



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년09월19일
 (11) 등록번호 10-2023275
 (24) 등록일자 2019년09월11일

(51) 국제특허분류(Int. Cl.)
 G06Q 30/02 (2012.01) G06Q 30/06 (2012.01)
 (52) CPC특허분류
 G06Q 30/0254 (2013.01)
 G06Q 30/0255 (2013.01)
 (21) 출원번호 10-2018-0026213
 (22) 출원일자 2018년03월06일
 심사청구일자 2018년03월06일
 (65) 공개번호 10-2019-0105746
 (43) 공개일자 2019년09월18일
 (56) 선행기술조사문헌
 KR1020150027442 A*
 KR1020160027263 A*
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
 세종대학교산학협력단
 서울특별시 광진구 능동로 209 (군자동, 세종대학교)
 (72) 발명자
 신동일
 서울특별시 강남구 압구정로 347, 26동 1207호(압구정동, 한양아파트)
 신동규
 서울특별시 강남구 언주로 201, 1903호(도곡동, SK리더스뷰)
 신경일
 광주광역시 남구 용대로 115, 104동 102(봉선동, 무등파크1차아파트)
 (74) 대리인
 양성보

전체 청구항 수 : 총 5 항

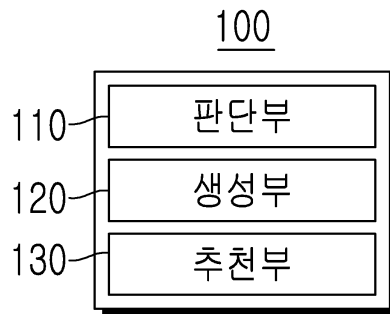
심사관 : 이충근

(54) 발명의 명칭 **FP Growth를 이용한 제품 추천 방법 및 시스템**

(57) 요약

FP Growth를 이용한 제품 추천 방법 및 시스템이 개시된다. 일 실시예에 따른 제품 추천 시스템에 의하여 수행되는 제품 추천 방법은, 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계; 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계; 및 상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 단계를 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류
G06Q 30/0631 (2013.01)

이 발명을 지원한 국가연구개발사업

| | |
|----------|--|
| 과제고유번호 | 1711055356 |
| 부처명 | 과학기술정보통신부 |
| 연구관리전문기관 | 정보통신기술진흥센터 |
| 연구사업명 | ICT유망기술개발지원 |
| 연구과제명 | IOT 기반의 고객데이터 수집 및 자동인식을 통한 빅데이터분석 클라우드 고객센싱 서비스 |
| 시스템 구축 | |
| 기여율 | 1/1 |
| 주관기관 | 주식회사소프트자이온 |
| 연구기간 | 2017.05.01 ~ 2018.04.30 |

명세서

청구범위

청구항 1

제품 추천 시스템에 의하여 수행되는 제품 추천 방법에 있어서,

데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계;

상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계; 및

상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 단계

를 포함하고,

상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는,

FP Growth 알고리즘을 사용하여 상기 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출하고, 상기 사용자의 구매 내역 정보를 복수 번 스캔함에 따라 패턴을 탐색하고,

상기 사용자의 구매 내역 정보를 스캔함에 따라 트랜잭션이 포함하고 있는 아이템마다 서포트 값을 계산하고, 상기 계산된 서포트 값을 토대로 서포트 값이 최소 임계치(minsup) 이상인 아이템을 추출하고, 상기 트랜잭션에서 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 상기 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성하고, 상기 서포트 값이 최소 임계치 이상인 아이템을 기 설정된 기준으로 정렬하여 테이블에 보관하고,

상기 사용자의 구매 내역 정보를 재스캔함에 따라 트랜잭션에 포함된 아이템 중 테이블에 보관되어 있는 아이템을 추출하고, 상기 추출된 아이템을 재정렬하여 트리에 삽입하고, 상기 추출된 아이템을 트랜잭션에 root부터 삽입하고, 상기 삽입을 하는 도중에 동일한 아이템이 추출될 경우 노드를 합치고, 노드에 저장하는 서포트 값을 증가시키는 과정을 완료함에 따라 트리 또는 서브 트리를 구축하고, 상기 트리 또는 상기 서브 트리를 계속적으로 생성하고, 서포트 값을 추출함에 따라 최소 임계치 이상의 패턴들을 추출하고,

1)상기 트리 또는 상기 서브 트리를 기준으로 상기 트리 또는 상기 서브 트리의 이전 트리에서 새로운 트리 또는 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 2)상기 선택된 아이템을 포함한 패턴의 서포트 값이 최소 임계값 이하일 때 상기 트리 또는 상기 서브 트리의 생성을 중지하는 단계

를 포함하고,

상기 분할 정복 기법은, 상기 테이블에서 바텀업 방식으로 상기 아이템을 prefix로 하고, 상기 아이템을 포함하는 트랜잭션의 다른 아이템들의 서포트 값을 계산하고, 상기 계산된 서포트 값이 상기 최소 임계치보다 크거나 같을 경우, 서브 트리가 생성되는

제품 추천 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

제1항에 있어서,

상기 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계는,

상기 데이터를 바이너리 코드로 변환하는 전처리 과정을 수행하고, 상기 전처리 과정이 수행됨에 따라 바이너리 코드로 변환된 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 확인하는 단계

를 포함하는 제품 추천 방법.

