

데이터 마이닝을 이용한 사용자 선호도 기반 상황인지 시스템 설계

안후영^o 박영호

숙명여자대학교 멀티미디어학과

hyahn85@sookmyung.ac.kr, yhpark@sookmyung.ac.kr

A Design of an Context Aware System based on User Preferences using Data Mining Techniques

Hoo-Young Ahn^o Young-Ho Park

Department of Multimedia Science, Sookmyung Women's University

요 약

유비쿼터스 컴퓨팅에서의 플랫폼 기술은 상황인지(context-awareness) 기술과 재구성형 네트워킹(reconfiguration networking) 기술이 융합되어 가면서 병행하여 발전되고 있으며 최근에는 이종간의 다른 리소스를 가지고 있는 모바일 플랫폼간의 자율적인 공유를 통한 보다 큰 개념의 유비쿼터스 서비스가 제공 되고 있다.본 논문은 사용자의 선호도를 고려한 서비스 기술과 서비스 추론 기술을 제안한다. 특히 본 논문에서는 데이터마이닝 기법을 이용하여 사용자 선호도에 기반한 서비스를 제공한다. 본 논문은 상황인지 시스템에서 온톨로지를 활용한 고도화된 서비스 추론 엔진과 함께 데이터마이닝기법을 이용하여 사용자의 과거 이력 분석을 통해 최적의 서비스등 다른 분야의 방법들을 함께 결합시킴으로서 상황 인지 시스템에서의 새로운 사용자 선호도 기반 서비스 패러다임을 제공하는 것을 목적으로 한다.

1. 서론

유비쿼터스 컴퓨팅 환경은 인터넷과 네트워크 환경 다양한 형태의 컴퓨팅 디바이스와 센서들의 발전으로 인해 삶의 형태와 질을 획기적으로 변화시키고 있으며 또 기업의 비즈니스 형태에도 많은 변화를 가져오고 있다 유비쿼터스 컴퓨팅 환경에서의 모든 인간 활동은 네트워크화된 사회 문화적 시스템 하에서 이루어 질 것이다 따라서 이러한 시스템과 인간과의 효율적인 인터페이스가 필수적으로 요청되어 질 것이다. 모바일 기기는 이러한 인간과 유비쿼터스 환경을 효율적으로 연계[1]시켜 줌으로써 사회 문화적인 활동의 효율성을 높여 줄 것이다

모든 컴퓨터, 컴퓨팅 디바이스, 혹은 센서가 네트워크로 서로 연결되어 언제 어디서나 사용할 수 있는 환경을 말하는 유비쿼터스 컴퓨팅 환경으로 인간의 컴퓨팅 환경은 발전해 가고 있다. 이러한 컴퓨팅 패러다임의 변화는 개인형 모바일 플랫폼 기술에도 크게 작용하고 있다.

유비쿼터스 컴퓨팅에서의 플랫폼 기술은 상황인지(context-awareness) 기술과 재구성형 네트워킹(reconfiguration networking) 기술이 융합되어 가면서 병행하여 발전되고 있으며 최근에는 이종간의 다른 리소스를 가지고 있는 모바일 플랫폼간의 자율적인 공유를 통한 보다 큰 개념의 유비쿼터스 서비스가 제공 되고 있다.

본 논문에서는 사용자의 선호도를 고려한 서비스 기술과 서비스 추론[2][3][4] 기술을 제안한다. 특히 데이터마이닝 기법을 이용하여 사용자 선호도에 기반한 서비스를 제공한다. 데이터 마이닝 기법이란 대규모로 저장된 데이터 안에서 체계적이고 자동적으로 통계적 규칙이나 패턴을 찾아내는 것이다. 데이터 마이닝 기법은 통계학 분야에서의 탐색적 자료분석, 가설 검증, 다변량 분석, 시계열 분석, 일반 선형 모형 등의 방법론과 데이터베이스 쪽에서 발전한 OLAP(On-Line Analytic Processing), 인공지능 분야에서 발전한 SOM, 신경망,

전문가 시스템 등의 기술적인 방법론이 쓰인다 특히 데이터 마이닝의 응용 분야로 신용 평점 시스템의 신용평가 모형 개발, 사기탐지 시스템 장바구니 분석, 최적 포트폴리오 구축과 같이 다양한 산업 분야에서 광범위하게 사용되고 있다.

