할 수 있다. 이로부터 연관 규칙을 생성하게 되면 최상위 레벨의 정보를 통해 하위 레벨의 객체 이동 정보나 위치 정보, 상태 정보를 빠르게 획득 가능하다. 또한, 다차원 레벨을 갖는 연관 규칙 마이닝을 수행할 때 메타 규칙의 생성은 제한적이고 유용한 규칙만을 효율적으로 생성 가능하도록 할 수 있다. 따라서, 생성된 메타 규칙을 이용하여 많은 양의 데이터에서 질의를 효과적으로 수행 할 수 있을 뿐만 아니라, 데이터베이스의 저장 효율을 높이고, 객체간의 숨겨진 연관 관계를 발견하는데 있어 효율적인 방법이다.

1. 서 론

RFID는 무선 주파수(RF, Radio Frequency)를 이용하여 대상을 식별할 수 있는 기술로 유통, 물류, 의료, 항공, 자동차 등의 산업분야에 다양하게 응용되고 있으며, 재고 관리 및 공급망 관리 등에서 널리 응용되고 있다. RFID 응용의 주요 과제는 방대한 양의 데이터를 생성하기 때문에 데이터를 어떻게 다루고 분석할 것인지가 중요하다. RFID 데이터는 개별 아이템별로 같은 위치, 간 정보를 공유하며 이동하고 따라서 객체의 변동에 대한 유연성을 제 할 수 있다. 각 개별 아이템들은 아이템, 박스, 팔레트, 컨테이너 형태의 포함관계를 갖고 이동하는 것이 기존의 거래 아이템들과 차이가 있기 때문에 거래에서 발생하는 데이터들에 대하여 연관 규칙을 찾는 방법이 기존의 Apriori 기반 알고리즘 또는 FP-Growth 알고리즘과는 달리 포함 관계정보를 효율적으로 나타낼 수 있는 확장된 연관 규칙 생성이 필요하다. 따라서 본 논문에서는 RFID 데이터에 대한 연관 규칙을 효율적으로 생성하기 위해 단일 개념 레벨에서 연관 규칙을 찾는 방법과는 달리 다단계 개념 레벨에서 RFID 데이터로부터 의미 있는 정보를 발견하고자 하며, 이로부터 연관 규칙을 생성하게 되면 최상위 레벨의 정보를 통해 하위 레벨의 객체 이동 정보나 위치 정보, 상태 정보를 빠르게 획득 가능하다. 이때, RFID 데이터는 최하위 레벨의 데이터 항목의 개수가 많고 상위 레벨의 데이터 항목의 개수가 많지 않은 특징을 가지므로 각 레벨마다 최소지지도의 값을 적당하게 선택하여 가지치기 유용한 정보들을 잃어버릴 위험을 막고 상위 레벨의 지지도가 너무 낮은 경우 중간 결과가 많이 생성되는 것을 막을 수 있다. 또한, 객체간의 연관 관계를 나타낼 때, 계층적인 포함 정보를 나타낼 수 없던 기존의 방법과는 달리 최상위 레벨로부터 최하위 레벨까지 부분적인 순서 집합 및 메타 규칙을 효율적으로 생성할 수 있다. 따라서, 생성된 메타 규칙을 이용하여 많은 양의 데이터에서 질의를 효과적으로 수행 할 수 있을 뿐만 아니라, 포함관계 정보를 통해 객체 간의 보다 더 정확한 정보를 획득 가능하고 객체간의 발생된 이상 현상을 발견하는데 있어 효과적인 방법을 제 하고자 한다. 본 논문의 구성은 다음과 같다. 2장에서는 RFID 데이터와 다단계 연관 규칙 마이닝 기법, 그리고 계층 기반(Hierarchy-based) 마이닝 기

법에 대하여 살펴보고, 3장에서는 본 논문에서 제안하고자 하는 RFID 데이터로부터 메타 규칙 생성을 이용한 마이닝 방법을 기술하고, 4장에서는 제 된 마이닝 기술을 이용하여 생성된 메타 규칙을 통해 메타 질의 생성 결과를 분석하고자 한다. 마지막으로 5장에서는 향후 연구 방향과 결론을 기술한다.

