

홈 네트워크 시스템 상에서 모바일 멀티미디어 서비스를 위한 효과적인 데이터베이스 설계 방안

송혜주[†] · 박영호^{**} · 김정태^{***} · 백의현^{****}

요약

최근 멀티미디어 콘텐츠의 영향력이 점차 커지면서, 자신이 소유한 모바일 디바이스를 통해 멀티미디어 서비스를 제공받고자 하는 사용자의 요구가 증가되고 있다. 본 논문에서는 향후 홈 네트워크 시스템에 제공될 모바일 멀티미디어 서비스를 위한 데이터베이스 설계기법을 제안한다. 이를 위해 먼저, 서비스 운용 시 요구되는 에트리뷰트들을 추출하여 릴레이션을 구축하고 데이터베이스를 디자인한다. 특히, 데이터베이스 운영 시 발생할 수 있는 중복이나 갱신이상 현상을 막기 위하여, 릴레이션에 존재하는 다치 종속성(MVD, Multi Valued Dependency)을 밝혀내고 정규화(Normal Form) 이론에 기반 한 데이터베이스를 디자인을 제안한다. 실험에서는 다양한 정규형을 만족하는 분할된 릴레이션의 질의 수행결과를 통해, 데이터 중복현상과 갱신이상의 발생을 비교, 분석함으로써, 제안한 방법이 효율적임을 입증한다.

키워드 : 홈 네트워크, 홈네트워크 서비스, 데이터베이스, 다치 종속성

An Efficient Database Design Method for Mobile Multimedia Services on Home Network Systems

Hye-Ju Song[†] · Young-Ho Park^{**} · Jung-Tae Kim^{***} · Eui-Hyun Paik^{****}

ABSTRACT

Recently, users who want to be provided mobile devices, such as PDP, PMP, and IPTV connected wireless internet, with multimedia contents are increasing due to an influence of multimedia contents. In the paper, we propose an efficient database design method for managing mobile multimedia services on home network systems. For this, we build relations using attributes required while providing multimedia services, and then design a database. Specially, we propose a database design method based on normalization theory to eliminate redundancies and update anomalies caused by a non trivial multi valued dependency in relations. In the experiments, we compare and analyze occurrence frequencies of data redundancies and update anomalies through query executions on the relation decomposed into normal forms. The results reveal that our database design is fairly effective.

Key Words : Home Network, Home Network Service, Database, MVD(Multi-Valued Dependency)

1. 서론

최근 멀티미디어 콘텐츠의 영향력이 점차 커지면서, 자신이 소유한 모바일 디바이스를 통해 멀티미디어 서비스를 제공받고자 하는 사용자의 요구가 증가되고 제공받고 있다. 홈 네트워크 시스템에서 멀티미디어이다. 현재 대부분의 사용자들은 필요한 멀티미디어 콘텐츠를 전문적인 멀티미디어 서비스 공급업체로부터 서비스를 제공받는다면, 가족 구성원이 이용한 모든 멀티미디어 서비스들을 홈 네트워크 공급

업체에서 통합 과금할 수 있고, 멀티미디어 콘텐츠를 별도의 멀티미디어 서비스 업체에 추가가입 없이 가족 구성원이 편리하게 이용할 수 있다는 장점을 가진다.

기존의 홈 네트워크에서 데이터베이스를 활용한 연구들은 홈 네트워킹 된 기기의 내장형 데이터베이스 관리시스템이나, 다양한 정보 가진 기기들에서 발생하는 데이터와 가진 기기 및 센서들의 운용상태를 관리, 제어, 조작하는데 집중되어왔다. 기기 중심의 데이터베이스는 모바일 디바이스나 멀티미디어 콘텐츠 등을 위한 데이터베이스 설계를 전혀 고려하지 않았으므로, 멀티미디어 서비스를 효과적으로 관리할 수 없다. 또한, 현재 홈 네트워크 서비스는 하나의 홈 네트워크 아이디를 이용해 모든 서비스가 가족 구성원이 사용하는 유무선 기기들에게 제공되고 있다. 이러한 환경에서 멀티미디어 서비스를 위한 데이터베이스를 디자인 할 때,

* 본 연구는 한국전자통신연구원(ETRI)에서 홈 네트워크 위탁과제의 지원으로 수행되었습니다.

† 준회원 : 숙명여자대학교 멀티미디어학과 석사과정

** 정회원 : 숙명여자대학교 멀티미디어학과 조교수

*** 정회원 : 한국전자통신연구원 유비쿼터스 홈서비스 연구팀 연구원

**** 정회원 : 한국전자통신연구원 유비쿼터스 홈서비스 연구팀 팀장

논문접수 : 2007년 4월 6일, 심사완료 : 2007년 6월 18일

각 요소들 간 존재하는 관계성을 고려하지 않는다면, 릴레이션 상태를 일관성 있게 유지하고, 관련된 데이터들간의 독립성을 유지할 수 없다. 따라서, 구체적인 멀티미디어 서비스의 유형을 분류하고, 이를 기록, 관리 할 수 있는 데이터베이스의 디자인 방안이 필요하다.

본 논문의 목적은 구체적인 사용자 서비스를 효과적으로 지원하기 위해 서비스 운영 중 발생하는 데이터를 효과적으로 관리하는 방법을 제안하는 것이다. 이를 위해 가족 구성원들 각각에게 제공되는 멀티미디어 서비스와 이들이 소유한 모바일 디바이스에 제공되는 서비스를 중복 없이 관리할 수 있는 제 4 정규형(4NF, Fourth Normal Form)[1, 2, 3]기반의 데이터베이스 디자인을 제시하고자 한다. 제 4 정규형은 BCNF(Boyce-Codd Normal Form)의 기준을 포함하는 스키마로 의미 있는 다치 종속성(Non-trivial Multi-valued Dependency)이 존재하지 않는다. 이는 기존의 홈 네트워크의 기기정보뿐만 아니라 멀티미디어 콘텐츠 서비스 관련 정보를 결합하여 사용자 중심의 멀티미디어 홈 네트워킹 서비스를 제공할 수 있는 효과적인 방안이다.

본 논문에서는 홈 네트워크 시스템에서 사용자 중심의 효과적인 멀티미디어 서비스를 위한 데이터베이스 디자인에 관하여 다음과 같은 공헌을 제시한다.