본 논문은 유비쿼터스 환경뿐만 아니라 개인의 성향과 상황정보[5][6]를 고려한 개인화된 멀티미디어 추천 시스템과 같은 다양한 응용에서도 활용이 가능한 기술이므로 사회적 문화적으로 활용 용도가 매우 클 것으로 사료된다.

- 유비쿼터스 환경에 다양한 서비스를 제공하는 상황인지 시뮬레이터 시스템을 구현함
- 기존의 상황인지 서비스 시스템에서는 단순히 디바이스 정보 및 환경 정보를 고려하여 서비스를 제공하였으나 제안하는 시스템에서 사용자의 과거 데이터들에 데이터마이닝 기법을 사용하여 분석함으로써 사용자 선호도 기반 서비스를 제공함
- 본 논문은 유비쿼터스 환경뿐만 아니라 개인의 성향과 상황 정보를 고려한 개인화된 멀티미디어 추천 시스템과 같은 기술로서 활용 용도가 매우 클 것으로 사료됨

본 논문에서 제안하는 방법은 데이터 마이닝 기술에 기반하여 사용자 선호도를 고려한 서비스를 지원할 수 있도록 하며, 사용자 의도 기반 자율서비스를 위한 사용자 선호도 관리 기능을 가능하게 한다 사용자 선호도가 반영된 다양한 서비스들을 생성할 뿐만 아니라 사용자에게 양질의 서비스를 제공하기 위해서 매우 중요한 일이다. 또한, 본 연구는 유비쿼터스 환경뿐만이 아니라 개인의 성향과 상황 정보를 고려한 개인화된 멀티미디어 추천 시스템과 같은 기술로서 활용도가 매우 클 것으로 사료된다. 다음 장에서는 제안하는 방법과 관련된 기존논문들에 대해 소개한다.

본 논문의 구성은 다음과 같다 2장에서는 관련 논문들에 대해 살펴보고, 3장에서는 모바일 오브젝트 상황인식 시스템의 설계에 대해서 설명한다 4장에서는 제안하는 기법을 이용하여 구현한 구현 시스템에서 대해서 기술한다. 마지막으로 5장에서는 결론 및 향후 논문 방향을 제시한다.

2. 관련 연구

상황인지 분야에는 다양한 논문이 있고 상용으로 이용 가능한 상황인지 시스템도 많이 제안되어 있다대

표적인 상용이미지 검색시스템은 MobiHealth, Scalet, ServiceGlobe, Gaia 시스템이 있다.

MobiHealth (Mobile Healthcare)[7]는 Proactive Health Care Systems 기술 분야의 논문으로서 사전 건강관리로 병을 미리 예방 각자의 가정에서 건강관리를 받을 수 있게 하는 시스템이다 그림 1은 MobiHealth 시스템의 구성도를 보인다 제안하는 MobiHealth 시스템은 다섯 개의 방으로 이루어져 있으며 적외선 센서 컴퓨터, 바이오 센서, 비디오 카메라 등으로 구성된다

