

Ubiquitous Computing and Statistics; What's the Connection?

Sung-Hae Jun¹⁾ and Hongsuk Jorn²⁾

Abstract

Mark Weiser introduced ubiquitous computing in his article titled 'The computer for 21st Century' in 1991. This has been new paradigm after internet. Now, the rapid development of mobile computer, wireless network, and intelligent system has supported ubiquitous computing environment. In the related area of information science, the researchers have studied on ubiquitous computing. But in the field of Korea statistics, this research has not been worked yet. So, we proposed the connection between statistics and ubiquitous computing in this paper. As an example, we showed an efficient cache hoarding for ubiquitous computing using statistical methods. In experimental results, we verified our proposed issue.

Keywords : Ubiquitous Computing; Cache Hoarding; Collaborative Filtering.

1. 서론

1990년대 이후 전 세계는 다양한 분야에서 인터넷의 직접적인 영향을 받고 있다. 인터넷은 세상의 컴퓨터가 각자 고유한 주소(IP address)를 가지고 연결되어 있는 거대한 네트워크이다. 21세기가 되면서 이러한 인터넷의 패러다임은 빠르게 유비쿼터스 컴퓨팅(ubiquitous computing) 환경으로 바뀌고 있다(Hunter, 2002). 유비쿼터스 환경은 2000년대 정보 기술(information technology) 분야에서 중요한 연구과제의 하나로써 떠오르고 있다. 최근 들어와서 통계학은 지능형 시스템(intelligent system), 생명정보학(bioinformatics) 등 인접 학문에 많은 도움을 주고 있다. 특히, 데이터마이닝(data mining) 프로젝트에서 통계학 전공자들은 IT 산업에서 필요로 하는 많은 일들을 하고 있다. 여론 조사, 품질 관리, 생물 및 의학 데이터 분석, 마케팅 등 많은 분야에서 통계학이 활용되고 있지만, 특히 최근에는 정보 통신 분야에서도 통계학의 필요성이 더욱 증가되고 있다.

2003년에 들어서면서 국내에서도 한국정보과학회를 비롯한 많은 정보과학 관련 학회에서 유비쿼터스 특집을 다루면서 유비쿼터스의 패러다임 환경에서 정보과학 관련 학문들이 나아가야 하는 방향에 대한 논의가 심도 있게 이루어지고 있다. 2004년 1월 현재 국내 통계학과와 절반이 훨씬 넘는 대학에서 정보통계학, 정보분석학 등으로 통계학 전공 명칭에 정보과학 분야가 추가되면서 전통적인 통계학 분야에 정보기술이 추가되고 있다. 이러한 변화의 시점에서 인터넷 이후 정보 기술 분야의 새로운 패러다임들 중의 하나로서 강력하

1) Full Time Lecturer, Department of Statistics, Chongju University, Chungbuk 360-764, Korea.

E-mail : shjun@cju.ac.kr

2) Professor, Department. of Statistics, Inha University, Incheon 402-751, Korea.

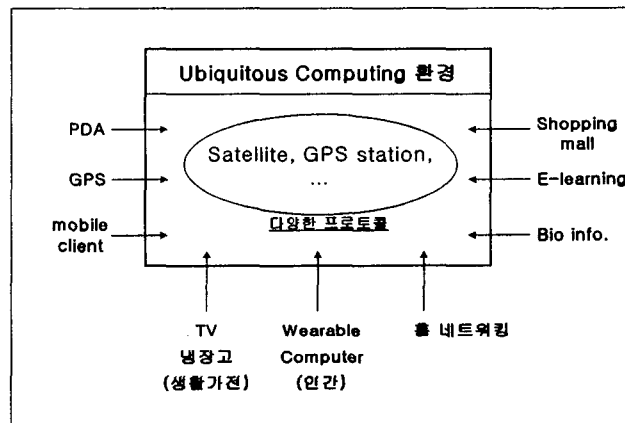
E-mail : hsjorn@anova.inha.ac.kr

고 대두되고 있는 유비쿼터스 컴퓨팅 환경에 통계학이 접목되어 이바지할 수 있는 가능성의 일부를 본 논문에서는 찾아보고, 실험으로 가능성을 확인하였다. 즉, 통계학이 21세기 정보기술의 새로운 주자로서 주목받고 있는 유비쿼터스 환경에서 어떠한 역할을 담당할 수 있고, 이러한 역할을 수행하기 위해서는 어떠한 연구가 이루어져야 하는지에 대한 연구 방향을 제안하였고 이를 통해 정보과학 분야에서 통계학의 중요성을 확인하였다.