청구항 7

제품 추천 시스템에 의하여 수행되는 제품 추천 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램에 있어서,

데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계;

상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계; 및

상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 단계

를 포함하고,

상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는,

FP Growth 알고리즘을 사용하여 상기 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출하고, 상기 사용자의 구매 내역 정보를 복수 번 스캔함에 따라 패턴을 탐색하고,

상기 사용자의 구매 내역 정보를 스캔함에 따라 트랜잭션이 포함하고 있는 아이템마다 서포트 값을 계산하고, 상기 계산된 서포트 값을 토대로 서포트 값이 최소 임계치(minsup) 이상인 아이템을 추출하고, 상기 트랜잭션에서 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 상기 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성하고, 상기 서포트 값이 최소 임계치 이상인 아이템을 기 설정된 기준으로 정렬하여 테이블에 보관하고,

상기 사용자의 구매 내역 정보를 재스캔함에 따라 트랜잭션에 포함된 아이템 중 테이블에 보관되어 있는 아이템을 추출하고, 상기 추출된 아이템을 재정렬하여 트리에 삽입하고, 상기 추출된 아이템을 트랜잭션에 root부터 삽입하고, 상기 삽입을 하는 도중에 동일한 아이템이 추출될 경우 노드를 합치고, 노드에 저장하는 서포트 값을 증가시키는 과정을 완료함에 따라 트리 또는 서브 트리를 구축하고, 상기 트리 또는 상기 서브 트리를 계속적으로 생성하고, 서포트 값을 추출함에 따라 최소 임계치 이상의 패턴들을 추출하고,

1)상기 트리 또는 상기 서브 트리를 기준으로 상기 트리 또는 상기 서브 트리의 이전 트리에서 새로운 트리 또는 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 2)상기 선택된 아이템을 포함한 패턴의 서포트 값이 최소 임계값 이하일 때 상기 트리 또는 상기 서브 트리의 생성을 중지하는 단계

를 포함하고,

상기 분할 정복 기법은, 상기 테이블에서 바텀업 방식으로 상기 아이템을 prefix로 하고, 상기 아이템을 포함하는 트랜잭션의 다른 아이템들의 서포트 값을 계산하고, 상기 계산된 서포트 값이 상기 최소 임계치보다 크거나 같을 경우, 서브 트리가 생성되는

컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 8

제품 추천 시스템에 있어서,

데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 판단부;

상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 생성부; 및

상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 추천부

를 포함하고,

상기 생성부는,

FP Growth 알고리즘을 사용하여 상기 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출하고, 상기 사용자의 구매 내역 정보를 복수 번 스캔함에 따라 패턴을 탐색하고,

상기 사용자의 구매 내역 정보를 스캔함에 따라 트랜잭션이 포함하고 있는 아이템마다 서포트 값을 계산하고, 상기 계산된 서포트 값을 토대로 서포트 값이 최소 임계치(minsup) 이상인 아이템을 추출하고, 상기 트랜잭션에서 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 상기 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성하고, 상기 서포트 값이 최소 임계치 이상인 아이템을 기 설정된 기준으로 정렬하여 테이블에 보관하고,

상기 사용자의 구매 내역 정보를 재스캔함에 따라 트랜잭션에 포함된 아이템 중 테이블에 보관되어 있는 아이템을 추출하고, 상기 추출된 아이템을 재정렬하여 트리에 삽입하고, 상기 추출된 아이템을 트랜잭션에 root부터 삽입하고, 상기 삽입을 하는 도중에 동일한 아이템이 추출될 경우 노드를 합치고, 노드에 저장하는 서포트 값을 증가시키는 과정을 완료함에 따라 트리 또는 서브 트리를 구축하고, 상기 트리 또는 상기 서브 트리를 계속적으로 생성하고, 서포트 값을 추출함에 따라 최소 임계치 이상의 패턴들을 추출하고,

1)상기 트리 또는 상기 서브 트리를 기준으로 상기 트리 또는 상기 서브 트리의 이전 트리에서 새로운 트리 또는 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 2)상기 선택된 아이템을 포함한 패턴의 서포트 값이 최소 임계값 이하일 때 상기 트리 또는 상기 서브 트리의 생성을 중지하는 것을 포함하고,

상기 분할 정복 기법은, 상기 테이블에서 바텀업 방식으로 상기 아이템을 prefix로 하고, 상기 아이템을 포함하는 트랜잭션의 다른 아이템들의 서포트 값을 계산하고, 상기 계산된 서포트 값이 상기 최소 임계치보다 크거나 같을 경우, 서브 트리가 생성되는

제품 추천 시스템.

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

제8항에 있어서,

상기 판단부는,

상기 데이터를 바이너리 코드로 변환하는 전처리 과정을 수행하고, 상기 전처리 과정이 수행됨에 따라 바이너리 코드로 변환된 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 확인하는

것을 특징으로 하는 제품 추천 시스템.

발명의 설명

기술 분야

[0001] 아래의 설명은 제품을 추천하는 기술에 관한 것이다.

배경 기술

[0002] 현대 사회는 하드웨어와 인터넷의 발달로 인하여 유비쿼터스 컴퓨팅에서 IoT로 진화되고, 개인 소유의 스마트폰을 보유하는 시대가 되었다. 이로 인하여 소비패턴 또한 과거와 달리 오프라인구매보다는 온라인 구매를 선호하는 시대가 되어가고 있다. 실제 한 백화점의 경우 20대와 30대의 경우 오프라인보다 온라인에서 물건을 구매하는 사람이 월등하게 증가하였으며, 이는 구입할 물건의 사양, 사용기나 가격 등의 정보를 먼저 찾아보고 최적의 경로를 통해 최적가로 합리적으로 구매하려는 소비자의 특성으로 인하여 점차 증가되는 추세이다. 온라인 쇼핑몰을 이용하여 소비자가 제품을 구매할 경우 구매 기록이 남게 된다.