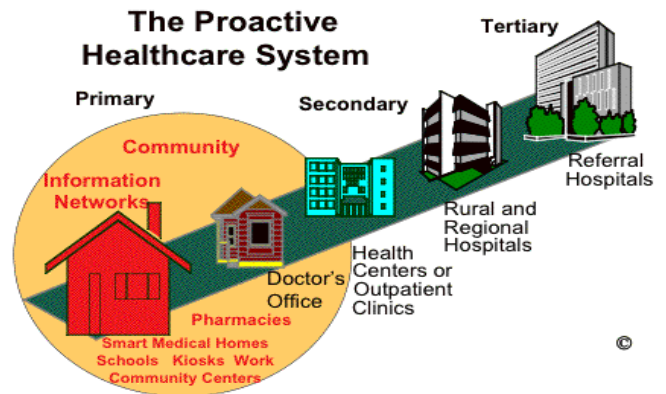


그림 1. University of Rochester의 MobiHealth

이를 통해 집에서 행동하는 여러 가지 모습(걸음 모양, 행동 원형들, 수면포즈, 기타 일반 행위) 에서 컴퓨터는 자료를 수집해 먼저 혈압, 맥박 등 기본적인 건강을 측정한다. 위의 시스템의 장점은 기존의 PC 사용 환경을 모바일 환경에서 서비스 가능하도록 하였다는 것이다. 그러나 환경과의 상호작용 등에 대한 고려가 없다는 단점을 가진다.

Scalet[8]는 Sensor Network 기술 분야의 논문으로서 Scalet에서는 상황인식 컴퓨팅 환경의 모든 측면을 다루고 있지는 않다. 단지, 응용에 상황정보를 어떻게 제공할 것인가에 대한 문제에만 역점을 두고 있다 특히, 이질적인 플랫폼간에 상황정보 전송을 위해 논문을 하였다. Scalet은 SOAP over HTTP와 WSDL을 활용하여 플랫폼간의 호환성을 유지하였다. 센서 노드와 휴대용 PC 간의 상황정보는 Scalet 인프라의 SOAP over HTTP를 통해 전송되고, 가용한 웹 서비스는 WSDL을 이용하여 정의되고 알려진다. 그림 2는 Scalet의 시스템 구조를 보인다. Scalet의 장점은 SOAP over HTTP와 WSDL을 활용하여 플랫폼간의 호환성을 유지하는 것이며 단점은 상황인식 컴퓨팅 환경의 모든 측면을 다루고 있지는 않

다는 것이다.

ServiceGlobe[9]는 고객 단말의 스크린 해상도와 지원 색상수, 네트워크 접속 방법 등과 같은 상황 정보 인식 웹 서비스를 위한 플랫폼으로 Java로 구현 되었으며, XML, SOAP, UDDI, WSDL 등과 같은 표준에 기반하고 있다. 고객의 상황정보는 SOAP over HTTP를 통해 서비스 플랫폼으로 전송된다. 서비스 플랫폼에서는 SOAP 헤더에 포함된 상황정보를 추출하고 처리하고 고객의 상황에 적절한 서비스를 제공한다. 그림 3은 ServiceGlobe의 상황정보 처리 절차를 보인다.

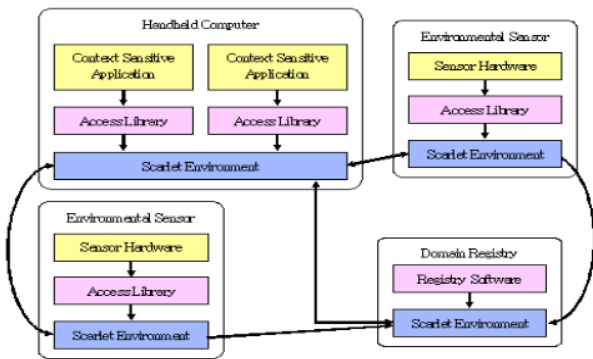


그림 2. Scalet의 구조

ServiceGlobe의 장점은 다양하고 이질적인 고객 단말의 능력과 고객의 위치 등과 같은 상황정보를 고려하여 더 나은 상황인식 웹 서비스를 제공하는 것이며 단점은 고 사양의 컴퓨팅 파워가 요구되며 컨텍스트 정보의 활용이 미비하다는 것이다.