2. 유비쿼터스 컴퓨팅

2.1 유비쿼터스 컴퓨팅의 정의

제록스(Xerox)사의 와이저(M. Weiser)는 유비쿼터스 컴퓨팅을 ‘실세계에 존재하는 모든 개체들을 기능적, 공간적으로 연결하여 사용자에게 필요한 정보나 서비스를 실시간으로 제공할 수 있는 기반 기술’로 정의하였다(Weiser, 1991; www1). 유비쿼터스 컴퓨팅의 정의에서 모든 개체에는 물론 인간도 포함된다. 인간은 자신이 가질 수 있는 컴퓨터(wearable computer)를 이용하여 언제든 유비쿼터스 환경에 일원이 될 수 있다. 물론 생활 가전 기기의 작은 컴퓨터부터 고성능 워크스테이션까지 모든 컴퓨터들도 이에 해당 된다. 미래에는 크고 작은 다양한 컴퓨터들이 현실세계에서 유, 무선 통신망으로 연결되어 사용자가 필요로 하는 정보를 즉시 제공할 수 있는 환경이 완벽하게 구현될 것이다. 현재는 개인 휴대단말기(PDA, personal digital assistant), 위치 측정시스템(GPS, global positioning system) 등과 같은 모바일 컴퓨터로서 유비쿼터스 환경에 대한 출발이 이루어지고 있다. 이를 위해서는 사용자가 시간적으로나 공간적으로 어떠한 거부감이나 불편함이 없이 편리하게 정보 서비스를 제공 받을 수 있는 기술 환경이 이루어져야 한다(장세이 외, 2003; Weiser, 1993). 다음 그림은 현재의 유비쿼터스 컴퓨팅 환경에 대한 개념도이다. 이러한 개념도는 정보기술의 발전에 따라 얼마든지 바뀔 수 있게 된다.



<그림 1> 유비쿼터스 컴퓨팅 환경

위의 그림을 통해 서로 다른 기종의 컴퓨터들이 다양한 프로토콜(protocol)로 연결되어 인공위성이나 GPS 서버를 통하여 서로간의 정보를 주고받을 수 있게 됨을 알 수 있다. 외출 중에도 인간은 자신의 PDA를 통하여 가정에서 일어나는 일들을 통제할 수 있게 되고 또한 자신이 주로 이용

하는 인터넷 쇼핑몰로부터 자신만을 위한 정보 서비스도 수시로 받게 되며, 자신의 자동차에 장착된 컴퓨터로부터 사고 발생시 자동으로 가장 가까운 응급 구조 센터로 연락을 취할 수 있게 된다. 이러한 것들은 유비쿼터스 컴퓨팅 환경에서의 극히 일부의 것에 지나지 않는다. 이보다도 훨씬 다양한 서비스를 유비쿼터스 환경으로부터 제공받을 수 있게 된다.