- 홈 네트워크 시스템에서 새로운 멀티미디어 서비스를 제공하기 위한 데이터베이스 적용방안을 제안: 이를 위하여, 멀티미디어 콘텐츠의 새로운 속성들을 저장, 검색, 갱신할 수 있는 효과적인 데이터베이스 관리방안을 제안한다.
- 멀티미디어 서비스를 체계적으로 관리하기 위한 데이터베이스 디자인 시, 다치 종속성(MVD, Multi-Valued Dependency)[1, 2]을 제거한 결합 없는 데이터베이스 디자인 방안을 제안: 이를 위하여, 제 4 정규형 기반의 데이터베이스 설계 기술을 소개한다.
- 바람직하지 못한 튜플(Tuple)의 발생 예를 실험을 통해 성능 검증: 이를 위하여, 다양한 정규형을 만족하는 릴레이션의 성능평가를 수행하여 제안된 방법의 사용이 릴레이션 설계 시 서비스 데이터의 중복을 제거할 수 있는 방법임을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 홈 네트워크의 기본적인 시스템 구성도와 멀티미디어 서비스에 대해 소개하고, 멀티미디어 서비스의 종류와 홈 네트워크에서 데이터베이스의 중요성에 대하여 논의한다. 3장에서는 멀티미디어 서비스를 위한 효과적인 데이터베이스 스키마 설계를 제안하고, 4장에서는 데이터베이스의 실험을 통해 성능평가결과를 제시한다. 5장에서는 홈 네트워크에서의 관련된 연구를 소개한다. 마지막으로 6장에서는 본 논문의 결론을 내린다.

2. 연구배경

2.1 홈 네트워크 시스템에서의 멀티미디어 서비스

홈 네트워크 시스템은 그 구성이 매우 다양하다. 본 논문

에서는 홈 네트워크와 홈 네트워크에서 멀티미디어 서비스를 다음과 같이 정의하고자 한다.

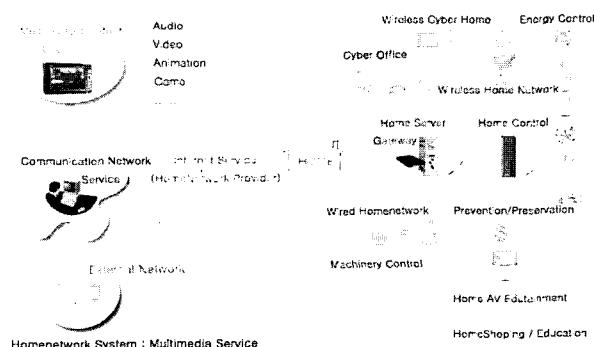
정의 1. 홈네트워크[4]는 초고속 네트워크 인프라를 기반으로 다양한 IT기술이 접목되어 서비스를 창출하는 복합 산업 분야로 가정 내의 정보가전 기기가 네트워크로 연결되어 기기, 시간, 장소에 구애받지 않고 다양한 서비스를 제공할 수 있는 디지털 가정환경이다.

정의 2. 멀티미디어 서비스(Multimedia Service)는 초고속 인터넷 망을 이용하여 기존의 다채널 방송프로그램과 비디오, 그래픽, 오디오, 게임, 애니메이션과 같은 멀티미디어 콘텐츠를 가정의 TV, 컴퓨터 등의 유무선 디바이스에 제공하는 서비스이다.

홈 네트워크 시스템은 멀티미디어 서비스를 제공하는 제공자와 사용자 간의 유무선 통신선로, 유무선 사업자, 인터넷업자, 가전기기업자, 통신업자, 건축업자들이 서로 공존하며 시스템을 구축한다. 이러한 환경에서 구체적인 멀티미디어 서비스를 제공하기 위한 다양한 디바이스의 토폴로지(topology)가 가능하다.

(그림 1)은 이러한 연결 형태의 예를 보이고 있다. 홈 네트워킹[4]의 기본 구조는 내부와 외부 네트워크를 연결하는 홈 게이트웨이(Home Gateway), 전화선·전력선·무선 등 가정 내 통신망, 정보기기를 제어하며 상호 연동시키는 미들웨어, 홈 네트워킹 기능이 추가된 정보기기 등으로 구성된다. 사용자들은 서비스를 제공받는 홈과 서비스 제공자들이 초고속 네트워크로 연결되어있어, 홈 게이트웨이를 통해 주방기기 제어, 에너지 관리, 사이버 오피스, 방범, 보안 등의 다양한 서비스를 제공받는다. 댁내에 설치된 홈 게이트웨이는 가정 내의 모든 센서 및 가전제품들과 커뮤니케이션을 통해 장비들 간의 상이한 프로토콜을 엮어주는 교량역할을 수행 한다.

사용자는 댁내의 네트워크화 된 가전기기들과 무선으로 연결된 디바이스로 멀티미디어 콘텐츠 사업자들이 제공하는 다양한 서비스를 제공받는다. (그림 1)의 시스템 구성처럼, 홈 네트워크 서비스를 제공하는 인터넷 사업자가 통신기간 사업자의 초고속 인터넷 선로를 이용하여 멀티미디어를 중심의 서비스를 댁내에 제공하는 역할을 한다. 가정에서 멀티미디어 서비스를 제공받기 위해서는 먼저 멀티미디어 콘텐츠가 저장된 외부 네트워크에서 홈서버로 콘텐츠를 전송



(그림 1) 홈 네트워크 시스템에서 멀티미디어 서비스 구성도

한다. 홈 서버에서는 해당 콘텐츠 서비스를 요청한 디바이스로 콘텐츠를 다시 전송하는 과정을 통해 사용자가 멀티미디어 서비스를 제공받을 수 있다.

3. 관련연구

본 장에서는 홈 네트워크 시스템에서 데이터베이스를 활용한 연구를 소개하고, 효과적인 스키마 설계에 대하여 언급한다.

3.1 홈 네트워크 시스템에서 데이터베이스를 활용한 연구

홈 네트워크 시스템은 다수의 정보가전기들로부터 전달되는 다량의 정보를 효과적으로 저장할 수 있는 데이터베이스와, 데이터베이스에 저장된 정보를 이용하여 통계분석, 고장진단, 정보 조회 등을 인터넷을 통해 서비스 할 수 있는 다양한 종류의 응용 프로그램으로 구성된다. 현재까지 홈 네트워크의 연구들은 주로 HTTP 프로토콜을 사용하여 홈 네트워크 기기간의 제어와 상호 운용하는 UPnP(Universal Plug and Play)[15, 17]와 Java를 기반의 Jini[16]등과 같은 미들웨어의 호환이나 네트워크를 응용한 기기간의 네트워크 연동, 유무선 통신기술, 데이터 전송규격, 보안유지 등에 관련하여 중점적으로 이루어져왔다.