아준다. Context History는 이전 상황정보들을 기록하고 있는 데이터베이스이며 Context Consumer는 상황정보를 사용하는 응용 서비스이다. 그림 3은 GAIA 상황인식 서비스 구조를 보인다. GAIA의 장점은 응용이 다양한 상황정보를 얻고 추론할 수 있게 해주는 것이며 단점은 중앙집중식 데이터베이스를 활용하여 구축된 프레임워크로 이동형 환경에서 적용하기에 취약하다는 것이다.

3. 시스템 설계

본 장에서는 데이터 마이닝 기법을 이용한 지능형 상황인식 시스템의 시스템 설계 방식에 대해 설명한다. 3.1 절에서는 모바일 오브젝트 상황인식 시스템의 구조와 기능에 대해 설명하고, 3.2 절에서는 데이터마이닝 기법을 이용한 사용자 선호도 추출 방식에 대해 설명한다. 3.3 절에서는 사용자 선호도 기반 서비스 제공 방식과 대표적인 사례에 대해 설명한다.

3.1 모바일 오브젝트 상황인식 시스템

모든 컴퓨터, 컴퓨팅 디바이스, 혹은 센서가 네트워크로 서로 연결되어, 언제 어디서나 사용할 수 있는 환경을 말하는 유비쿼터스 컴퓨팅 환경으로 인간의 컴퓨팅 환경은 발전해 가고 있다. 이러한 컴퓨팅 패러다임의 변화는 개인형 모바일 플랫폼 기술에도 크게 작용하고 있다.

특히 본 논문에서 제안하는 상황 인지 기반 자율 서비스를 지원하는 서비스 추론 기술 개발과 사용자 의도 기반 자율서비스를 위한 사용자 선호도 관리 모듈의 기술 고도화 개발은 유비쿼터스 환경에 다양한 서비스들을 생성할 뿐만 아니라 사용자에게 고급의 서비스를 제공한다. 그림 4는 모바일 기기들의 자율서비스 프레임워크의 구조를 보인다.

3.2 데이터마이닝 기법을 이용한 사용자 선호도 추출

본 논문에서는 데이터마이닝 기법을 이용하여 사용자 선호도에 기반한 서비스를 제공한다. 데이터 마이닝 기법이란 대규모로 저장된 데이터 안에서 체계적이고 자동적으로 통계적 규칙이나 패턴을 찾아내는 것이다. 데이터 마이닝 기법은 통계학 분야에서의 탐색적 자료분석, 가설 검정, 다변량 분석, 시계열 분석, 일반 선형 모형 등의 방법론과 데이터베이스 쪽에서 발전한 OLAP(On-Line Analytic Processing), 인공지능 분야에서 발전한 SOM, 신경망, 전문가 시스템 등의 기술적인 방법론이 쓰인다. 특히 데이터 마이닝의 응용 분야로 신용 평점 시스템의 신용평가모형 개발, 사기탐지 시스템

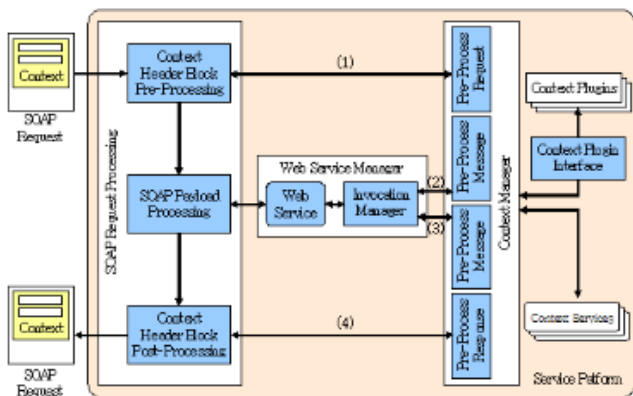


그림 3. 상황정보 처리 절차

GAIA[10]는 Context Provider는 다른 센서 또는 다른 데이터 소스로부터 상황정보를 수집하여 응용에 제공한다. Context Synthesizer는 Context Provider로부터 수집한 상황정보를 상위 개념의 상황정보로 추론하고 추상화하여 응용에 제공한다. Context Provider Lookup Service는 상황정보를 제공하는 Context Provider를 찾

장바구니 분석, 최적 포트폴리오 구축과 같이 다양한 산업 분야에서 광범위 하게 사용되고 있다

3.3 사용자 선호도 기반 서비스

사용자 선호도 기반 서비스 제공이란 데이터 마이닝을 통해 특정 사용자의 과거 데이터들에 대한 분석을 수행하여 선호도에 기반한 서비스를 제공하는 시스템을 의미한다. 사용자 선호도 기반 서비스 제공의 대표 사례는 다음과 같다.