[0003] 한편, 실제 과거부터 마케팅에서 고객의 데이터를 수집한 후 이를 분석하여 실제 고객을 관리하는데 많이 사용하였으며, 대표적으로 CRM(Customer Relationship Management)을 이용하여 많은 기업들은 고객과 관련된 데이터를 분석 및 통합해 고객 개개인의 특성에 맞게 마케팅을 하는데 적용되었다. 최근 머신러닝 및 딥러닝이 사람으로부터 각광을 받으면서 마케팅 분야에도 기존에 사람이 하던 일의 일부분을 머신러닝이나 딥러닝을 적용하여 기계가 자동으로 일을 처리하고 있다. 기업에서는 수익을 창출하기 위해 신규 고객을 모으고, 기존고객의 이탈을 방지하며, 고객들이 필요로 하는 것을 제시하여야 한다.

[0004] 데이터 마이닝을 이용하여 마케팅을 하는 대표적인 사례로 고객세분화 기호도 분석에 따른 타겟 마케팅이 있고, 수요예측을 통한 제품서비스 공급 준비가 있으며, 구매행동 데이터분석을 통한 마케팅 유통 계획안 개선 등이 있다. 고객들이 온라인 쇼핑몰을 통해 물건을 구매하는 경우 수많은 제품 중 원하는 제품을 찾기가 어려운 경우가 다수 발생한다. 또한 기업 입장에서는 분석을 통해 고객이 필요로 한 제품을 추천하여 더 많은 제품을 판매하여 수익을 창출해야 한다.

[0005] 따라서 이러한 문제점을 해결하기 위해 축적된 고객들의 데이터에 기반하여 고객들의 구매내역을 분석하여 각 고객이 원하는 제품을 추천하는 추천 서비스를 제공할 필요가 있다.

발명의 내용

해결하려는 과제

[0006] FP Growth 를 이용하여 고객의 구매 내역을 분석하여 고객에게 제품을 추천하는 방법 및 시스템을 제공할 수 있다.

과제의 해결 수단

[0007] 제품 추천 시스템에 의하여 수행되는 제품 추천 방법은, 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계; 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계; 및 상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 단계를 포함할 수 있다.

[0008] 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는, FP Growth 알고리즘을 사용하여 상기 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출하는 단계를 포함할 수 있다.

[0009] 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는, 상기 사용자의 구매 내역 정보를 2번 스캔함에 따라 패턴을 탐색하는 단계를 포함할 수 있다.

[0010] 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는, 상기 구매 내역 정보를 스캔함에 따라 트랜잭션에서 minsup 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 상기 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성하는 단계를 포함할 수 있다.

[0011] 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계는, 상기 서브 트리를 기준으로 상기 서브 트리의 이전 트리에서 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 상기 선택된 아이템을 포함한 패턴

턴의 서포트(support)가 minsup 이하일 때 상기 서브 트리의 생성을 중지하는 단계를 포함할 수 있다.

- [0012] 상기 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계는, 상기 데이터를 바이너리 코드로 변환하는 전처리 과정을 수행하고, 상기 전처리 과정이 수행됨에 따라 바이너리 코드로 변환된 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 확인하는 단계를 포함할 수 있다.
- [0013] 제품 추천 시스템에 의하여 수행되는 제품 추천 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램은, 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 단계; 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 단계; 및 상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 단계를 포함할 수 있다.
- [0014] 제품 추천 시스템은, 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단하는 판단부; 상기 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성하는 생성부; 및 상기 연관규칙에 기반하여 획득된 결과를 통하여 상기 사용자에게 새로운 제품을 추천하는 추천부를 포함할 수 있다.
- [0015] 상기 생성부는, FP Growth 알고리즘을 사용하여 상기 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출할 수 있다.
- [0016] 상기 생성부는, 상기 사용자의 구매 내역 정보를 2번 스캔함에 따라 패턴을 탐색할 수 있다.
- [0017] 상기 생성부는, 상기 구매 내역 정보를 스캔함에 따라 트랜잭션에서 minsup 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 상기 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성할 수 있다.
- [0018] 상기 생성부는, 상기 서브 트리를 기준으로 상기 서브 트리의 이전 트리에서 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 상기 선택된 아이템을 포함한 패턴의 서포트(support)가 minsup 이하일 때 상기 서브 트리의 생성을 중지할 수 있다.
- [0019] 상기 판단부는, 상기 데이터를 바이너리 코드로 변환하는 전처리 과정을 수행하고, 상기 전처리 과정이 수행됨에 따라 바이너리 코드로 변환된 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 확인할 수 있다.

발명의 효과

- [0020] 일 실시예에 따른 제품 추천 시스템은 FP Growth를 적용하여 획득된 결과를 통해 고객 각각의 구매 내역을 확인하고 고객의 구매력이 높은 제품을 추천하는 서비스를 제공할 수 있다.
- [0021] 일 실시예에 따른 제품 추천 시스템은 빠른 속도로 분석이 가능하고, 연산량을 감소시킬 수 있다.