홈 네트워크 시스템에서 데이터베이스를 활용한 연구로는, 네트워크화 된 기기의 내장형 데이터베이스 관리시스템[9]이나, 다양한 정보 가진 기기들에서 발생하는 데이터와 운영상태를 관리, 제어하는 방식이 대부분이었고, 멀티미디어 서비스 제공시 발생할 수 있는 구체적인 정보를 데이터베이스에 저장, 검색, 갱신, 관리 하기 위해 데이터베이스를 적용한 사례는 없었다. 따라서 본 논문에서는 홈 네트워크 시스템에서 멀티미디어 콘텐츠 서비스를 제공하기 위한 사용자 중심의 데이터베이스 디자인을 제안한다.

3.2 홈 네트워크 시스템에서 효율적인 스키마 설계에 대한 연구

미래에 존재할 홈 네트워크 시스템은 구체적인 서비스 제공과 일관성 있는 데이터 관리를 위하여 좋은 스키마 디자인을 우선적으로 고려해야 한다. 잘못된 스키마는 튜플(tuple)에서 중복된 정보나 갱신 이상, 가짜 튜플(spurious tuple)의 문제를 발생시켜 많은 비용 지출을 초래할 수 있다. 그러나, 현재까지 홈 네트워크 시스템에서 스키마를 적용한 연구는 없었다. 따라서 본 논문에서는 구체적인 사용자 서비스를 중복 없이 제공하기 위해 다치 종속성(MVD, Multi-Valued Dependency)을 제거한 결함 없는 데이터베이스 디자인 방안을 제안 하였다는 점이 기존연구와 다르다. 4.2절에서 효과적인 데이터베이스 디자인에 관하여 논하도록 한다.

4. 멀티미디어 서비스를 위한 데이터베이스 디자인

홈 네트워크 서비스를 이용하는 사용자는 모바일 디바이스들로 멀티미디어 콘텐츠를 제공받기 위하여, 서비스에 대

한 정보를 저장, 갱신, 검색, 관리 할 수 있는 데이터베이스가 필요하다. 기존에는 기기나 상태정보를 위한 데이터베이스만 존재하므로, 본 장에서는 멀티미디어 서비스를 제공한다고 가정할 때, 실제 요구되는 항목들로 애트리뷰트들을 추출하여 새로운 데이터베이스를 구축한다. 이 때, 데이터베이스는 홈 네트워크 사업자나 멀티미디어 콘텐츠 사업자 등, 누가 관리하느냐에 따라서 상이한 스키마로 설계될 수 있다. 본 장에서는 멀티미디어 사업자들의 콘텐츠를 홈 네트워크 서비스를 통해 제공해 줄 수 있는 멀티미디어 콘텐츠 제공업자 중심의 데이터베이스를 디자인한다. 데이터베이스는 향후 실제 서비스에서의 적용이 가능하다.

홈 네트워크 서비스를 이용하는 사용자는 여러 모바일 디바이스들로 멀티미디어 콘텐츠를 제공받기 위하여, 서비스에 대한 정보를 저장, 갱신, 검색, 관리 할 수 있는 데이터베이스가 필요하다. 이 때, 데이터베이스는 홈 네트워크 사업자나 멀티미디어 콘텐츠 사업자 등, 누가 관리하느냐에 따라서 상이한 스키마로 설계될 수 있다. 본 장에서는 멀티미디어 사업자들의 콘텐츠를 홈 네트워크 서비스를 통해 제공해 줄 수 있는 멀티미디어 콘텐츠 제공업자 중심의 데이터베이스를 디자인한다. 본 장에서 설계하는 데이터베이스의 구성은 홈 네트워크 시스템에서 멀티미디어 서비스를 제공한다고 가정할 때, 실제 요구되는 항목들로 애트리뷰트들을 추출하여 데이터베이스를 구축하며, 향후 실제 서비스에서의 적용이 가능하다.

4.1 데이터베이스 구성

멀티미디어 서비스를 위한 데이터베이스에는 가입자의 정보, 모바일 디바이스, 디바이스 별 가입 서비스, 멀티미디어 서비스 이용에 관한 정보들이 존재한다. 홈 네트워크 시스템에서 사용자는 PDA(Personal Digital Assistant)나 노트북과 같은 다양한 모바일 디바이스로 멀티미디어 서비스를 제공받을 수 있다. 이를 위해, 사용자가 모바일로 멀티미디어 서비스를 제공받기 위한 디바이스 릴레이션MM_DEVICE이 요구된다. 기본키(Primary Key)로 홈 네트워크 아이디(HN_id)와 디바이스 아이디(Device_id)를 가지며, 디바이스 이름(Device_name), 디바이스의 상태(Device_status), 비디오와 같은 서비스 이용시 필요한 디바이스 해상도(Resolution) 정보를 가진다.

OWNER_INFO 는 사용자에 관한 정보를 가지고 있는 릴레이션이다. 릴레이션을 구성하는 애트리뷰트로는 사용자가 속한 맥내 홈네트워크 아이디(HN_id), 사용자 아이디(Own_id), 사용자 주민등록번호(SSN), 사용자 이름(Own_name)을 가진다. 기본키는 HN_id과 Own_id이다. 사용자는 자신에게 부여된 아이디로 서비스를 이용하므로, 주민등록 번호 정보를 이용하여 미성년자의 성인물 등의 접속을 금지할 수 있다.

DEVICE_SERVICE는 각 디바이스별 가입한 서비스 타입과 사용자 데이터를 가지는 릴레이션으로 Ser_type은 각 디바이스에서 이용할 수 있는 서비스 타입을 의미한다. 기본키는 HN_id, Device_id, Own_id, Ser_type이다.

MM_SERVICE는 멀티미디어 서비스 이용내역에 관한 릴레이션이다. 릴레이션을 구성하는 애트리뷰트로 서비스 번호

(Serial_num), 네트워크 서비스 제공업체 아이디와 이름(HSP_id, HSP_name), 콘텐츠 아이디와 이름(Cont_id, Cont_name), 콘텐츠 제공업자 아이디(Cont_prov_id), 콘텐츠 장르(Cont_type), 단위 가격(Unit_price), 단위(Unit), 서비스 시작시간(Running_start), 서비스 이용시간(Running_time)을 가진다. 여기서 가격단위란, 서비스 이용요금을 시간단위로 과금할 것인가, 재생 단위 등으로 할 것인가에 관련된 가격책정의 기준을 의미하며, 서비스 타입은 비디오, 오디오등의 타입으로 구분될 수 있다. 기본키는 HN_id, Device_id, Own_id, Serial_num 애트리뷰트들의 조합으로 구성된다.