2. 관련 연구

2.1 RFID 데이터

RFID 데이터는[1] 동적으로 관찰(observations)결과를 생성하고, 객체의 위치 및 포함관계가 간에 따라 변화되는 특징을 가지고 있다. 따라서 모든 RFID 관련 데이터는 추적(tracking), 모니터링(monitoring), 어플리케이션 통합을 포함하는 관련 어플리케이션 단계에서 정보를 적당한 데이터 모델로 표현하는 것이 필수적이다. RFID 데이터의 정보 표현 방법[2]은 다음과 같이 나타낼 수 있다. 첫째, (EPC, Location, time) 형태에서 같은 위치에서 일정 간을 머물게 될 때 여러 개의 튜플 생성으로 인한 중복된 정보를 제거하기 위하여 (EPC, location, time_start, time_end)의 형태로 간 범위로 그룹화 한다. 둘째, RFID 데이터는 각 아이템들이 다른 위치로 그룹화 되어 이동되고 머물게 된다. 따라서, 그룹화된 튜플 형태(EPClist, location, time_in, time_out)로 구성하므로 저장 공간을 최소화하고, 리더기로 부터 데이터를 읽을 때 발생하는 불일치성을 제거하므로 데이터 클리닝을 수행할 수 있다. 셋째, 같은 위치로 이동하는 그룹화된 아이템들의 EPClist의 크기를 제거하므로 더 압축된 정보를 얻을 수 있다. 예로, 50개의 EPClist cd들이 두개의 그룹 선반(shelf)와 트럭(truck)으로 이동된다면 두 개의 일반화된 식별자(generalized identifiers, gids)로 나타내어 질의 처리를 효과적으로 처리 할 수 있다. 넷째, 질의는 추상화의 상위 레벨에서 이루어지고 만약 상위 레벨에서 발견된 흥미있는 패턴들과 연관되어 있다면 하위 레벨의 개별 아이템들에서 중요한 요약된 정보를 얻을 수 있다. 따라서, 더 적은 데이터 집합에서 질의를 수행하고 더 자세한 정보가 필요한 경우에 있어서 초기(original)데이터를 폐지 한다. 그림1은 규칙 생성을 위한 RFID 데이터 변환 과정이다.

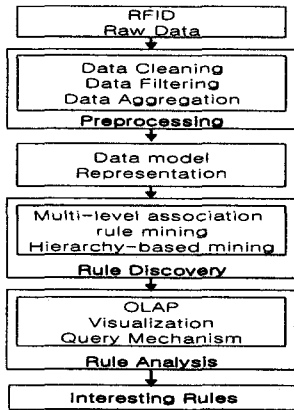


그림 1 RFID 데이터 규칙 생성 마이닝 단계

2.2 다단계 연관 규칙

연관 규칙 알고리즘[3]은 Apriori기반의 알고리즘들로 계층이 없는 항목들간의 지지도와 신뢰도를 기반으로 규칙을 생성하지만, 항목들간에 계층이 존재하는 데이터베이스에서는 다단계 연관 규칙 탐사 방법이 사용된다. 트랜잭션 데이터 집합에서 연관 규칙 마이닝은 다음과 같이 나타낼 수 있다. (1) 트랜잭션 데이터 집합을 T 라 할 때, T 는 $\langle T_i, \{A_p, \dots, A_q\} \rangle$ 이다. 이때 T_i 는 트랜잭션 식별자(Transaction identifier)이고, $A_i \in I$ (for, $i = p, \dots, q$)이다. 그리고 이때 I 는 항목 데이터 집합에서 모든 데이터 항목의 집합을 말한다. (2) 항목 데이터 집합의 서술(Description) D 는 각 항목 I 에서의 서술들을 포함하고, $\langle A_i, description_i \rangle$ 이고 여기서 $A_i \in I$ 이다. 각 항목의 지지도와 신뢰도는 다음과 같다.