특정 영화관에서는 개개인의 모바일 오브젝트 사용자의 영화 관람 이력을 수집해 놓으며 데이터마이닝 기법을 통해 특정 개인이 선호하는 영화 장르를 도출해 놓는다. 영화 관람 이력 수집은 모바일 오브젝트 사용자가 영화를 관람하고 나왔을 때 관람한 영화에 대한 견해를 문자 서비스 등을 통해 받아 수집한다 추후에 사용자가 해당 영화관에 방문했을 시에 영화관의 스크린에서는 사용자의 모바일 오브젝트 고유 번호를 감지하고 두 기기 간의 통신이 이루어지며 사용자의 근처에 있는 스크린들에서는 특정 개인이 선호하는 장르의 영화들에 대한 상영 정보를 제공해 준다. 특히 스크린과 특정 모바일 오브젝트가 통신하는 방법은 블루투스, 핸드폰 바코드, 핸드폰의 고유 번호 등을 통해 오브젝트들을 식별한다

4. 시스템 구현

본 장에서는 시스템 구현에 대하여 설명한다 4.1 절에서는 시스템 구성도를 설명하고 4.2 절에서는 상황 인지 및 데이터 마이닝 기법을 이용한 지능형 서비스 제공 방식을 설명한다 4.3 절에서는 데이터마이닝을 이용한 선호도 기반 서비스 제공에 대해 설명한다

4.1 시스템 구성도

모바일 오브젝트 상황 인지 기능지원 상황 정보 데이터베이스 프레임워크 기술 개발은 상황 데이터베이스 표준 스펙 개발, 질의 프로세서 개발, 외부 기기들과의 통신을 위한 데이터 포장 과정인 래퍼(wrapper) 및 인코더(encoder) 구현의 세 가지 모듈의 개발을 필요로 한다. 그림 4는 모바일 기기들의 자율서비스 프레임워크의 구조도이다. 본 논문에서는 유비쿼터스 환경에 다양한 서비스를 제공하는 상황인지 시뮬레이터 시스템을 구현하고, 사용자의 과거 데이터들에 데이터마이닝 기법을 사용하여 분석함으로써 사용자 선호도 기반 서비스를 제공한다.

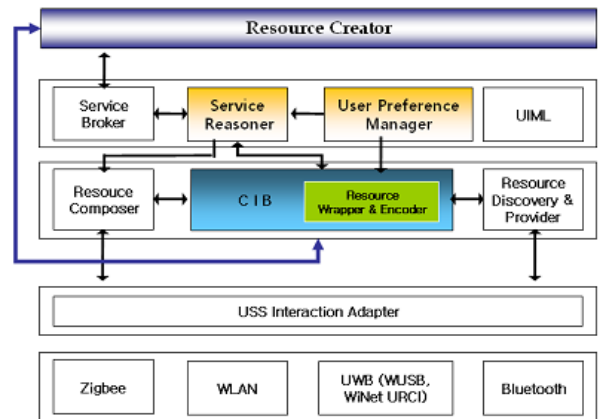


그림 4. 시스템 구성도

유비쿼터스 지능공간의 모바일 오브젝트 상황인지 기술 개발에 있어서 모바일 오브젝트들의 상황인지 기술로서 서비스 추론 기능의 고도화 작업과 사용자 선호도 관리 기능의 고도화 작업이 필요하다 또한 모바일 오브젝트의 사용자가 과거에 사용한 디바이스의 정보와 같은 과거 사용 이력들을 기술하여 데이터마이닝 기법 [11][12]을 통한 최적의 사용자 선호도를 반영한 서비스를 제공하도록 한다