(그림 2)의 MM_SERVICE 릴레이션은 부분 종속성(Partial Dependency)이 존재하지 않기 때문에 제 2정규형(2NF, Second Normal Form)[11,12]을 만족한다. 그러나 제 3정규형(3NF, Third Normal Form)을 만족하지 못한다. HSP_id를 통해 이행적 종속성(Transitive Dependency)[11, 13]이 발생하기 때문이다. 제 3정규형은 릴레이션 스키마 R에서 후보키(Candidated Key)나 어떤 키의 부분집합이 아닌 애트리뷰트들의 집합 Y가 존재하여 $X \rightarrow Z$ 와 $Z \rightarrow Y$ 가 만족될 때, 이행적 종속성인 $X \rightarrow Y$ 를 허용하지 않는다. 여기서 X, Y, Z는 R의 애트리뷰트들의 부분집합이며 X와 Y 사이의 함수적 종속성을 $X \rightarrow Y$ 와 같이 표기한다[12]. MM_SERVICE 릴레이션에서는 식 (1)과 같은 이행적 종속성이 성립한다.

$$\{HN_id \rightarrow HSP_id, HSP_id \rightarrow HSP_name\} \\ \models HN_id \rightarrow HSP_name \quad (1)$$

홈 네트워크 아이디가 네트워크 서비스 업체를 결정하므로, 종속성 $HN_id \rightarrow HSP_name$ 는 HSP_id를 통해 이행적임을 알 수 있다. 이와 같은 함수적 종속성(Functional Dependency)이 존재하면, 문제가 발생한다. 서비스가 사용자에게 제공될 때마다, 서비스에 관련된 튜플에서 같은 정보를 반복하게 되는 중복과 이러한 릴레이션에서 발생하는 갱신이상 문제이다. 따라서 하나의 릴레이션을 새로운 릴레이션들로 분해하여 문제가 있는 함수적 종속성을 제거해야 한다.

(그림 3)는 (그림2)의 릴레이션 MM_SERVICE를 제 3정규형으로 정규화 한 모습이다. MM_SERVICE는 사용한 서비스 정보, HN_PROVIDER는 네트워크 서비스 제공업체의 정보, MM_CONTENTS는 멀티미디어 콘텐츠에 관한 구체적인 정보이다. MM_SERVICE릴레이션의 외래키인 Cont_id

MM_DEVICE	(HN_id, Device_id, Device_name, Device_status Resolution)
OWNER_INFO	(HN_id, Own_id, SSN, Own_name)
DEVICE_SERVICE	(HN_id, Device_id, Own_id, Ser_type)
MM_SERVICE	(HN_id, Device_id, Own_id, Serial_num, HSP_id, HSP_name, Ser_type, Cont_id, ContName, Cont_prov_id, Cont_type, Unit_price, Unit, Running_start, Running_time)

(그림 2) 홈 네트워크에서의 데이터베이스 스키마

MM_SERVICE	(HN_id, Device_id, Own_id, Serial_num, HSP_id, Ser_type, Cont_id, Running_start, Running_time)
HN_PROVIDER	(HSP_id, HSP_name)
MM_CONTENTS	(Cont_id, ContName, Cont_prov_id, Cont_type, Unit_price, Unit)

(그림 3) 제 3정규형으로 정규화 된 릴레이션

는 MM_CONTENTS를 참조한다. 릴레이션에서의 밑줄은 기본키를 구성하는 애트리뷰트를 의미한다. 이와 같은 스키마는 이행적 종속성이 제거되었으므로 제 3정규형을 만족한다. 제 3정규형을 만족하는 스키마는 이행적 종속성이 제거되어 같은 정보를 반복하는 중복을 줄일 수 있다.

4.2 다치 종속성이 제거된 데이터베이스 디자인

지금까지 함수적 종속성과 관련한 데이터베이스 디자인에 대해서 논의하였다. 그러나 많은 경우 릴레이션들은 함수적 종속성만으로 표현할 수 없는 제약조건들을 가진다. 본 절에서는 구성원들이 소유한 다중 디바이스에 제공되는 멀티미디어 서비스를 제공하기 위해 다치 종속성(MVD, Multi-Valued Dependency)을 제거한 데이터베이스를 디자인 한다. 그 이유는 다치 종속성이 존재하는 경우, 사용자에게 제공되는 구체적인 서비스를 구분해 낼 수 없기 때문이다. 또한, 중복된 정보로 인한 갱신이상 현상(Update Anomaly Problem)은 데이터의 일관성과 신뢰성을 결여 시킬 수 있다. 이러한 문제를 해결하기 위해 제 4 정규형 기반의 데이터베이스 디자인을 제안한다.

제 4 정규형(4NF, Fourth Normal Form)[1, 2, 3]은 BCNF(Boyce-Codd Normal Form)의 기준을 포함하는 스키마로 의미 있는 다치 종속성(Non-trivial Multi-valued Dependency)이 존재하지 않는다. 릴레이션 스키마 R의 함수적 종속성 및 다치 종속성들의 집합을 F라 하고, F를 만족하는 R의 모든 릴레이션 인스턴스에서 성립하는 함수적 종속성 또는 다치 종속성들의 완전한 집합 F^+ 가 있다. 만약, F에 대한 폐쇄 F^+ 의 모든 의미 있는 다치종속성 $X \twoheadrightarrow Y$ 가 존재할 때, X가 R의 슈퍼키(Super Key)이면 R은 F에 대한 제 4 정규형이라고 정의된다[10]. 여기서 R은 릴레이션 스키마이며, X, Y는 R의 부분집합이다. 또한, X와 Y 사이의 다치종속성을 $X \twoheadrightarrow Y$ 와 같이 표현한다. 다치 종속성[1, 2, 8, 14]은 한 릴레이션 내에서 세 애트리뷰트 그룹이 존재할 때 일어난다. 릴레이션 스키마 R의 임의의 릴레이션 상태를 r이라 할 때, r에 대해 $t1[X] = t2[Y]$ 를 만족하는 두 튜플이 존재한다면, 식 (2)와 같은 성질을 가진 두 개의 튜플 $t3$ 와 $t4$ 도 반드시 존재해야한다[10]. 여기서 Z는 $R - (X \cup Y)$ 이다.

$$t3[X] = t4[X] = t1[X] = t2[X] \\ t3[Y] = t1[Y] \text{ and } t4[Y] = t2[Y] \\ t3[Z] = t2[Z] \text{ and } t4[Z] = t1[Z] \quad (2)$$

$X \twoheadrightarrow Y$ 가 성립하면 대칭성에 의해 $X \twoheadrightarrow Z$ 가 성립하며, Y와 Z는 서로 독립적인 관계이다. 이 때, X가 Y와 Z를 다중 결정(Multidetermine)한다고 말하며, $X \twoheadrightarrow Y \mid Z$ 와 같이 표기한다. 위의 정의는, X에 의해 결정되는 Y값들은 오직 X에 의해서만 결정되며, R의 나머지 애트리뷰트들인 Z의 값들과는 관계없다는 것을 나타낸다. 따라서 임의의 두 튜플이 다른 Y 값을 갖지만 동일한 X값을 가진다면 동일한 X값을 가지는 각 Z값들에 대해 이 Y값들을 중복해야 한다. 상기 정의를 (그림 4)의 DEVICE_SERVICE 릴레이션에 적