- **지지도(Support)** : 패턴(Pattern) A 는 하나의 항목 A_i 또는 결합 항목 집합(set of conjunctive items) $A_1 \wedge \dots \wedge A_j$ 이고, $A_1, \dots, A_j \in I$ 이다. 이때 집합 S 에서 패턴 A 의 지지도(support)는 $\sigma(A/S)$ 이다.
- **신뢰도(Confidence)**: 집합 S 에서 $A \rightarrow B$ 의 신뢰도, $\phi(A \rightarrow B/S)$ 는 집합 S 에서 패턴 A 가 발생, $\sigma(A/S)$, 했을 때 집합 S 에서 패턴 B 가 발생할 가능성, $\sigma(A \wedge B/S)$ 을 말한다.

또한, 다단계 연관 규칙 마이닝[4][5]에서는 서로 다른 레벨마다 최소 지지도(min_support, σ')와 최소 신뢰도(min_confidence, ϕ')를 다르게 하여 생성되는 강한 규칙(strong rule)을 통해 유용한 정보를 얻을 수 있다.

- **Large** : 집합 S 의 레벨 l 에서 패턴 A 의 지지도(support), $\sigma(A/S)$ 가 최소지지도 임계치(min_support threshold, σ')보다 클 때, $(\sigma(A/S) \geq \sigma'(A/S))$
- **High** : 규칙 $(A \rightarrow B/S)$ 의 신뢰도(confidence)가 레벨 l 에서 최소신뢰도 임계치(min_confidence threshold, ϕ')보다 클 때, $(\phi(A \rightarrow B/S) \geq \phi'(A \rightarrow B/S))$

• **Strong** : 집합 S 에서 A 와 B 의 조상(ancestor), 즉 상위 레벨 항목에 대응되는 레벨의 항목들에 대하여 현재 레벨에서 $(A \wedge B/S)$ 가 large이고, $(A \rightarrow B/S)$ 의 신뢰도가 현재 레벨에서 high 일때, 규칙 $(A \rightarrow B/S)$ 는 Strong이다.

즉, 강한 규칙(strong rule)은 현재 레벨에서 $\sigma(A \wedge B/S) \geq \sigma'$ 이고, $\phi(A \rightarrow B/S) \geq \phi'$ 이다.