기존 방식의 지능형 서비스 명세는 서비스나 응용 개발자에 의해 작성되어 서비스 기반구조에서 실행된다 이러한 방식은 동적 환경에서 개별 사용자에게 최적화된 서비스를 제공하기가 어려울 뿐만 아니라 사용자에게 대한 정보가 공유됨으로 인해서 보안상의 문제가 발생할 수 있다. 따라서 서비스를 제공받는 개인 사용자에게 의한 서비스 시나리오의 작성과 사용자 단말기에서 필요한 지능형 서비스를 실행하는 새로운 서비스 제공 방식을 제공할 수 있다.

4.2 상황 인식 시스템

유비쿼터스 시스템에서의 모바일 기기들의 자율서비스 프레임워크를 위한 상황인지 기반 기술개발을 위해서는 서비스 추론 소프트웨어 설계 및 개발, 사용자 선호도 관리 소프트웨어 설계 및 개발 모바일 오브젝트 상황 인지 기능지원 상황 정보 데이터베이스 프레임워크 기술 개발 분야로서 크게 세 개의 모듈들이 필요하다

서비스 추론 소프트웨어 설계 및 개발에서는 주어진 상황에 적합한 서비스 디바이스를 자동 검색하는 기능이 필요하다. 따라서 온톨로지를 이용한 서비스 분석 가용 디바이스 정보 요청 질의어 생성 사용자 선호도를 반영한 규칙 정의 및 관리, 선호도 정보를 고려한 디바이스 추론 등의 기능을 구현해야 한다

사용자 선호도 관리 소프트웨어 설계 및 개발에서는 모바일 오브젝트의 사용자가 과거에 사용한 디바이스들의 정보와 같은 과거사용 이력들을 상황 데이터베이스에 기술하며, 위의 데이터들을 사용자의 프로파일과 선호도 정보 관리해 주는 소프트웨어를 말한다

그러나 위의 기능들을 개발하는데 있어서는 사용자 선호도 관리에 있어서 표준 및 획득 방법의 부재 사용자 선호도 관리에 있어서 표준 및 획득 방법의 부재 모바일 플랫폼에서 적용 가능한 상황인지를 위한 알고리즘의 부재, 모바일 콘텍스트에 대한 정형화된 표현 방법의 부재라는 문제점들이 있다 특히 사용자의 과거 사용 이력들을 분석하여 위의 기법들을 개발하는데 있어서는 사용자 선호도 관리에 있어서 표준 획득 방법이 부재 한다는 문제점을 가진다

따라서 본 논문에서는 수집된 사용자의 모바일 오브젝트 과거 이력을 이용하여 데이터마이닝 기법[13]을 적용한 표준화된 사용자 선호도 관리[14] 방법을 제안한다

4.3 데이터마이닝을 이용한 사용자 선호도 추출

데이터 마이닝을 이용한 선호도 기반 서비스 제공은 다음과 같이 이루어진다. 데이터 마이닝의 계급 분류(classification hierarchies), 연관 규칙(associative rules), 순차 관계(sequential patterns), 시간의 흐름에 따른 방식(patterns within time series), 분류 및 구분(categorization and segmentation) 방식을 사용하여 사용자 선호도를 분석하여 다음과 같은 서비스를 제공한다

자원 관리(Service-Oriented Resource Management) 서비스를 제공한다 다양한 스마트 객체들이 제공하는 서비스들은 이질적인 환경에서 동작한다 서비스 개발자들이 이러한 이질적인 환경에 대한 모든 정보를 획득하고 인지하는 것은 불가능하기 때문에 스마트 객체에 대응하는 서비스 수준에서의 관리가 필요하다 따라서 스마트 객체를 서비스 수준으로 추상화하고 서비스 접근 제어를 통한 서비스 및 자원관리 서비스를 제공한다