용하면, 식 (3)과 같은 두개의 독립적인 일대다(1:N) 관계가 존재함을 알 수 있다.

$$\{HN_id, Device_id\} \rightarrow \{Own_id \mid Ser_type\} \quad (3)$$

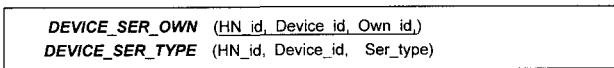
택내의 모바일 디바이스는 여러 사용자를 가질 수 있으며, 디바이스에서 이용할 수 있는 멀티미디어 서비스의 타입 역시 비디오, 오디오 등 다양하기 때문에 다치 종속성이 발생한다. 이와 같이 두 개 이상의 독립적 애트리뷰트들이 같은 릴레이션 스키마에 존재하면, 튜플에 대한 연산이 다른 튜플에 영향을 미치고, 릴레이션의 의미가 달라질 수 있다. 예를들어, 택내의 모바일 디바이스에 새로운 서비스를 추가한다면, 사용자와 서비스 타입간의 독립성을 유지하기 위해 각 사용자마다 튜플을 삽입해야 한다. 이러한 릴레이션은 중복이 발생할 뿐 아니라, 하나라도 빠뜨리면 다치 종속성을 위반하고, 일관성을 잃게 된다. 다치 종속성을 제거하기 위해서는, DEVICE_SERVICE와 같은 릴레이션 스키마를 허용하지 않는 제 4 정규형 기반의 데이터베이스 설계기술이 필요하다.

(그림 4)는 릴레이션 DEVICE_SERVICE을 제 4 정규형을 만족하는 두개의 릴레이션으로 분해한 것이다. 두개의 릴레이션은 식 (4)와 같은 집합을 형성한다.

$$\begin{aligned} &DEVICE_SER_OWN \\ &= \{ HN_id, Device_id \cup Own_id \} \\ &DEVICE_SER_TYPE \\ &= \{ DEVICE_SER_OWN - Own_id \} \end{aligned} \quad (4)$$

이것은 다치 종속성인 HN_id, Device_id → Own_id를 기반으로 분해한 것으로, 그 분해집합은 비부가적 조인(Nonadditive Join) 특성을 가진다. (그림 5)와 같이 DEVICE_SERVICE 릴레이션을 제 4 정규형 기반의 DEVICE_SER_OWN과 DEVICE_SER_TYPE 릴레이션으로 분해하면, 다치 종속성이 제거되어 저장 공간을 절약하고 갱신이상을 피할 수 있다. 제 4 정규형을 만족하는 릴레이션은 BCNF(Boyce-Codd Normal Form) 또한 만족한다.

제 5정규형[10]은 스키마를 두개의 스키마로 비부가적 조



(그림 4) 제 4 정규형으로 정규화 된 릴레이션

Set 1: Query1:	SELECT *	FROM	MM_SERVICE;
Query2:	DELETE	FROM	MM_SERVICE
		WHERE	HN_id = 'HN02';
	SELECT *	FROM	MM_SERVICE;
Set 2: Query1:	SELECT *	FROM	MM_SERVICE;
Query2:	SELECT *	FROM	MM_SERVICE;
	SELECT *	FROM	HN_PROVIDER;
	SELECT *	FROM	MM_CONTENTS;
Set 3: Query1:	SELECT *	FROM	DEVICE_SERVICE;
Query2:	SELECT *	FROM	DEVICE_SER_TYPE;
	SELECT *	FROM	DEVICE_SER_OWN;

(그림 5) 질의집합

인 분해할 수는 없지만, 세 개 이상의 릴레이션 스키마로 비부가적 조인 분해할 수 있는 경우, 다중분해를 수행하여 만들 수 있다. 실제 제 5정규형은 찾아내기 매우 어려운 독특한 제약조건이며 현실에서 조인종속성을 고려하는 경우는 드물다. 제안한 릴레이션 스키마는 조인 종속성을 가지고 있지 않으므로, 본 논문에서는 제 5정규형을 적용하지 않는다. 그 외 기타 종속성과 정규형 또한 제약조건을 고려하지 않더라도, 제안하고 있는 방법에 영향을 미치지 않으므로 적용하지 않는다.

사용자가 요청한 멀티미디어 서비스의 구체적인 정보가 기록된 데이터베이스는 다양한 형태의 적용이 가능하다. 만약, 사용자가 어떠한 서비스를 어떻게 이용하였는지 구체적인 데이터를 찾기 위해서는 MM_SERVICE 릴레이션과 MM_CONTENTS 릴레이션의 조인을 수행하여 튜플을 찾을 수 있다. 이와 같은 데이터를 활용하면, 데이터마이닝(Data Mining) [5]이나 상황인식(Context-Aware)[6, 7]기술을 적용하여 사용자가 주로 요청하는 서비스의 연관성을 찾아내고 분석할 수 있다. 또한 트리거(Trigger)[18]를 적용해, 특정 이벤트나 시간에 맞추어 정보를 제공할 수 있는 사용자 중심의 홈 네트워크 서비스도 가능하다. 사용자의 직접적인 서비스 요청에 의해서만 동작하는 수동적인 환경이 아닌 사용자가 의식하지 않고도 서비스를 제공받을 수 있는 능동적인 서비스 환경이 가능해 질 수 있다.

5. 성능평가

본 장에서는 각 정규형을 만족하는 스키마들의 질의처리에 따른 갱신이상과 중복의 발생빈도를 비교하고자 한다. 본 논문에서는 이를 위해 4장에서 언급한 다양한 정규형을 만족하는 스키마들의 질의결과로써 나타난 데이터베이스 사이즈와 튜플수를 비교, 분석 하였다. 실험은 CPU 3.20 GHz 상에서 My SQL 5.0.0a 데이터베이스관리 시스템을 사용하여 수행하였다. 실험은 또한, 질의 수행결과가 바로 이전에 수행된 결과로 영향 받지 않게 하기 위하여, 각 질의 수행 이후에, DBMS 버퍼를 초기화 하였다.

5.1 질의

본 논문에서는 실험을 위하여 (그림 5)와 같은 세 가지 질의 집합을 사용하였다.