2.3 계층 기반 마이닝 기법

계층적 개념(hierarchical concept)[6]은 추상화된 다단계 레벨에서 지식 발견에 사용되고, 일반화된 개념으로 하위 레벨의 개념 집합으로부터 상위 레벨로 순차적인 매핑을 정의하고 있다. 이들 계층적 개념은 트리(tree) 형태로 구성되며, 이때 각 노드들은 개념(concept)을 나타낸다. 계층적 개념에서 일반화와 세분화는 숨겨진 데이터간의 관계를 제공하고, 여러 관점으로 데이터의 의미를 보여준다. 일반화(generalization)는 최하위 레벨의 원 데이터(raw data)를 상위 레벨에서 다룰 때 유용하게 사용되며 발견된 패턴들에 대한 이해를 더욱 명적이고 데이터가 갖는 의미의 이해를 쉽게 할 수 있다. 또한, 일반화는 데이터 압축의 장점을 제공하여 데이터의 입출력 동작을 적게 하므로 마이닝을 효과적으로 수행하도록 한다. 이러한 계층적 개념의 4가지 유형은 다음과 같다. 첫째, 스키마 정의 계층(schema-defined hierarchy)은 데이터베이스 스키마에서 속성들 사이에 전체 혹은 부분 순서 즉, 속성들 사이의 의미 관계를 표현하고, 이들 관계들은 데이터간의 메타 데이터 정보를 나타낼 수 있다. 둘째, 집합 그룹핑 계층(set-grouping hierarchies)으로 상수나 범위 값으로 그룹화 하여 스키마 정의 계층에 대한 재정의 표현을 제공한다. 셋째, 연산 유도 계층(operation-derived hierarchies)은 연산을 기본으로 하여 이메일 주소(e-mail address), URL과 같이 복잡한 데이터 형태로부터 정보에 대한 계층 개념을 통해 정보를 추출할 수 있는 유형이다. 마지막으로 규칙 기반 계층(rule-based hierarchies)은 규칙의 정의와 현재 데이터베이스를 기반으로 규칙 집합에 의해 정의된 전체 개념 또는 부분 개념을 나타낼 수 있다. 계층 기반 마이닝 기법은 위에서 설명한 방법들을 이용하여 어플리케이션의 유형에 맞는 계층 개념을 적용하여 표현 할 수 있고, 계층 기반 연관 규칙 생성 방법을 그림2에서 살펴보면 다음과 같다. 노드 C와 노드 G는 같은 레벨의 동일한 부모 노드를 갖고 있는 연관 규칙을 생성하지만, 노드 C와 노드 I는 서로 다른 레벨의 계층 구조를 갖고, 이때 노드 C와 노드 J와 유사한 연관 규칙을 생성한다. 따라서, 계층적 개념을 갖는 항목들간의 연관 규칙 생성은 계층이 존재하지 않는 단일 레벨에서 연관성을 찾는 방법과는 다르게 규칙을 생성하게 된다.

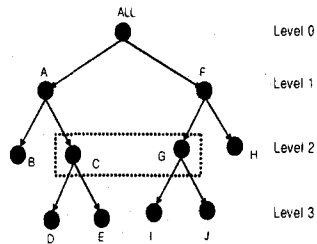


그림 2 계층 개념도

3. 메타 규칙 생성

3.1 메타 규칙

데이터 마이닝을 처리할 때 발생하는 문제는 너무도 방대한 양의 규칙과 패턴들을 생성할 뿐만 아니라 생성된 규칙과 패턴들이 실제 사용자에게 유용한 정보로 제공되지 못한다는 것이다. 이러한 문제들을 해결하고자 지 된 메타쿼리 데이터 마이닝 [7][8](metaquery-directed data mining)방법이 제안되었으나 단일 개념 레벨에서 적용될 뿐 다단계 개념 레벨을 갖는 연관 규칙 발견에는 적당하지 않다. 본 장에서는 다단계 레벨에서 연관 규칙 생성을 효율적으로 수행하기 위해 메타 규칙 생성을 기반으로 한 마이닝 방법을 제안하고자 한다. 메타 규칙의 규칙 형식은 (1)과 같이 나타낸다.

$$P_1 \wedge P_2 \wedge \dots \wedge P_l \rightarrow Q_1 \wedge Q_2 \wedge \dots \wedge Q_r \quad (1)$$

여기서, P_i 와 ($i=1, \dots, l$), Q_j ($j=1, \dots, r$)는 예 된 술어(instantiated predicates) 또는 술어 변수들(predicate variables)이다. 위의 규칙 형식에 따른 메타 규칙의 예는 (2)와 같다.

$$P_1(s, x) \wedge P_2(c, y) \rightarrow Q(s:S, c:C,) \quad (2)$$

다른 개념 레벨에서 데이터 관계의 항목 집합들 사이의 연관 규칙을 나타내는 단일 변수 메타 규칙(single variable meta-rule)의 형식은 (3)과 같다.

$$P_1(t:rel, x_1) \wedge \dots \wedge P_m(t, x_m) \rightarrow Q_1(t, y_1) \wedge \dots \wedge Q_n(t, x_n) \quad (3)$$

(3)의 규칙에서 공통 변수 t 는 관계 키(rel)를 의미한다.