도면의 간단한 설명

- [0022] 도 1은 일 실시예에 따른 제품 추천 시스템의 구성을 설명하기 위한 블록도이다.
- 도 2는 일 실시예에 따른 제품 추천 시스템의 제품 추천 방법을 설명하기 위한 흐름도이다.
- 도 3은 일 실시예에 따른 제품 추천 시스템의 사용자의 구매 내역에 대한 트랜잭션을 나타낸 예이다.
- 도 4는 일 실시예에 따른 제품 추천 시스템에서 트리를 생성하는 과정을 나타낸 것이다.
- 도 5는 일 실시예에 따른 제품 추천 시스템에서 사용자의 구매 내역을 나타낸 예이다.
- 도 6은 일 실시예에 따른 제품 추천 시스템에서 전처리 과정을 수행하는 것을 설명하기 위한 도면이다.
- 도 7은 일 실시예에 따른 제품 추천 시스템에서 전처리 과정이 수행된 데이터 셋을 나타낸 도면이다.
- 도 8은 일 실시예에 따른 제품 추천 시스템에서 FP Growth 알고리즘을 적용한 제품 추천 목록을 나타낸 예이다.
- 도 9는 일 실시예에 따른 제품 추천 시스템에서 제품을 추천한 예이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.
- [0024] 도 1은 일 실시예에 따른 제품 추천 시스템의 구성을 설명하기 위한 블록도이고, 도 2는 일 실시예에 따른 제품 추천 시스템의 제품 추천 방법을 설명하기 위한 흐름도이다.
- [0025] 제품 추천 시스템(100)은 판단부(110), 생성부(120) 및 추천부(130)를 포함할 수 있다. 이러한 구성요소들은 제품 추천 시스템(100)에 저장된 프로그램 코드가 제공하는 제어 명령에 따라 프로세서에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다. 구성요소들은 도 2의 제품 추천 방법이 포함하는 단계들(210 내지 230)을 수행하도록 제품 추천 시스템(100)을 제어할 수 있다. 이때, 구성요소들은 메모리가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.
- [0026] 프로세서는 제품 추천 방법을 위한 프로그램의 파일에 저장된 프로그램 코드를 메모리에 로딩할 수 있다. 예를 들면, 제품 추천 시스템(100)에서 프로그램이 실행되면, 프로세서는 운영체제의 제어에 따라 프로그램의 파일로부터 프로그램 코드를 메모리에 로딩하도록 제품 추천 시스템을 제어할 수 있다. 이때, 프로세서 및 프로세서가 포함하는 판단부(110), 생성부(120) 및 추천부(130) 각각은 메모리에 로딩된 프로그램 코드 중 대응하는 부분의 명령을 실행하여 이후 단계들(210 내지 230)을 실행하기 위한 프로세서의 서로 다른 기능적 표현들일 수 있다.
- [0027] 단계(210)에서 판단부(110)는 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단할 수 있다. 판단부(110)는 데이터를 바이너리 코드로 변환하는 전처리 과정을 수행하고, 전처리 과정이 수행됨에 따라 바이너리 코드로 변환된 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 확인할 수 있다.
- [0028] 단계(220)에서 생성부(120)는 판단된 사용자의 구매 내역에 따라 연관규칙을 생성할 수 있다. 생성부(120)는 FP Growth 알고리즘을 사용하여 사용자의 구매 내역 정보를 스캔함에 따라 패턴을 추출할 수 있다. 생성부(120)는 사용자의 구매 내역 정보를 2번 스캔함에 따라 패턴을 탐색할 수 있다. 생성부(120)는 구매 내역 정보를 스캔함에 따라 트랜잭션에서 minsup 이상의 아이템들로 트리(Global FP Tree)를 구축하고, 구축된 트리에서 분할 정복 기법을 사용하여 또 다른 아이템을 선택하여 적어도 하나 이상의 서브 트리(Conditional FP Tree)를 생성할 수 있다. 생성부(120)는 서브 트리를 기준으로 서브 트리의 이전 트리에서 새로운 서브 트리의 생성이 불가능한 상태일 때 또는 선택된 아이템을 포함한 패턴의 서포트(support)가 minsup 이하일 때 서브 트리의 생성을 중지할 수 있다.
- [0029] 단계(230)에서 추천부(130)는 연관규칙에 기반하여 획득된 결과를 통하여 사용자에게 새로운 제품을 추천할 수 있다.
- [0030] 도 3은 일 실시예에 따른 제품 추천 시스템의 사용자의 구매 내역에 대한 트랜잭션을 나타낸 예이다.
- [0031] FP Growth(Frequent Patterns Growth)는 Apriori 알고리즘을 개선한 알고리즘으로 데이터가 가진 패턴이 얼마나 많이 발생하는지를 분석하는 패턴 마이닝에 쓰이는 기법 중 하나이다. 실시예에서는, FP Growth(Frequent Patterns Growth)를 이용하여 제품을 추천하는 방법을 설명하기로 한다. 여기서, 사용자(예를 들면, 고객)의 구매 내역을 표현할 수 있는 제품의 수를 k 라고 할 때, 공집합을 제외한 부분집합의 개수를 $2^k - 1$ 개, 연관규칙의 개수를 $3^k - 2^{k+1} + 1$ 개이다. 예를 들면, 제품이 10개라고 하였을 때, 부분집합의 개수는 1023개, 연관규칙의 개수는 57002개이다. 제품의 수가 많아질수록 연관규칙의 경우의 수도 기하급수적으로 증가하게 된다.
- [0032] 제품 추천 시스템은 데이터의 속성에 기초하여 사용자의 구매 내역 정보를 판단할 수 있다. 제품 추천 시스템은 연관규칙 생성 알고리즘을 이용하게 된다. 제품 추천 시스템은 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성할 수 있다. 종래의 부분집합인 지지도 계산량을 감소시키기 위한 다양한 방법이 존재하였다. Apriori 알고리즘은 모든 가능한 항목집합의 개수를 줄이는 방법을 사용하고, DHP 알고리즘의 경우 트랜잭션 개수를 줄이는 방법을 사용하며, FP Growth 알고리즘은 비교하는 수를 줄이는 방법을 사용한다. Apriori와 동일한 결과를 보여주지만 Apriori에 비해 실행시간을 급격하게 감소시킨 알고리즘으로 FP Growth 알고리즘을 주로 이용한다. 본 발명의 제품 추천 시스템은 이러한 FP Growth 알고리즘을 이용하여 제품을 추천할 수 있다. Apriori 알고리즘은 패턴을 찾기 위해 할당된 트랜잭션의 아이템 수개의 스캔을 해야 하는 반면, FP Growth 알고리즘을 이용할 경우, 2번의 스캔만 수행하면 된다. 이로 인하여, FP Growth 알고리즘을 이용할 경우 실행 시간을 감소시킬 수 있으며, 패턴 후보인 candidate를 생성하지 않는다는 특징이 있다. 다시 말해서, 제품 추천