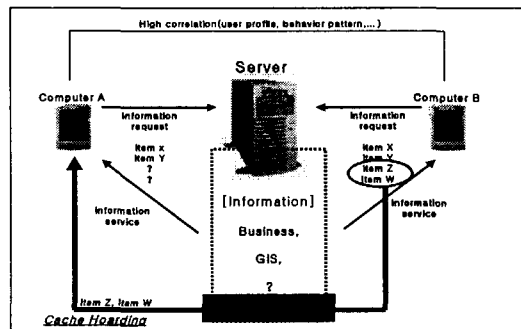
2.2 통계학과 관련된 최근 연구 동향

유비쿼터스 컴퓨팅에 관한 연구는 거의 모든 분야에서 연구되어 지고 있기 때문에 일일이 열거하기는 어렵다. 유비쿼터스 컴퓨팅은 컴퓨터가 어디에 존재하든, 성능이 우수하든 안 하든 인간도 알지 못하는 상황에서 인간의 삶을 지배할 수 있게 된다. 따라서 유비쿼터스 환경에서 동작하는 컴퓨터의 소프트웨어들은 궁극적으로 각 개인의 행위에 대한 패턴을 파악하여 개인별 특성에 따라 서로 다른 서비스를 할 수 있어야 한다. 즉 행위 데이터에 대한 분석을 통하여 지능화된 정보 서비스를 할 수 있게 된다. 본 논문에서는 유비쿼터스 컴퓨팅 환경에서의 통계학의 역할에 관한 제안이기 때문에 이 절에서의 연구동향은 주로 통계학과 관련이 있는 연구 분야에 대하여 알아보았다. 유, 무선 네트워크 환경에서 인간은 과거와는 비교할 수 없을 정도로 다양하고 풍부한 정보를 제공받고 있지만 아직도 인간은 네트워크 환경에서 자신들이 필요로 하는 정보를 검색하기 위하여 불필요한 노력을 많이 하고 있다. 불필요한 노력을 하게 되는 가장 큰 이유는 정보의 양이 기하급수적으로 증가하고 이와 비례하여 더욱 복잡해지고 있지만 인간의 정보처리 능력은 한계를 보이고 있다. 따라서 인간의 작업량을 줄여주고 또한 인간보다도 정교하게 정보를 검색하고 동시에 인간과의 상호 작용이 아주 유연한 지능형 시스템이 연구, 개발되고 있다(백성욱, 2003; Joerding, 1998; Mase, 2002; Writtig, 1995). 국내 통계학 분야에서도 모바일 환경에서 클라이언트들의 효과적인 캐시 적재를 위한 학습 기법에 대한 연구가 이루어지고 있다(전성해 외, 2003).

3. 향후 유비쿼터스의 발전과 통계학의 역할

유비쿼터스 환경에서 인간을 포함하여 네트워크에 연결된 클라이언트들의 정보 공유 및 서비스를 위한 자동화된 지능 시스템을 위하여 통계학의 역할은 크게 데이터 마이닝 기술로 요약할 수 있다. 데이터 마이닝은 접속 고객의 소비성향, 사기유형 등의 비즈니스 지능화 뿐만 아니라, 서버와의 네트워크 단절 기간에도 GPS와 같은 모바일 클라이언트가 마치 서버에 연결되어 있는 것처럼 정보 서비스를 받을 수 있도록 네트워크 단절 전에 미리 클라이언트 캐시에 필요한 정보를 방송(broadcasting)할 수 있는 규칙의 구축 등과 같은 것들은 모두 통계적 모형 기반의 데이터 마이닝 기술이 필요하게 된다. 비즈니스 지능화를 위한 데이터 마이닝 기법으로는 로지스틱 회귀모형, 군집화 기법, 신경망(neural networks) 등이 있고, 또한 클라이언트 캐시 적재를 위한 기법으로는 연관성 규칙(association rules), 협업 추천 모형(collaborative filtering)등이 있다. 이 밖에도 대부분의 통계 분석 기법들이 적용될 수 있을 것이다. 이제는 비즈니스 거래 데이터를 포함하여 많은 트랜잭션(transaction) 데이터가 테라바이트(terabyte) 단위를 넘어서고 있다. 때문에 대용량 데이터 분석을 통한 자동화된 규칙 생성을 위한 통계학의 역할은 더욱 필요하게 될 것이다. 본 논문에서는 눈부시게 발전하고 있는 유비쿼터스 컴퓨팅 환경에서 통계의 역할 중에서 이동 컴퓨터를 위한 효과적인 캐시 적재에서의 통계학의 역할에 대하여 알아보았다. 최근의 컴퓨터 하드웨어 기술의 발달에 의해 노트북, PDA 등 사용자가 이동하면서 사용할 수 있는 작은 컴퓨터의 생산이 가