또한 커뮤니티 기반의 코디네이션 메커니즘(Community 기반 Coordination Mechanism)으로서 커뮤니티에 참여하는 멤버들 간의 상호 동작을 지원하고 커뮤니티 멤버들 간의 상호동작을 조율하기 새로운 코디네이션 메커니즘에 서비스를 제공한다

따라서 본 논문에서는 상황 인지 시스템에 위의 세 가지 특성을 접목시키는 독창적인 아이디어이므로 논문

주제 및 방법은 창의성과 독창성이 매우 높다고 할 수 있다.

5. 결론 및 향후 논문

제안하는 방법은 유비쿼터스 지능공간의 모바일 오브젝트 상황인지 기술 개발에 있어서 모바일 오브젝트들의 상황인지 기술로서 서비스 추론 기능의 고도화 작업과 사용자 선호도 관리 기능의 고도화 기술을 논문한다또한 모바일 오브젝트의 사용자가 과거에 사용한 디바이스의 정보와 같은 과거 사용 이력들을 기술하여 데이터마이닝 기법을 통한 최적의 사용자 선호도를 반영한 서비스를 제공하도록 한다

본 논문은 상황인지 시스템에서 온톨로지를 활용한 고도화된 서비스 추론 엔진과 함께 데이터마이닝기법을 이용하여 사용자의 과거 이력 분석을 통해 최적의 서비스 등 다른 분야의 방법들을 함께 결합시킴으로서 상황인지 시스템에서의 새로운 사용자 선호도 기반 서비스 패러다임을 제공한다

6. 참고문헌

- [1] Choonsung Shin, Hyoseok Yoon, Woontack Woo, "User-Centric Conflict Management for Media Services Using Personal Companions," *ETRI Journal*, Vol.29, No.3, 2007
- [2] Grobelnik M. Fortuna, B. and D. Mladenic, "Background knowledge for ontology construction", in Proc. *Contexts and Ontologies*, 2006
- [3] G. Chen and D. Kotz., "A survey of context-aware mobile computing research". *Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College*, November 2000
- [4] D. Estrin, D. Culler, K. Pister, and G. Sukhatme. Connecting the physical world with pervasive networks. *IEEE Pervasive Computing*, vol.1, No.1, pp.59-69, January-March 2002
- [5] A. Schmidt and K. Van Laerhoven. How to build smart appliances. *IEEE Personal Communications*, pp.66-71, 2001
- [6] M. A. Orgun W. Lin and G. J. Williams.,

- "Multilevels hidden markov models for temporal data mining," in Proc. *the KDD 2001 Workshop on Temporal Data Mining San Francisco, CA, USA, August 2001*
- [7] "MobiHealth", <http://www.futurehealth.rochester.edu>
- [8] "Scalet", <http://www.iit.edu>
- [9] "Gaia", <http://www.uiuc.edu>
- [10] <http://www-db.in.tum.de/research/projects/sg>
- [11] Rakesh Agrawal , Tomasz Imieliński , Arun Swami, "Mining association rules between sets of items in large databases," in proc. *the 1993 ACM SIGMOD international conference on Management of data*, p.207-216, May 25-28, 1993, Washington, D.C., United States
- [12] Ramakrishnan Srikant , Rakesh Agrawal," Mining Generalized Association Rules," in proc. *the 21th International Conference on Very Large Data Bases*, p.407-419, 1995
- [13] Raymond T. Ng , Laks V. S. Lakshmanan , Jiawei Han , "Alex Pang, Exploratory mining and pruning optimizations of constrained associations rules," in proc. *the 1998 ACM SIGMOD international conference on Management of data*, p.13-24, June, 1998, Seattle, Washington, United States
- [14] Brian Lent , Arun N. Swami , Jennifer Widom, "Clustering Association Rules," in proc. *the Thirteenth International Conference on Data Engineering*, pp.220-231, April , 1997