시스템은 FP Growth 알고리즘을 통하여 candidate를 생성하지 않는 대신 트리와 노드 링크라는 특별한 자료 구조를 사용한다.

[0033] 제품 추천 시스템은 판단된 사용자의 구매 내역 정보에 따라 연관규칙을 생성할 수 있다. 제품 추천 시스템은 사용자의 구매 내역에 대하여 첫 번째 스캔을 수행할 수 있다. 이때, 사용자의 구매 내역은 트랜잭션 데이터를 의미할 수 있다. 제품 추천 시스템은 사용자의 구매 내역을 스캔함에 따라 트랜잭션이 포함하고 있는 아이템마다 서포트(반발도, support)를 계산할 수 있다. 제품 추천 시스템은 계산된 서포트를 토대로 서포트가 최소 임계치(minsup)이상인 아이템을 추출할 수 있다. 제품 추천 시스템은 트랜잭션에서 minsup이상의 아이템들로 Global FP Tree를 생성하고, 생성된 트리에서 분할 정복 기법을 사용하여 또 하나의 아이템을 선택해서 새로운 서브 트리인 Conditional FP Tree를 생성할 수 있다. 이와 같이, 아이템의 길이를 하나씩 늘려가면서 서브 트리를 복수 개를 생성할 수 있다. 이때, 트리 또는 서브 트리의 이전 트리에서 더 이상 서브 트리를 구축할 수 있을 때, 또는, 선택한 아이템을 포함하는 패턴의 서포트 값이 minsup보다 작을 경우에 트리 및 서브 트리의 생성을 중지할 수 있다.

[0034] 구체적으로, 제품 추천 시스템은 서포트가 minsup이상인 아이템을 추출함에 따라 추출된 아이템을 기 설정된 기준(예를 들면, 내림차순, 올림차순 등)으로 정렬하여 테이블에 보관할 수 있다. 제품 추천 시스템은 사용자의 구매 내역을 두 번째 스캔(재스캔)할 수 있다. 제품 추천 시스템은 사용자의 구매 내역을 다시 한번 스캔함으로써 트랜잭션에 포함된 아이템 중 테이블에 보관되어 있는 아이템을 추출할 수 있고, 추출된 아이템을 재정렬하여 트리에 삽입할 수 있다. 제품 추천 시스템은 트랜잭션마다 root부터 삽입을 할 수 있으며, 새로운 아이템이 나타나면 head of node-links라는 포인터가 그 아이템 노드를 가리키도록 한다. 그리고, 삽입을 하는 도중 동일한 아이템이 나올 경우 노드를 합치고, 노드에 저장하는 서포트 값을 증가시킬 수 있다. 예를 들면, 서포트 값을 기 설정된 크기(예를 들면, 1)만큼 증가시킬 수 있다. 노드 링크는 새로운 노드가 생길 때마다 이전 아이템 링크에서 계속 연결시킬 수 있다. 이러한 과정을 완료함에 따라 트리를 구축할 수 있다. 이때, 구축된 트리를 Global Tree라고 한다. 제품 추천 시스템은 한번에 해결하기 어려운 문제를 작은 단위의 부분 문제로 분할하여 해결하는 분할 정복 기법으로 테이블에서 바텀업 방식으로 아이템을 하나씩 prefix로 하고, 아이템을 포함하는 트랜잭션의 다른 아이템들의 서포트 값을 계산하고, minsup(최소 임계치)보다 크거나 같을 경우, 서브 트리를 생성할 수 있다. 이때, 서브 트리를 conditional Tree라고 한다. 이러한 방식으로 트리 및 서브 트리를 계속적으로 생성하고, 서포트 값을 산출하면서 패턴이 minsup 이상의 패턴들이 도출될 수 있다. 제품 추천 시스템은 minsup 이상의 패턴을 추출할 수 있다. 예를 들면, 제품 추천 시스템의 사용자의 구매 성향을 추출할 수 있다.