능해 졌다. 이러한 휴대용 컴퓨터들은 어느 위치에서도 전역적 데이터 서비스에 접속이 가능한 무선 통신 장치를 갖추고 있다. 하지만 네트워크 연결에 대한 비용문제 등 여러 가지 문제점 때문에 계속적으로 네트워크와 연결되지 못하고 단절되는 경우가 종종 발생하게 된다. 하지만 이때에도 모바일 클라이언트는 계속적인 서비스 요구가 발생하게 되고 이러한 것들을 클라이언트 캐시가 도와주어야 한다. 이 때 해당 사용자에게 필요한 아이템들이 해당 모바일 컴퓨터의 캐시에 적재되어 있어야 한다. 또한 모바일 컴퓨터가 정적이거나 이동 가능한 모바일 서버의 많은 공유 데이터베이스에 효율적으로 접속할 수 있게 하는 모바일 데이터베이스 시스템에 대한 연구가 최근까지 대단히 많이 이루어지고 있다. 특히 모바일 환경에서의 클라이언트-서버 컴퓨팅에 대한 연구도 이루어졌다(Alonso, 1993; Chrysanthis, 1993; Dunham et. al., 1995; Hunter, 2002). 사용자는 이동 중에 인터넷 서핑이나 이메일 확인을 할 수 있다. 그러나 무선통신을 하는 모바일 컴퓨터는 무선통신의 비용과 무선 네트워크를 사용할 수 없는 경우의 발생 때문에 종종 네트워크와 단절된다. 가령 사용자가 모바일 컴퓨터를 가지고 비행기를 타게 되는 경우처럼 말이다. 단절 될 때의 클라이언트 캐시에서 필요한 파일을 적재하는 방법에 대한 연구(Kistler et. al. 1992)와 사용자의 간섭 없이 자동으로 클라이언트의 캐시에 데이터를 적재해 주는 방법에 대한 연구(Kuenning et. al. 1997)가 현재 수행되고 있다. 개인의 모바일 컴퓨터가 네트워크와 단절 기간 동안에도 사용자들은 단절 이전과 똑같은 정보 서비스를 받고 싶어 한다. 어떤 경우에는 사용자는 현재 자신의 모바일 컴퓨터가 네트워크와 단절된 상황인지 아닌지 모를 정도로 정보 서비스를 계속적으로 받았으면 한다. 이러한 계속적인 정보 서비스가 가능하기 위해서는 각자의 모바일 컴퓨터의 캐시에 적재된 정보에 의해 수행되어지는 단절 중의 작업을 위한 효과적인 캐시 적재 모형의 필요성이다. 캐시 적재 프로세스는 미래에 단절되기 전에 클라이언트 캐시에서 필요로 하는 데이터 아이템들의 적재하는 것이다. 그러나 사용자가 미래에 어떤 파일을 필요로 하게 될지를 서버가 아는 것은 매우 어렵다. 또한 이러한 작업은 자동화 되어져야 한다. 자동화 적재는 사용자의 어떠한 간섭이 없이도 적재 집합을 예측하여 만들어 내는 프로세스이다. 효과적인 캐시 적재를 위하여 통계적 모형에 기반한 데이터 마이닝 전략을 사용할 수 있다. 데이터 마이닝 기법을 이용하면 대량의 트랜잭션 데이터로부터 숨겨진 패턴을 찾아내어 새로운 현상이나 미래의 행위를 예측하는 모델링 작업을 수행할 수 있다. 모바일 컴퓨터가 네트워크와 단절이 발생했을 때 이 클라이언트의 미래의 요청에 대한 아이템 예측을 위한 규칙을 생성하기 위하여 본 논문에서는 모바일 클라이언트의 서비스 요청(request) 이력(history) 데이터에 협업 추천(collaborative filtering) 모형의 적용을 제안하였다.



<그림 2> 모바일 클라이언트의 캐시 적재를 위한 추천 시스템

협업 추천은 기본적으로 클라이언트들이 아이템의 요구에 대한 정보를 기반으로 하여 특정 클라이언트가 특정 아이템에 대한 유용성(utility) 혹은 선호도(preference)를 예측해 내는 것을 목적으로 한다. 본 논문에서 제안 적용하는 협업 추천은 <그림 2>와 같은 구조를 지니고 있다. <그림 2>는 모바일 클라이언트의 캐시 적재를 위한 협동 추천 시스템의 구조를 나타내고 있다. 즉 서로 유사한 모바일 클라이언트의 아이템 요구 이력 정보를 이용하여, 미래의 네트워크 단절에 대비한 캐시 적재를 수행하게 된다. 특히 본 논문에서는 클라이언트의 요구 이력 정보 뿐만 아니라 모바일 클라이언트의 위치 정보까지 고려하였다. 기존의 연구에서는 모바일 클라이언트의 요청 아이템에 대한 예측 모형을 위한 학습 데이터로서 현재까지의 클라이언트의 요구 이력 정보만을 사용하였다(Saygin et. al. 2000).