여러 관계들(relations) 사이의 관계성을 나타내는 다중 변수 메타 규칙은 조인 관계를 기반으로 수행된다. 다중 변수 메타 규칙에 대한 형식을 (4)와 같이 나타낼 수 있다.

$$\begin{aligned} &P_1(s, x) \quad P_2(c, y) \quad Q(s:S, c:C,) \\ &P_1(s, x) \wedge P_2(c, y) \rightarrow Q_1(s:S, c:C,) \\ &P_1(s, x) \wedge P_2(c, y) \rightarrow Q_2(s:S, c:C,) \\ &P_1(s, x) \wedge P_2(c, y) \rightarrow Q_3(s:S, c:C,) \end{aligned} \quad (4)$$

세 개의 관계(relation) s, c, z 에 대한 조인 결과를 이용하여 다중 변수 메타 규칙은 Q_1, Q_2, Q_3 중의 하나를 생성할 수 있다.

3.2 연관 규칙을 위한 메타 규칙 생성

다차원 레벨을 갖는 연관 규칙 마이닝을 수행할 때 메타 규칙의 생성은 제한적이고 유용한 규칙만을 효율적으로 생성 가능하도록 할 수 있다. 그림 3은 간 정보를 공유하며 이동하는 RFID 데이터에 대하여 속성들 사이의 의미 관계를 표현하는 스키마 정의 계층(schema-defined hierarchy)과 RFID 데이터 내에 숨겨진 데이터의 정보들을 이용하여 데이터의 개별 속성값을 유도할 수 있는 연산 유도 계층(operation-derived hierarchy)을 혼합한 정보로 구성된 트리 구조로 나타 내었다. 그림3에서 T_1, T_2, \dots, T_n 은 각 개별 아이템 또는 포함 관계를 갖는 항목들간에 이동되는 간에 대한 정보를 나타내며, 이때 T_1 영역에 있는 모든 노드들은 동일한 간에 트랜잭션이 수행 되었음을 의미한다. 그림3의 (a)는 동일한 타임스탬프 T_1 과 동일한 레벨 3에 존재하는 항목들간의 연관 규칙을 갖는 것을 나타내며, 만약 동일한 레벨에 존재하는 항목의 수가 많을 경우