[0035] 도 4는 일 실시예에 따른 제품 추천 시스템에서 트리를 생성하는 과정을 나타낸 것이다. 도 4를 참고하면, 사용자별 구매 내역으로 생성된 FP Tree를 나타낸 것으로, 사용자(식별 정보, ID)는 5명이고, 사용자의 구매 내역에서의 아이템의 수는 6인 것을 도식화 하였다. 사용자의 구매 내역의 트랜잭션을 오름차순 또는 내림차순으로 정렬하여 차례로 트리에 삽입할 수 있다. 제품 추천 시스템은 연관규칙에 기반하여 획득된 결과를 통하여 고객에게 새로운 제품을 추천할 수 있다.

[0036] 도 5를 참고하면, 사용자의 구매 내역을 나타낸 것이다. 일례로, 제품 추천 시스템은 분석할 원본 데이터에서 아이디별 구매 아이템을 출력할 수 있다. 제품 추천 시스템은 연관 규칙 알고리즘을 적용하기 위하여 데이터를 바이너리 코드 형태로 변형하는 전처리 과정을 수행할 수 있다. 제품 추천 시스템은 데이터에 대한 전처리 과정을 완료함에 따라 도 7과 같은 형태의 데이터를 획득할 수 있다. 도 6을 참고하면, 데이터 셋의 속성을 뜻하며, 사용자가 제품을 구매한 이력이 있는지 여부에 대하여 t/f로 나타낸다. 예를 들면, t는 제품을 구매한 이력이 있을 경우, f는 제품을 구매한 이력이 없을 경우를 의미할 수 있다. 도 7을 참고하면, 중괄호 안에 숫자들이 포함되어 있는데 이것은 데이터에 해당하는 속성, 즉 제품을 뜻하는 것으로 0번부터 시작하게 된다. 그리고 각 중괄호는 각각의 한 사람의 제품에 대한 구매 내역을 의미한다. 만약 0번째 속성에 해당하는 제품을 구매한 이력이 있으면 t로 표현하며, 구매하지 않은 제품에 대해서는 따로 표기를 하지 않고 생략할 수 있다.

[0037] 제품 추천 시스템은 예를 들면, lowerBoundMinSupport 값을 0.001로 설정하고, minMetric의 값을 0.5로 설정한 후 해당 데이터에 FP Growth 알고리즘을 적용할 수 있다. 이때, lowerBoundMinSupport란 앞에서 말한 minsup에 해당하며 {전체 트랜잭션 * 0.001}에 해당하는 값을 minsup로 설정한다는 뜻이고, minMetric은 추천에 대한 신뢰도의 최소 임계치로 신뢰도가 0.5이상인 값만 출력한다는 것을 의미한다.

[0038] 도 8은 일 실시예에 따른 제품 추천 시스템에서 FP Growth 알고리즘을 적용한 제품 추천 목록을 나타낸 예이다. purchaseProduct는 구매한 제품, purchaseQuantity는 상기 제품을 구매한 사람의 숫자를 의미한다.

recommendProduct는 purchaseProduct를 구매한 사람이 구매한 제품, 다시 말해서, 이는 추천 제품이 된다. 그리고 recommendQuantity는 purchaseProduct와 recommendProduct를 구매한 사람의 수를 뜻한다. Confidence는 recommendQuantity/purchaseQuantity의 연산식을 가지며 제품 추천에 대한 신뢰도이다. 제품 추천 시스템은 결과 빠른 속도로 분석이 가능하였으며, 몇몇의 제품을 구매한 경우 특정 제품을 자주 구매하는 것을 알 수 있었고, 이와 같이 획득된 결과를 통해 해당 제품을 구매한 사용자에게 특정 제품을 추천할 수 있다. 도 9는 일 실시예에 따른 제품 추천 시스템에서 제품을 추천한 예이다. 9(a)는 구매한 제품인 제품 BGMP310이고, 9(b)는 BGSP321 여성 흰색 바지를 나타낸 것이다. 예를 들면, 제품 BGMP310은 여성 검정 바지이고, BGSB435는 여성 모 더니티 블라우스 상품이며, 이러한 두 제품을 구매한 경우 BGMP310과 유사한 스타일의 흰색 바지를 추천할 수 있다.

[0039] 일 실시예에 따른 제품 추천 시스템은 기존의 방법이 트랜잭션도 많이 읽어야 하고, 연산량이 많이 필요하였던 문제점을 FP Growth 알고리즘을 사용하여 해결할 수 있다. 다시 말해서, 제품 추천 시스템은 2번의 스캔만으로 결과를 획득할 수 있으며, 연산량도 적어도 분석 속도가 빠르다. 제품 추천 시스템은 FP Growth를 적용하여 획득된 결과를 통해 고객 각각의 구매 내역을 확인하고 고객의 구매력이 높은 제품을 추천하는 서비스를 제공할 수 있다.