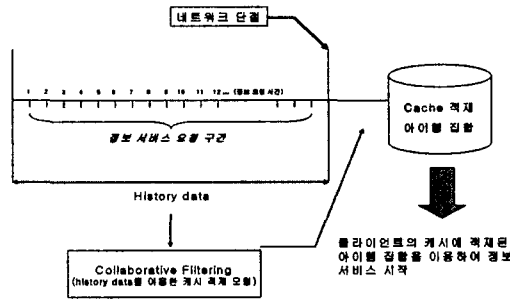
4. 실제 적용 예

유비쿼터스 컴퓨팅 환경에서 통계적 모형의 적용 예를 살펴보기 위하여 본 논문에서는 협업 추천 모형을 이용한 모바일 클라이언트의 효과적인 캐시 적재 모형을 구축하였고, 실험을 통하여 기존 모형들의 성능비교를 수행하였다. 기본적인 유비쿼터스 컴퓨팅 환경인 현재의 무선 네트워크에서는 지역적 통신 두절, 시스템적 대역폭의 축소 등 여러 가지 이유로 잦은 네트워크 단절이 발생한다. 이로 인한 이동 컴퓨팅의 문제점들을 해결하기 위한 하나의 전략으로서 네트워크 단절을 대비하여 사전에 네트워크에 연결된 모바일 클라이언트의 캐시에 향후에 필요로 할 정보를 미리 적재한다. 현재 주로 사용되고 있는 적재 방법은 각 클라이언트 컴퓨터가 가장 최근에 서버에게 요청했던 정보를 서버로부터 클라이언트에 방사(broadcasting)하는 LRU(least recently used)방식이다. 하지만 이러한 캐시 적재 전략은 실시간으로 변화하는 네트워크 환경과 사용자의 정보 요구에 대한 변화를 잘 반영하지 못하여 결국 네트워크 단절 상황에서 사용자에게 꼭 필요한 정보를 효과적으로 제공해 주지 못하게 된다. 본 논문에서는 이러한 상황에서 정보 아이템들 간의 상호 의존성에 기반한 협업추천 모형을 적용하여 기존의 방법에 비해 효과적인 정보 서비스가 가능하도록 하였다. 특정 모바일 클라이언트에 대한 필요 정보만 찾아내어 해당 클라이언트 컴퓨터에 적재하면 이 모바일 컴퓨터를 사용하는 사용자에게는 만족할 만한 정보 서비스를 기대할 수 있게 된다. 아울러 저장 공간의 제약을 갖는 모바일 컴퓨터의 한계도 극복할 수 있게 된다. 본 논문의 적용 실험에서는 CSIM Simulator(www2)를 이용하여 이동 컴퓨터의 과거 정보 요구에 대한 이력 데이터를 모의 생성하여 사용하였다. Cache hit ratio(14)를 이용한 객관적인 성능 평가를 통하여 제안 기법을 적용한 캐시 적재 성능이 기존의 것에 비해 우수함을 보였다. 비교 측도는 사용자가 필요로 하는 정보가 모형에 의해 만들어진 적재 데이터 집합에 얼마나 많이 있는지를 측정하는 것이다. 실험을 위하여 우선 모바일 클라이언트가 필요로 하는 정보 아이템의 전체 집합은 다음 식과 같이 정의하였다. I_N 은 전체 N 개의 원소로 이루어진 아이템 집합이다.

$$I_N = \{i_1, i_2, \dots, i_N\}, \quad i_j \in I_N \quad (4.1)$$

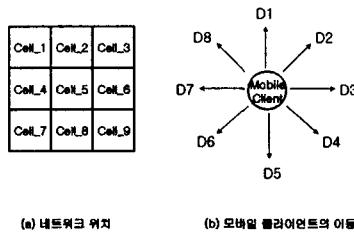
식 (4.1)에서 i_j 는 j 번째에 해당하는 아이템을 나타낸다. 본 실험에서는 편의상 전체 아이템 집합의 크기 N 을 9로 하였다. <그림 3>은 이러한 이력 정보 데이터(history data)의 구조이다. <그림 3>은 캐시 적재 아이템들의 이력 데이터의 구조와 캐시 정보 서비스의 과정을 나타내고 있다. 이력 데이터는 모바일 클라이언트가 네트워크에 연결되어 있는 동안 서버에게 요청했던 정보 아이

템들을 가지고 있다. 이것을 이용하여 협업 추천 모형을 통한 모바일 클라이언트의 캐시 적재 데이터 집합(hoarding data set)을 결정할 수 있게 된다.