SET 1은 멀티미디어 서비스 릴레이션에서 데이터를 삭제할 때 발생하는 갱신이상현상을 보여주는 질의들의 집합이다. 첫번째 질의에서는 (그림 2) MM_SERVICE 스키마 릴레이션에서 SELECT 연산을 수행하여 저장된 데이터베이스에서의 중복현상을 보이고, 두번째 질의에서는 DELETE 연산을 수행한다. 세번째 질의에서는 데이터 삭제 후, 다시 SELECT 연산을 통해 질의결과를 반환한다.

SET 2는 멀티미디어 서비스 릴레이션에 존재하는 중복현상을 보여주는 질의들의 집합이다. 첫번째 질의에서는 (그림 2)의 MM_SERVICE 스키마 릴레이션에서 SELECT 연산을

수행하며, 두번째 질의에서는 (그림 3)의 MM_CONTENTS, HN_PROVIDER과 MM_SERVICE로 분해된 스키마 릴레이션에서 SELECT 연산을 수행한다.

SET 3은 디바이스와 서비스 릴레이션에 존재하는 중복 현상을 보여주는 질의들의 집합이다. 첫번째 질의에서는 (그림 2)의 DEVICE_SERVICE스키마 릴레이션에서 SELECT 연산을 수행하며, 두번째 질의에서는 (그림 4)의 DEVICE_SER_TYPE과 DEVICE_SER_OWN로 분해된 스키마 릴레이션에서 SELECT 연산을 수행한다.

5.2 데이터베이스

본 실험은 가공데이터(Syntactic Data)를 사용하여 수행한다. 홈 네트워크 시스템 상에서 멀티미디어 서비스에 대한 실제 데이터가 존재하지 않기 때문에 실험을 위한 각 튜플은 임의로 생성하였다.

SET 2의 질의집합에서는 데이터베이스 크기의 효율성을 실험하기 위해, 서비스 릴레이션에서 튜플의 수가 4, 8, 16, 32, 64로 스캐어하게 증가하는 다섯가지 데이터 집합을 생성하였다. SET 3의 질의집합에서의 실험 데이터는 하나의 맥내에는 네 명의 가족구성원과 네 개의 모바일 디바이스가

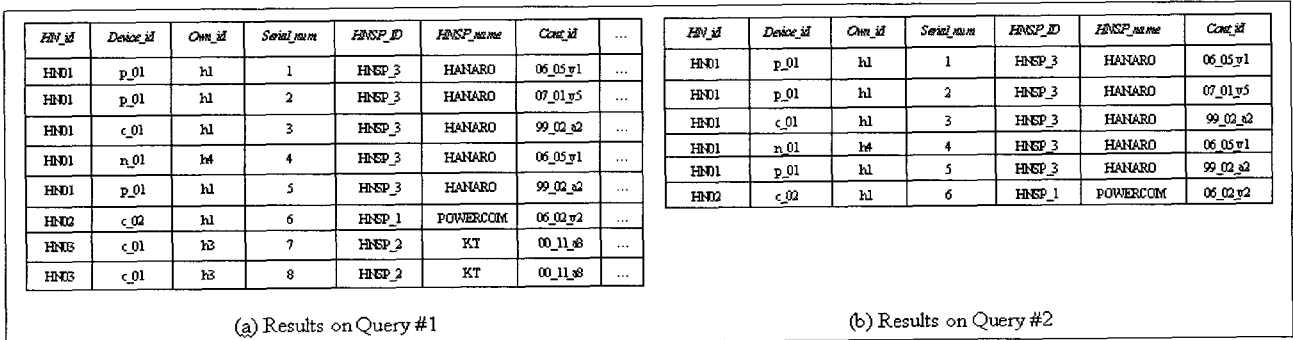
존재한다고 가정한다. 또한, 튜플 수의 증가에 따른 질의 결과 성능을 비교하기 위해서, 멀티미디어 서비스 타입과 가족 구성원의 수를 리니어 하게 증가시킨다.

데이터베이스의 크기는 실험목적에 따라 조정할수 있으나, 본 논문에서 설명하고자 하는 목적에 영향을 미치지 않는다. 그러므로, 실제 실험데이터에 대한 실험은 향후연구로 한다.

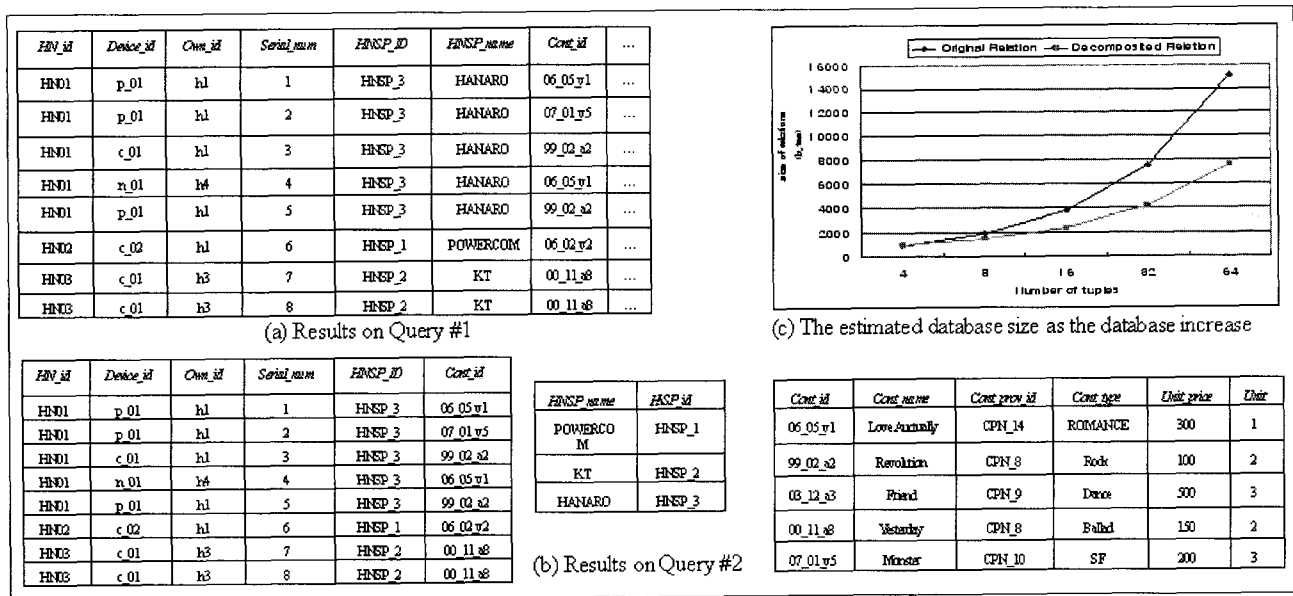
5.3 실험결과 및 분석

(그림 6)은 SET1의 질의수행 결과들을 보인다. (그림 6(a))는 MM_SERVICE 릴레이션에서 데이터의 중복으로, HN_ID가 'HN02'인 사용자 정보를 삭제할 때 발생하는 갱신 이상 현상을 보여준다. 만약, HN_ID가 'HN02'인 홈 네트워크 업체를 이용하는 사용자의 튜플이 삭제되면, 홈 네트워크 업체와 관련된 튜플 역시 삭제된다. (그림 6(b))는 Query2의 질의수행을 통해 이러한 갱신이상의 결과를 보인다. 갱신이상은 데이터베이스의 무결성을 상실시키고, 데이터베이스가 무결성이 상실된 채로 사용된다면, 부정확 또는 잘못된 연산을 초래할 수 있게 된다.