[0040] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0041] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

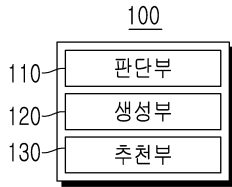
[0042] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0043] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

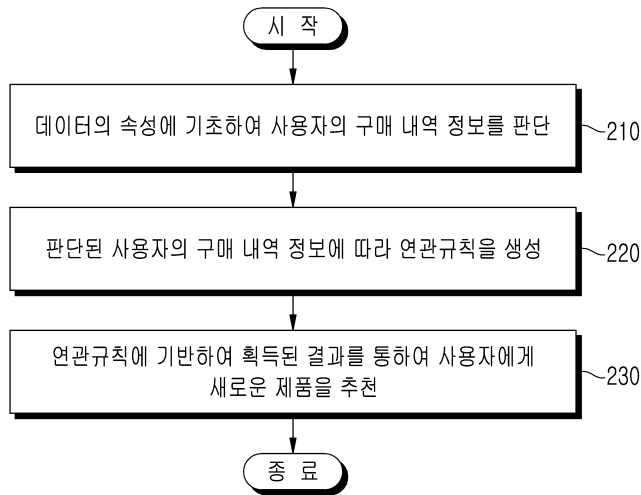
[0044] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

도면1



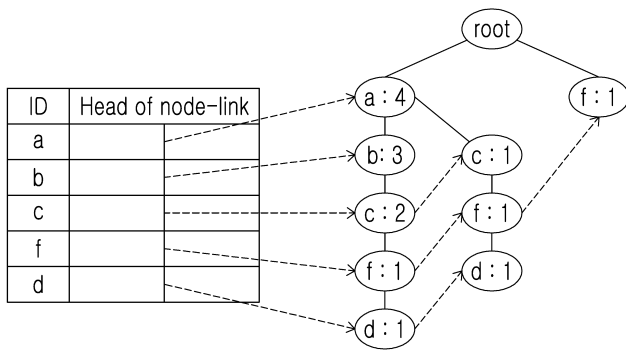
도면2



도면3

| ID | Item | Frequent Item |
|----|------------------|---------------|
| 1 | a, b, c, d, e, f | a, b, c, f, d |
| 2 | a, b, c | a, b, c |
| 3 | a, b | a, b |
| 4 | a, c, d, f | a, c, d, f |
| 5 | f | f |

도면4



도면5

| | |
|----------|---------|
| a0310213 | BGSV106 |
| a0310213 | BGSP326 |
| a0310213 | BGST539 |
| a0310213 | BGST538 |
| a0310213 | BGSP326 |
| a0625029 | B4331 |
| a070408 | B6039 |
| a070408 | B6025 |
| a070408 | B6728 |
| a070408 | B6001 |
| a070408 | B6714 |
| a070408 | B6025 |
| a070408 | B6059 |
| a10443 | BFFB423 |
| a10443 | BFFS205 |
| a10443 | BFFP321 |
| a10443 | BFFB427 |
| a10443 | BGSP306 |
| a10443 | BGSP311 |
| a10443 | BGMT511 |
| a10443 | BGMB406 |
| a10443 | BGMP307 |

도면6

```

@attribute product_code_B6516 {f,t}
@attribute product_code_B3128 {f,t}
@attribute product_code_B6542 {f,t}
@attribute product_code_C0365 {f,t}
@attribute product_code_B6804 {f,t}
@attribute product_code_B6951 {f,t}
@attribute product_code_BFFO610 {f,t}
@attribute product_code_B1716 {f,t}
@attribute product_code_BFFO614 {f,t}
@attribute product_code_BFFK947 {f,t}
@attribute product_code_BFFT531 {f,t}
@attribute product_code_BFFS211 {f,t}
@attribute product_code_BFFO628 {f,t}
@attribute product_code_BFFP324 {f,t}
@attribute product_code_BFFP340 {f,t}
@attribute product_code_BGST502 {f,t}
@attribute product_code_BGSO611 {f,t}
@attribute product_code_BGSJ128 {f,t}
    
```