<그림 3> 이력 정보의 데이터 구조

즉 모형을 구축하는 일정 시점 이전까지의 클라이언트의 요청(request) 아이템들의 이력을 이용하여 미래 시점에서 모바일 클라이언트의 캐시에 적재할 아이템들을 결정하게 된다. 이 때 클라이언트 요청 로그(client request log) 데이터는 연결 기간과 단절 기간(disconnected and connected periods)으로 나뉘어져 있는 구조가 된다. 추가적으로 본 논문에서는 모바일 컴퓨터의 이동성(mobility) 정보를 클라이언트 캐시의 적재 집합 결정에 반영하기 위하여 모바일 컴퓨터의 위치 정보(각도, 방향, 속도 등)와 모바일 컴퓨터가 현재 위치하는 주변 환경의 정보 등을 고려하여 제안 기법을 이용하여 모바일 컴퓨터의 캐시에 적재하는 아이템들을 예측하여 서버가 방사(broadcasting)할 수 있도록 하였다. <그림 4>는 클라이언트의 위치 정보에 대한 개념이다.



<그림 4> 모바일 클라이언트의 위치 정보의 구조

<그림 4>의 (a)는 이 실험에서의 전체 네트워크를 9개의 셀로 나누었다. 각 셀을 대표하는 아이템은 <표 1>과 같이 결정하였다.

<표 1> 각 셀과 해당되는 대표 아이템

| | | | | | | | | | |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| cell | Cell_1 | Cell_2 | Cell_3 | Cell_4 | Cell_5 | Cell_6 | Cell_7 | Cell_8 | Cell_9 |
| item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

<표 1>에 의해 만약 특정 클라이언트가 Cell_3에 위치하게 되면 아이템 3이 추가로 더 캐시에 적재되어야 한다. 그리고 (그림 4)의 (b)는 특정 모바일 클라이언트가 다음 셀로 이동하게 되는 방향을 D1부터 45도 간격으로 D8까지 8방향으로 이동할 수 있도록 결정하였다. 이들 정보가 (그림 3)의 이력 정보에 의한 협업 추천 모형의 결과와 결합하여 단절되기 전에 모바일 컴퓨터의 캐시에 필요한 아이템들을 적재할 수 있게 한다. 실험에서는 각 모바일 클라이언트의 위치를 D1에서 D8로 순서적으로 이동하는 것으로 가정하였다. 앞에서 언급한 바와 같이 성능 평가는 캐시 적중률을 이용하였다. 즉 단절 기간 동안 클라이언트가 요구가 얼마나 자동화 적재(automatic hoarding) 기법에 의한 캐시 적재 아이템 집합에 있는지를 비교한다. 본 논문의 실험을 위한 시뮬레이션 데이터는 XMM-Newton SOC의 CSIM Simulator 소프트웨어를 이용하여 생성하였다(www2). 다음의 <표 2>은 본 논문의 실험에서 사용된 이력 데이터의 일부를 보여주고 있다.

<표 2> 모의 생성된 이력 데이터

| 실험회수 | training data (400개) | test data (100개) |
|------|-------------------------|-------------------------|
| 1 | 1,2,9,1,6,...,3,4,6,3,6 | 8,1,6,7,6,...,2,2,5,8,4 |
| 2 | 7,9,6,9,5,...,9,2,5,8,6 | 1,8,5,3,1,...,1,2,7,7,2 |
| ! | ! | ! |
| 99 | 3,2,5,9,8,...,5,6,3,5,6 | 9,9,2,3,5,...,3,1,2,3,8 |
| 100 | 9,8,8,7,7,...,6,4,4,5,6 | 6,4,8,7,9,...,9,1,2,6,6 |