(그림 7)은 SET2의 질의 결과들을 보인다. (그림 7(a))는



(그림 6) Set 1 쿼리결과



(그림 7) Set 2 쿼리결과

Query 1의 질의수행 결과로써, MM_SERVICE 릴레이션에 존재하는 이행적 종속성으로 인해 저장된 데이터의 중복현상을 보인다. (그림 7(b))는 Query2의 질의수행 결과로 MM_SERVICE 릴레이션을 HN_PROVIDER, MM_CONTENTS, MM_SERVICE 릴레이션들로 분리함으로써 멀티미디어 콘텐츠와 멀티미디어 서비스, 홈네트워크 업체 데이터간의 이행적 종속성이 해결됨을 보인다. (그림 7(b))와 같이 제 3정규형으로 분해된 릴레이션은, 중복을 해결하고 (그림 6(c))에서 발생하는 갱신이상 문제를 해결 할 수 있다. (그림 7(c))는 (그림 7(a))와 7(b))에서 질의결과로 나타나는 데이터베이스 크기를 비교한 결과를 보인다. (그림 7(c))의 실험결과에 의하면, 분해되지 않은 릴레이션의 경우, 제 3정규형으로 분해된 릴레이션에 비해 튜플 수가 64일 때, 최대 1.98배까지 데이터베이스의 크기가 급격하게 증가하는 것으로 나타났다. 그 이유는, 분해되지 않은 릴레이션의 경우, 튜플수가 증가함에 따라 중복된 튜플수 역시 증가하여 데이터베이스의 크기가 급격하게 증가하기 때문이다. 이에 비하여, 제 3정규형으로 분해된 릴레이션은 중복을 해결하므로 이전의 방식에 비해 효율적인 가진다. 질의대상 문서수가 증가함에 따라 크기의 차이는 점차 커지므로 제안한 스키마 설계가 효율적으로 저장공간을 사용함을 보인다.

(그림 8)은 SET3의 질의 결과들을 보인다. 그림 8(a)는 DEVICE_SERVICE 릴레이션에 존재하는 다치종속성으로 인하여 발생한 중복현상을 보인다. 그러나, DEVICE_SERVICE 릴레이션을 DEVICE_SER_TYPE과 DEVICE_SER_OWN 릴레이션으로 분해하면, 제 4 정규형을 만족하게 되어 이러한 다치종속성을 제거할 수 있다. 이 분해는 저장 공간을 절약할 뿐 아니라 다치 종속성과 관련된 갱신이상을 피할 수 있다. (그림 8(b))는 Query2의 질의수행 결과를 통해 분해된 릴레이션에서 중복현상이 해결됨을 보인다. (그림 8(c))는 (그림 8(a)와 8(b))의 질의결과로 나타나는 튜플 수를 가족구성원 또는 멀티미디어 서비스 타입의 증가에 따라 비교한 결과를 보인다. (그림 8(c))의 실험결과에 따르면, (그림 3)의 분해 이전의 DEVICE_SERVICE 릴레이션 튜플 수가 제 4정규형으로 분해된 릴레이션에 비해, 한 대내에서 가족구성원이 4명 증가했을 때, 최대 1.97배까지 증가하는 것으로

나타났다. 그 이유는 다치 종속성으로 인한 갱신이상을 방지하기 위해서는 사용자와 서비스 타입간의 독립성을 유지해야 하는데, 이를 위해 각 가족 구성원이나 멀티미디어 서비스 타입마다 튜플을 삽입하여 중복현상이 발생하기 때문이다. 실제응용에서는 실험에서보다 질의대상 문서수가 수십 배로 증가하기 때문에, 그 성능의 차이는 점차 커지므로 제안한 스키마 설계가 더 좋은 질의 처리 능력을 보인다.

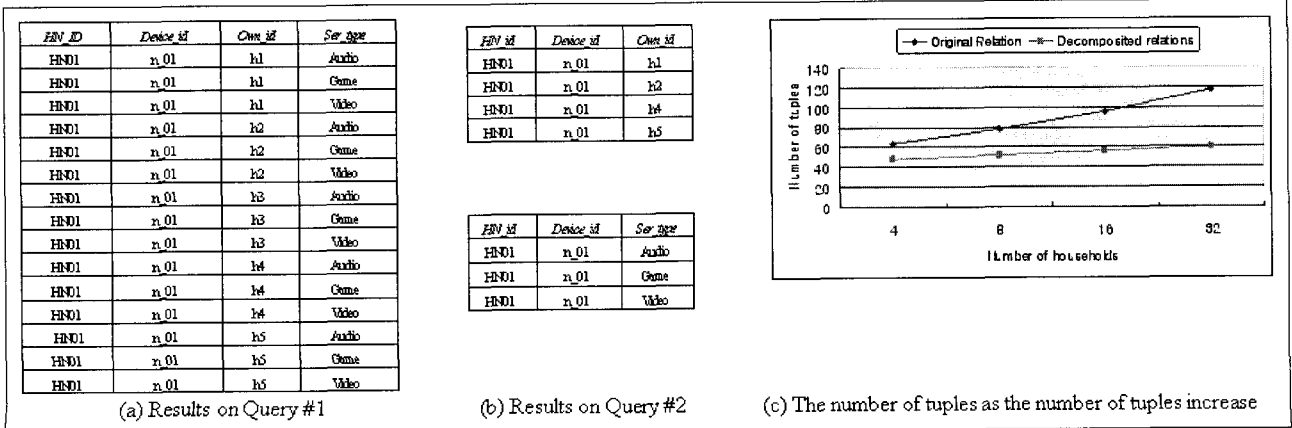
실험결과는 제안한 제 4규형 기반의 데이터베이스 디자인이 홈 네트워크 시스템에서 세분화된 멀티미디어 서비스를 사용자에게 제공해주기 위한 효과적인 방안임을 보이고, 본 논문의 4장에서 제안하고 있는 방법의 적합성을 입증하고 있다.