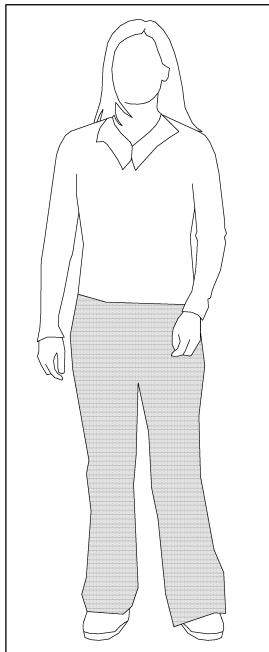
도면7

(10 t,1 t,2 t,3 t,4 t,5 t,6 t,7 t,8 t,9 t,10 t,11 t,12 t,13 t,14 t,15 t,16 t,17 t)
 (18 t,19 t,20 t,21 t,22 t,23 t,24 t,25 t,26 t,27 t,28 t,29 t,30 t,31 t,32 t,33 t,34 t,35 t,36 t,37 t,38 t,39 t,40 t,41 t,42 t,43 t,44 t,45 t,46 t,47 t,48 t)
 (35 t,39 t,55 t,56 t,57 t,58 t,59 t,60 t,61 t,62 t)
 (63 t,64 t,65 t)
 (66 t,67 t)
 (68 t)
 (70 t,71 t,72 t,73 t,74 t)
 (26 t,45 t,75 t,76 t,77 t,78 t,79 t,80 t,81 t,82 t,83 t,84 t,85 t,86 t,87 t,88 t,89 t,90 t,91 t,92 t,93 t,94 t,95 t,96 t,97 t,98 t,99 t,100 t,101 t,102 t,103 t,104 t,105 t,106 t,107 t,108 t,109 t,110 t,111 t,112 t,113 t,114 t,115 t,116 t,117 t,118 t,119 t,120 t,121 t,122 t,123 t,124 t,125 t,126 t,127 t,128 t,129 t,130 t,131 t,132 t,133 t,134 t,135 t,136 t,137 t,138 t,139 t,140 t,141 t,142 t,143 t,144 t,145 t,146 t,147 t,148 t,149 t,150 t,151 t,152 t,153 t,154 t,155 t,156 t,157 t,158 t,159 t,160 t,161 t,162 t,163 t,164 t,165 t,166 t,167 t,168 t,169 t,170 t,171 t,172 t,173 t,174 t,175 t,176 t,177 t,178 t,179 t,180 t,181 t,182 t,183 t,184 t,185 t,186 t,187 t,188 t,189 t,190 t,191 t,192 t,193 t,194 t,195 t,196 t,197 t,198 t,199 t,200 t,201 t,202 t,203 t,204 t,205 t,206 t,207 t,208 t,209 t,210 t,211 t,212 t,213 t,214 t,215 t,216 t,217 t,218 t,219 t,220 t,221 t,222 t,223 t,224 t,225 t,226 t,227 t,228 t,229 t,230 t,231 t,232 t,233 t,234 t,235 t,236 t,237 t,238 t,239 t,240 t,241 t,242 t,243 t,244 t,245 t,246 t,247 t,248 t,249 t,250 t,251 t,252 t,253 t,254 t,255 t)
 (13 t,256 t,257 t,258 t,259 t)
 (93 t,260 t)
 (96 t,79 t,206 t,261 t)
 (42 t,234 t,262 t)
 (24 t,25 t,26 t,46 t,80 t,100 t,107 t,155 t,163 t,175 t,263 t,264 t,265 t,266 t,267 t,268 t,269 t,270 t,271 t,272 t,273 t,274 t,275 t,276 t,277 t,278 t)
 (20 t,279 t,280 t)

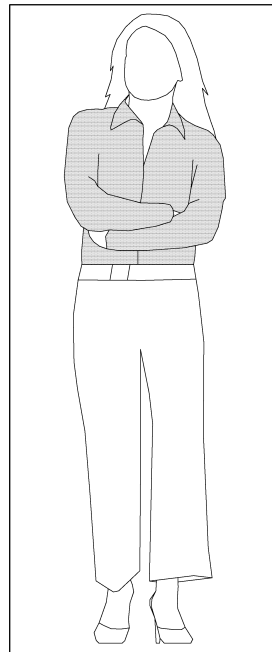
도면8

| purchaseProduct | recommendProduct | purchaseQuantity | recommendQuantity | Confidence |
|-----------------------------|------------------|------------------|-------------------|------------|
| B3128,B6532,B6048 | B3323 | 17 | 17 | 1.00000 |
| B5711,BFFP341,BGST514 | B3323 | 19 | 19 | 1.00000 |
| B6705,BGSP323,BGSB435 | BGSP321 | 18 | 18 | 1.00000 |
| B3128,BGSP326,B6705,BFFP313 | B3323 | 17 | 17 | 1.00000 |
| BGSP317,BGSP327 | BGSP321 | 22 | 21 | 0.95455 |
| BGSO611,BFFN508 | BFFN108 | 21 | 20 | 0.95238 |
| B2805,B6532,B6201 | B3323 | 21 | 20 | 0.95238 |
| B3128,BGMP310,BGMT502 | BGSP321 | 20 | 19 | 0.95000 |
| B1716,B6532,BFFP333 | B3323 | 20 | 19 | 0.95000 |
| B6705,BFFP341,BGST514 | B3323 | 19 | 18 | 0.94737 |
| B1716,B5917,B5919 | B2708 | 19 | 18 | 0.94737 |
| B3128,B5729,B4542,B2708 | B3323 | 19 | 18 | 0.94737 |
| B4295,B2708,B5918 | B5242 | 18 | 17 | 0.94444 |
| B3128,B3879,B5711,B1716 | B3323 | 18 | 17 | 0.94444 |
| B3128,BGSP326,B6705,BFFP341 | B3323 | 18 | 17 | 0.94444 |
| BGSP326,B6705,BFFP313 | B3323 | 29 | 27 | 0.93103 |
| B2805,B1716,B6532 | B3323 | 27 | 25 | 0.92593 |
| BFFP333,B6915 | B1716 | 24 | 22 | 0.91667 |
| BFFS203,BFFN508 | BFFN108 | 36 | 33 | 0.91667 |
| B2805,B3128,B5729,B2708 | B3323 | 24 | 22 | 0.91667 |
| B5711,B1716,B6532 | B3323 | 22 | 20 | 0.90909 |
| B3128,B5711,BGSP326,B6705 | B3323 | 22 | 20 | 0.90909 |
| B2805,BGSP303,B5508 | B3323 | 21 | 19 | 0.90476 |
| B5729,BGSP306,B5508 | B2805 | 21 | 19 | 0.90476 |
| B3128,B5508,BFFP313 | B3323 | 21 | 19 | 0.90476 |
| B5711,B4542,BFFP341 | B3323 | 21 | 19 | 0.90476 |
| B2805,B3128,B6705,B6045 | B3323 | 21 | 19 | 0.90476 |
| B2805,B3323,B6705,BGSB435 | BGSP321 | 21 | 19 | 0.90476 |
| B5917,B6763 | B2708 | 20 | 18 | 0.90000 |
| B2708,BFFN508 | BFFN108 | 20 | 18 | 0.90000 |
| B3128,B3323,B5201 | B2805 | 20 | 18 | 0.90000 |
| B5711,B2708,BGMP310 | B3323 | 20 | 18 | 0.90000 |
| B4542,BFFP341,BGST524 | BGSP321 | 20 | 18 | 0.90000 |

도면9



(a)



(b)