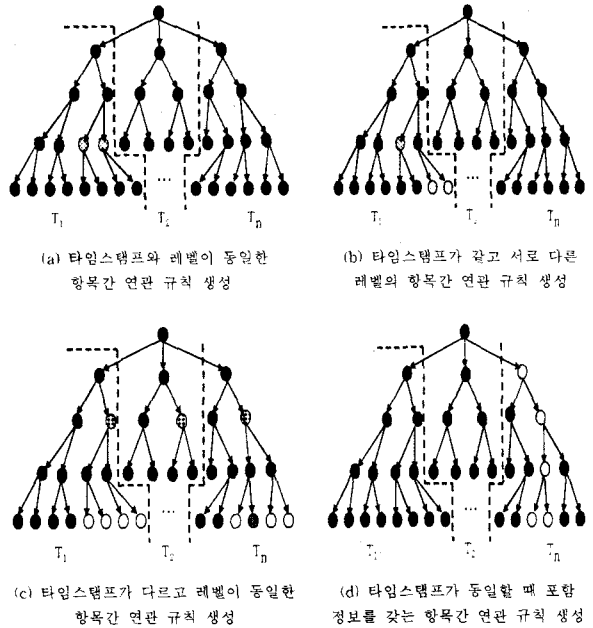


그림 3. 타임스탬프에 따른 RFID 데이터간 계층 개념도

에 특정한 한 항목에 대한 부모(parent) 항목과의 메타 규칙을 통해 모든 형제(sibling) 항목에 대한 연관 규칙을 생성할 필요 없이 일반화 키모로 규칙 생성을 적게 유도할 수 있다. (b)는 동일한 간 T_1 에서 레벨이 서로 다른 항목간의 연관 규칙을 생성하는 방법으로 (a)와 마찬가지로 형제 항목이 갖는 연관 규칙을 공유할 수 있다는 특징이 있다. (c)는 RFID 데이터가 포함 관계를 가지고 간 및 위치의 이동이 발생하는 경우에 있어서 간에 대한 값만 변할 뿐 항목간 포함 관계 정보 및 계층 정보에 대한 변동이 없기 때문에 간에 대한 일반화를 통해 저장 효율 및 질의 효율을 높이기 위한 연관 규칙 생성 방법이다. (d)는 동일한 간내에 발생된 항목들간의 순서(ordered) 정보를 이용하여 레벨간 부분순서 및 전체 순서 즉 항목간 포함 정보에 대한 연관 규칙을 생성할 수 있다. 그림4는 간 정보에 대한 일반화를 수행한 결과 트리 구조이다.

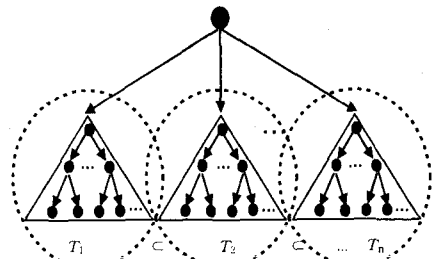


그림 4 타임스탬프의 일반화를 통한 RFID 계층 개념도

간 정보의 일반화를 통해 메타 규칙을 이용한 연관 규칙 생성 알고리즘은 그림5와 같이 나타낼 수 있다.

Input : RFID 데이터베이스(DB), 계층 개념(H), min_sup_threshold[l], min_con_threshold[l], 메타규칙(meta-rule)

Output : 메타 규칙에 따른 제한된 레벨 간 연관 규칙 생성

Process: (1) DB를 간 속성에 대하여 오름차순 정렬한다.
 (2) 동일한 간에 발생한 트랜잭션들에 대하여 일반화(generalization) 킨다. (T_Gid)
 (3) 생성된 메타규칙과 각 레벨의 σ' , φ' 값을 인자로 받는다.
 (4) 레벨 1로부터 최하위 레벨까지 하향식 탐색 방법을 이용하여 순회하며 최소지지도를 만족하는 길이가 k인 후보 항목들을 탐색한다.
 (5) (4)에서 생성된 후보 항목들에 대하여 입력 받은 메타 규칙을 적용하여 의미 있는 후보 항목들을 제외하고 가지치기 한다.
 (6) (5)에서 생성된 결과 항목들에 대하여 (4) (5)의 과정을 반복 수행한다.

그림 5 메타규칙 생성에 따른 제한된 연관 규칙 생성 알고리즘

4. 규칙 생성 분석

4.1 메타 질의

이 장에서는 RFID 데이터베이스로부터 제한된 연관 규칙을 생성하기 위해 생성 가능한 메타 규칙 유형을 살펴보고 메타 질의를 표현하고자 한다. 규칙1은 동일한 간에 발생한 형제 항목들간에 부모 노드와의 중복된 연관성을 제거하고 규칙 생성의 제한을 유도할 수 있는 메타 규칙 표현이다.

$$T_i.sibling(x, y) \rightarrow T_i.parent(x,), C=100\% \quad \text{규칙1}$$

규칙1에 대한 질의 형태의 예를 보면, 먼저 다단계 연관 규칙 생성 방법을 이용하여 Apriori기반 후보항목을 생성한 후 규칙 생성을 위한 x항목과 동일한 레벨에 있는 형제 항목들의 규칙을 모두 제거한다. 형제 항목을 생성 규칙에서 제거한 결과를 후보 항목으로 하여 또 다른 규칙을 유도하므로 일반화된 규칙 정보를 생성할 수 있게 되고, 그 결과 저장 효율 및 질의 수행 시간을 줄일 수 있게 된다.