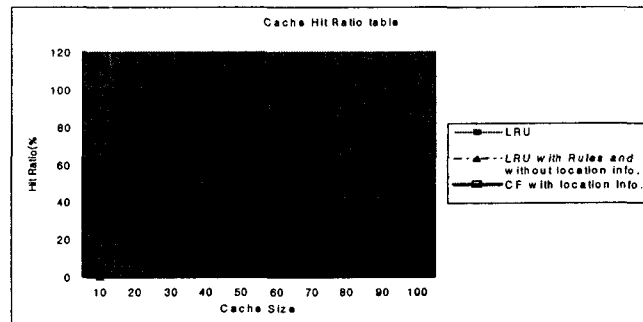
<표 2>은 총 100번의 실험을 위한 이력 데이터를 나타내고 있다. 전체 생성 데이터의 80%는 학습(training)에 이용하고 나머지 20%는 모형의 타당성을 검증하는 테스트(testing) 데이터로 이용하였다. 구체적인 협업 추천을 위한 방법으로는 통계적 상관 계수와 코사인 유사도(cosine similarity)를 이용하였다(Linden et. al. 2003). 다음의 <표 3>은 비교 모형들 간의 적중률에 대한 평균과 표준편차를 나타내고 있다.

<표 3> 적중률의 평균(mean)과 표준편차(s.d.)

| 가능한 캐시 저장 공간 | [1] LRU | [2] LRU with rules and without location info. | [3] CF with location info. |
|-----------------|-----------------|--|----------------------------|
| | mean (s.d.) | mean (s.d.) | mean (s.d.) |
| 10 | 0.0000 (0.0000) | 0.1961 (0.0147) | 0.2951 (0.0067) |
| 20 | 0.0612 (0.2419) | 0.2621 (0.0133) | 0.2877 (0.0071) |
| 30 | 0.1198 (0.3812) | 0.3465 (0.0258) | 0.4012 (0.0069) |
| 40 | 0.2531 (0.5011) | 0.4511 (0.0261) | 0.4951 (0.0078) |
| 50 | 0.3977 (0.7292) | 0.5541 (0.0278) | 0.5647 (0.0091) |
| 60 | 0.6445 (1.0527) | 0.6459 (0.0311) | 0.6719 (0.0143) |
| 70 | 0.7995 (1.3198) | 0.8248 (0.0392) | 0.8343 (0.0192) |
| 80 | 0.8488 (1.3742) | 0.8588 (0.0411) | 0.8577 (0.0230) |
| 90 | 0.9451 (1.4351) | 0.9466 (0.0501) | 0.9643 (0.0258) |

위의 표에서 가능한 캐시 저장 공간은 총 저장할 수 있는 아이템의 수를 나타낸다. 예를 들어 30은 저장 공간에 저장할 아이템의 수가 30개 밖에 안 된다는 것이다. 즉 제한된 저장 공간의 크기를 나타낸다. 동일한 저장 공간에서 가장 적중률이 높은 것은 제안 모형이었다. 물론 <표 3>의 [2]의 Saygin(2000) 전략도 좋은 결과를 보이고 있다. 하지만 제안 모형에 비해 표준 편차가 크게 나타나고 있어서 모형의 안정성 면에서는 제안 모형에 비해 떨어지고 있다. 캐시 저장 공간이 100

이 되면 저장 공간이 모든 아이팀 집합을 저장할 수 있게 되기 때문에 의미가 없다. 위의 실험결과에서 실질적으로 의미가 있는 것은 적은 캐시 저장 공간에서의 적중률이다. 왜냐하면 현재 유비쿼터스의 이동 컴퓨터는 늘 제한된 저장 공간만을 갖기 때문이다. 다음 그림은 비교 모형들 간의 적중률의 평균값을 시각적으로 표현하였다.



<그림 5> 3개의 모형간의 Cache hit ratio의 비교 결과

<그림 5>는 모바일 클라이언트의 캐시 저장 공간의 크기에 따른 필요 아이팀들의 적중률을 보여주고 있다. 가로 축의 캐시 저장 공간의 크기가 100이 되면 필요한 모든 아이팀들을 모두 저장할 수 있는 공간이 되기 때문에 모형간의 비교에 대한 의미가 없게 된다. 하지만 10, 20, 그리고 30과 같이 캐시의 저장 공간이 제한되어 있는 경우에는 각 모형에 의해 결정되어진 서로 다른 아이팀들이 저장되고 이들로부터 적중률을 계산하면 서로 다르게 된다. <표 3>과 <그림 5>에 의하면 제안 방법이 기존 방법들에 비해 보다 큰 캐시 적중률 값을 갖게 됨을 알 수 있었다.