6. 결론

본 연구는 홈 네트워크 시스템에서 멀티미디어 서비스를 위한 효과적인 데이터베이스 디자인 방안을 소개하였다. 구체적인 사용자 서비스를 지원하기 위해 서비스 운영 중 발생하는 데이터를 저장, 검색, 갱신 할 수 있는 효과적인 데이터베이스를 디자인 하는 것이다. 이를 위하여 사용자 서비스 릴레이션에 존재할 수 있는 다치 종속성을 밝혀내고 정규화(Normal Form) 이론에 기반 한 데이터베이스 디자인 방안을 제안하였다. 그리고 실험을 통하여 다양한 정규형을 만족하는 릴레이션 분할은 분석함으로써, 제안한 방법의 적합성을 입증하였다. 실험결과는 이론적 제 4 정규형 기반의 릴레이션 설계가 홈 네트워크 시스템에서 서비스 데이터의 중복을 제거할 수 있는 효과적인 방안임을 보였다. 향후연구에서는 실제 데이터를 실험으로 연구를 진행한다. 그리고 다양한 멀티미디어 데이터에 대한 인덱싱 방법을 연구하고, 효율적인 검색방법을 제안한다.

참고 문헌

- [1] R. Fagin, "Multivalued dependencies and a new normal form for relational databases," *ACM Tans on Database Systems* 2:3, pp. 262-278, Vol.2, No.3, Sep., 1977.
- [2] Catriel Beerl, Ronald Fagin, and John H. Howard, "A



(그림 8) Set 3 쿼리결과

Complete Axiomatization for Functional and Multivalued Dependencies in Database Relations," *In Proc. 1977 ACM SIGMOD Int'l Conf. on Management of Data*, pp. 47-61, Toronto, Canada, Aug., 1977.

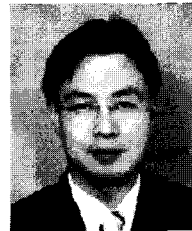
- [3] Carlo. Zaniolo, "A New Normal Form for the Design of Relational Database Schemata," *ACM Trans. on Database Systems(TODS)*, Sep., 1982.
- [4] B. Rose, "Home networks: a standards perspective," *Communications Magazine*, IEEE Vol.39, pp.78-85, Dec., 2001.
- [5] Ming-Syan Chen, Jiawei Han, and Philip Yu, "Data Mining: An Overview from a Database Perspective," *IEEE Transactions on Knowledge and Data Engineering(TKDE)*, Vol.8, Num. 6, pp.866-883, 1996.
- [6] Dey, A. K, et al., "A Conceptual Framework and Toolkit for Supporting the Rapid Prototyping of Context-Aware Computing," *HCI Journal*, Vol.16, pp.97-166, 2001.
- [7] Hong. J.I, et al., "An Infrastructure Approach to Context-Aware Computing," *HCI Journal*, Vol.16, pp.287-303, 2001.
- [8] S. Hartmann, and S. Link, Multi-valued dependencies in the presence of lists," *In Proc. the 23rd SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pp.330-341, 2004
- [9] McCarthy, Dennis R, and Umeshwar Dayal, "The Architecture of an Active Database Management System," *In Proc. 1989 ACM SIGMOD Int'l Conf. on Management of Data*, pp.215-224, Portland, June, 1989.
- [10] Jeffrey D. Ullman, "Principles of Database Systems, 2nd Edition." *Computer Science Press*, ISBN 0-914894-36-6 BibTeX, 1982.
- [11] E. F. Codd, "A relational model of data for large shared data banks," *Communications of the ACM*, Vol.13 No.6, pp. 377-387, June, 1970
- [12] E. F. Codd, "Extending the database relational model to capture more meaning," *ACM Trans. on Database Systems (TODS)*, Vol.4, pp. 397-434, Dec., 1979.
- [13] J. Biskup, U. Dayal, and P. Bernstein. "Synthesizing Independent Database Schemes," *In Proc. 1979 ACM SIGMOD Int'l Conf. on Management of Data*, pp.143-151, Boston, 1979.
- [14] P. C. Fisher, D.M. Tsou, "Whether a Set of Multivalued Dependencies Implies a Join Dependence is NP Hard," *SIAM Journal*, pp. 259-266, 1983
- [15] S. Lee, H.J. Choi, C.H. Kim, and S.M. Back, "A New Control Protocol for Home Appliances - LnCP," *In Proc. ISIE 2001*, Vol.1, pp. 286-291, 2001.
- [16] P. M. Ahmed Al-Theneyan and M. Zubair. "Enabling Jini for use across non-multicastable networks," *In ICASE Report*, No.2000-34, 2000
- [17] C. Ellison, "Device Security: 1 Service Template," *UPnP Forum*, Nov., 2003.
- [18] Al-Fayoumi, Nabeel, "Temporal Trigger Processing in the Trigger Man Active DBMS," *Univ. of Florida*, Aug., 1998.



송혜주

e-mail: thdgPwn@hotmail.com

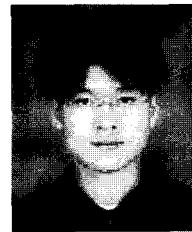
2007년 숙명여자대학교
멀티미디어학과졸업(학사)
2007년~현재 숙명여자대학교
멀티미디어학과(석사과정)
관심분야: 데이터베이스 시스템,
멀티미디어데이터베이스, 바이오 정보공학



박영호

e-mail: yhpark@sm.ac.kr

1992년 동국대학교 공과대학 컴퓨터공학과
(학사, 석사)
2005년 한국과학기술원 전산학과(공학박사)
1993년~1999년 한국전자통신연구원(ETRI)
교환전송연구단 선임연구원
2005년~2006년 한국과학기술원 첨단정보기술연구센터 연구원.
2006년~현재 숙명여자대학교 이과대학 멀티미디어학과 조교수
관심분야: 데이터베이스관리시스템, 정보검색, XML,
Telecommunication System



김정태

e-mail: jungtae_kim@etri.re.kr

2002년 Charles Sturt Univ. AUS Bachelor
of Information Technology(학사)
2004년 Monash Univ. AUS Master of
Network Computing(공학석사)
2004년 8월~현재 한국전자통신연구원
유비쿼터스홈서비스연구팀 연구원



백의현

e-mail: ehpaik@etri.re.kr

1984년 숭실대학교 전자계산학과(학사)
1987년 숭실대학교 전자계산학과
(공학석사)
1997년 숭실대학교 전자계산학과
(공학박사)

현재 한국전자통신연구원 유비쿼터스홈서비스연구팀장