discover association rules
 from candidate_item(level_l)
 where Timestamp = T_i and parent(x,z) and $\sigma'(x \wedge z/S) \geq \sigma'$ and $\varphi'(x \rightarrow z/S) \geq \varphi'$

delete association rules
 from candidate_item(level_l)
 where Timestamp = T_i and sibling(x,y)

규칙2는 타임스탬프의 일반화를 이용하여 규칙1에 대한 제한된 메타 규칙의 형태를 나타낸다.

$$T_i.parent(x, y) \wedge T_i.parent(x,y) \rightarrow T_i.ancestor(x,), C=100\% \quad \text{규칙2}$$

T_i 에서 x의 부모 정보를 통해 T_j 에서 생성된 x의 조상 정보를 중복되고 흥미 없는 규칙을 제한하여 효율적으로 연관 규칙 생성이 가능하다.

discover meta rules
 from $T_i.candidate_item(level_l)$, $T_j.candidate_item(level_l)$
 where $T_i.parent(x,y) = T_j.parent(x,y)$
 and $\sigma'(x \wedge y/S) \geq \sigma'$ and $\varphi'(x \rightarrow y/S) \geq \varphi'$

discover association rules
 from $T_j.candidate_item(level_l)$
 where Timestamp = T_j and ancestor(x,z) and $\sigma'(x \wedge z/S) \geq \sigma'$ and $\varphi'(x \rightarrow z/S) \geq \varphi'$

5. 결론

본 논문에서는 RFID 데이터베이스가 기존의 트랜잭션 데이터베이스와 달리 간과 위치 정보만을 다르게 갖는 중복된 많은 양의 튜플 생성으로 인해 저장 공간의 비효율성과 질의 복잡성 문제를 해결하고자, 타임스탬프에 의한 일반화 처리와 데이터의 연관 규칙 생성을 제한하기 위한 방법으로 메타 규칙을 이용한 마이닝 기법을 제 하였다. 또한, 제한된 메타 규칙을 적용한 메타 질의를 통해 보다 유용한 정보들을 효율적으로 처리할 수 있다. 향후 연구 방향으로 방대한 양의 데이터를 생성하는 RFID 데이터베이스의 저장효율 및 검색 효율을 높이고 데이터 내에 숨겨진 정보를 효율적으로 찾기 위한 다양한 마이닝 기법의 연구가 필요하다.

6. 참고문헌

[1] H.Gonzalez, J.Han, X.Li, D.Klabjan, "Warehousing and Analyzing Massive RFID Data Sets", Proceedings of the 22nd International Conference on Data Engineering(ICDE'06), 2006.4
 [2] F.Wang, P.Liu, "Temporal Management of RFID Data", Proceedings of the 31st VLDB Conference, Trondheim, Norway, pp.1128-1139, 2005
 [3] R.Agrawal, R.Srikant, "Fast algorithms for mining association rules", In Proc, VLDB, Santiago, Chile, pp487-499, 1994.9
 [4] J.Han, Y.Fu, "Discovery of Multiple-Level Association Rules from Large Databases", Proceedings of the 21st VLDB Conference Zurich, Switzerland, pp.420-431, 1995
 [5] J.Han, Y.Fu, "Mining Multiple-Level Association Rules in Large Databases", Technical Report, Univ. of MissouriRolla, URL:Http://www.umn.edu/ yongjian/pub/, 1997
 [6] G.Psaila, P.Luca Lanzi, "Hierarchy-based Mining of Association Rules in Data Warehouses", In Proc. of the ACM Symposium. on Applied Computing, Como, Italy, pp307-312, 2000.3
 [7] Y.Fu, J.Han, "Meta-Rule-Guided Mining of Association Rules in Relational Databases", In Int. Workshop on Knowledge Discovery and Deductive and Object-Oriented Databases, Singapore, 1995.12
 [8] W.Shen, K.Ong, B.Mitbander, C.Zaniolo, "Metaqueries for data mining", In U.M. Fayyad, G. Piatetsky-Shapiro, P.Smyth, and R.Uthrusamy, editors, Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1995