5. 결론

본 논문은 인터넷 이후의 정보기술의 새로운 패러다임인 유비쿼터스 컴퓨팅 환경에서 통계학이 어떠한 방향으로 어느 정도의 역할을 담당할 수 있을지에 대한 가능성을 살펴보고, 아울러 간단한 실험을 통하여 실제 적용 예를 단계적으로 살펴보았다. 연관성 기반의 협업 추천 모형을 이용한 모바일 클라이언트의 효율적인 캐시 적재 모형은 네트워크의 단절 동안의 효과적인 정보 서비스 문제와 모바일 컴퓨터의 적은 캐시 용량 문제를 해결할 수 있는 하나의 방법을 제공하였다. 다변량 분석 기법, Support Vector Machine, 그리고 통계적 학습 모형과 같은 더욱 정교한 통계 기법을 적용하여 이미 부분적으로 이루어지고 있고 앞으로 더욱 보편화될 유비쿼터스 컴퓨팅 환경에서 통계학의 역할이 클 것으로 생각된다. 많은 적용과 이에 대한 결과는 향후 연구과제로 남긴다.

참고문헌

- [1] 백성욱 (2003). 유비쿼터스 컴퓨팅 환경을 위한 지능형 미디어 기술, 「정보과학회지」, 제21권 제5호, 36-42.

- [2] 장세이, 우운택 (2003). 유비쿼터스 컴퓨팅 환경을 위한 센싱 기술과 컨택스트-인식 기술의 연구 동향, 「정보과학회지」, 제21권 제5호, 18-28.
- [3] 전성해, 류제복, 이승주 (2003). Data Mining Approach to Supporting Hoarding in Mobile Computing Environments, 「한국통계학회 2003 춘계 학술대회 논문집」, 13-18.
- [4] Alonso R. and Korth H.F. (1993). Database system issues in nomadic computing, In *Proceedings of the ACM SIGMOD International Conference on the Management of Data*.
- [5] Linden, G., Smith, B. and York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering, *IEEE Internet Computing*, Vol. 7, Issue 1, 76-80.
- [6] Chrysanthis P.K. (1993). Transaction processing in mobile computing environment, In *Proceedings of the IEEE Workshop on Advances in Parallel and Distributed Systems*.
- [7] Dunham M.H. and Helal A.S. (1995). Mobile computing and databases: Anything new?, *SIGMOD Record*, 24(4), 5-9.
- [8] Hunter R. (2002). World Without Secrets: Business, Crime, and Privacy in The Age of Ubiquitous Computing, John Wiley and Sons.
- [9] Jing J., Helal A. and Elmagarmid A.K. (1999). Client server computing in mobile environments, *ACM Computing Surveys*.
- [10] Joerding T. and Meissner K. (1998). Intelligent multimedia presentations in the Web: fun without annoyance, *Computer Network and ISDN Systems*, No. 30, 649-650.
- [11] Kistler J.J. and Satyanarayanan M. (1992). Disconnected operation in the coda file system, *ACM Transactions on Computer Systems*, 10(1), 3-25.
- [12] Kuenning G. and Popek G. (1997). Automated hoarding for mobile computers, In *Proceedings of the ACM Symposium on Operating Systems Principles*, St. Malo, France.
- [13] Mase K. (2002). Intelligent Interfaces for Information Agents: Systems, Experience, Future Challenges, *Lecture Note in Artificial Intelligence*, No. 2446, 10-13.
- [14] Saygin Y., Ulusoy O. and Elmagarmid A. (2000). Association Rules for Supporting Hoarding in Mobile Computing Environments, *IEEE 10th International Workshop on Research Issues on Data Engineering*.
- [15] Weiser M. (1991). The computer for 21st Century, *Scientific American*, Vol. 265, No. 3, 94-104.
- [16] Weiser M. (1993). Some computer science issues in ubiquitous computing, *Communications of ACM*, Vol. 36, No. 7, pp. 75-84.
- [17] Wittig H. and Griwodz C. (1995). Intelligent media agents in interactive television systems, *International Conference on Multimedia Computing and Systems*, 182-189.
- [18] www1, <http://www.ubiq.com/weiser>.
- [19] www2, <http://www.atl.lmco.com/proj/csim>.