

## 과금 데이터를 이용한 이동통신 네트워크의 특성 연구

김예호, 권중욱, 최형기

성균관대학교 전자전기컴퓨터공학부

[yhhkim | cwkwon | hkchoi]@ece.skku.ac.kr

### Measurement Analysis of Billing Record in the Cellular Networks

Yae Hoe Kim, Chong-Wook Kwon, Hyoung-Kee Choi

School of Information and Communication Engineering

Sungkyunkwan University

#### 요 약

현재 이동통신 네트워크는 통신 기술의 발전에 힘입어 음성 서비스를 넘어선 다양한 패킷 데이터 기반 서비스를 제공하고 있다. 하지만 현재 IEEE 802.11과 같은 이기종 네트워크에 대한 연구가 다양하게 진행되고 있는 것과 달리 이동통신 네트워크의 패킷 데이터 기반 서비스에 대한 연구는 이루어지지 않았다. 이러한 추세의 원인은 이동통신 네트워크가 사업자에 의해서 폐쇄적으로 운영되고 있기 때문이다. 우리는 지금까지 이기종 네트워크에 대한 연구를 바탕으로 예측만 하던 문제를 해결하기 위해 이 논문은 통해서 이동통신 네트워크의 패킷 기반 데이터 서비스의 특징을 분석한다. 실제 트래픽을 모두 분석하는 것은 규모의 문제로 불가능했기 때문에 우리는 서비스에 대한 과금 데이터를 사용했다. 과금 데이터에는 네트워크의 특징이 대부분 반영되어 있기 때문에, 실제 트래픽을 수집해서 분석하는 것과 같다. 분석결과로 이동통신 환경에서 사용자 행동 패턴과 서비스 유형별 트래픽 특성 및 네트워크의 특징을 보여준다. 이러한 결과는 향후 네트워크 진화에 의해서 예상되는 트래픽 예측 및 관리에 활용 가능하고, 네트워크 모델링에 있어서의 기반 지식을 제공한다.

#### 1. 서 론

이동통신 네트워크는 다양한 통신 기술 발전에 힘입어 빠른 속도로 발전하고 있다. 특히, 단말의 성능향상과 사용자의 다양한 요구에 의해서 이동통신 네트워크를 이용한 패킷 기반 데이터 서비스의 수요가 확대되고 있다. 관련 서비스로는 초기 서비스부터 지원되던 WAP (Wireless Application protocol)과 MMS (Multimedia Messaging System)는 물론 VOD (Video on Demand) 및 인터넷 접속 등을 비롯한 많은 서비스가 있다. 이러한 서비스들의 규모가 커지고 개별적인 서비스의 범위가 빠른 속도로 확대되고 있다. 이와 더불어 앞으로의 PCN (Packet core network)은 현재와 같이 이동통신 네트워크에서 패킷 서비스를 지원하기 위한 부분적인 기능이 아니라 전체 네트워크를 통합하는 백본 네트워크로 이용될 것이다.

하지만 기존의 연구들은 직접적으로 이동통신 네트워크의 PCN에 대해서 다루고 있지 않다. 이동통신

네트워크에서 패킷 데이터가 서비스된 시점의 문제와 서비스 제공자에 의해 폐쇄적으로 운영되는 문제 때문이다. 현재까지 진행된 연구들은 소규모의 무선랜 또는 2세대 이동통신 네트워크에 대한 연구가 대부분이다. 따라서 지금까지 우리들은 이동통신 네트워크의 PCN에 대해서 제한적으로 예상하고 있을 뿐이었다.

앞으로 이동통신 네트워크는 다른 이기종 네트워크와 더불어 통합 구조로 발전하게 될 것이다. 통합 환경에서 일어날 문제점과 그 대안을 예상하고 연구하기 위해서는 각각의 세부 네트워크의 특징을 이해해야 한다. 그 대안으로 우리는 이 논문에서 이동통신 네트워크의 PCN에 대한 특성을 분석했다. 본래 특성 분석을 위해서는 실제 네트워크의 트래픽을 모두 수집해서 분석을 시도한다. 하지만 우리는 규모의 문제로 인해서 제한적인 방법을 선택했다. 우리는 다양한 정보를 포함하고 있는 과금 데이터를 통한 분석방법을 사용했다. 과금 데이터의 정의는 2장에서 자세히 설명한다. 데이터의 분석 결과로 사용자 행동 패턴 및 네트워크의 특징을 정의했다. 이 논문의 결과는 향후 이동통신 네트워크의 PCN에 있어서 사용자 행동 패턴과 네트워크 트래픽을 예측하고 관리하는데 중요한 기준이 될 것이다.

<sup>1</sup> "본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음" (IITA-2008-C1090-0801-0028)

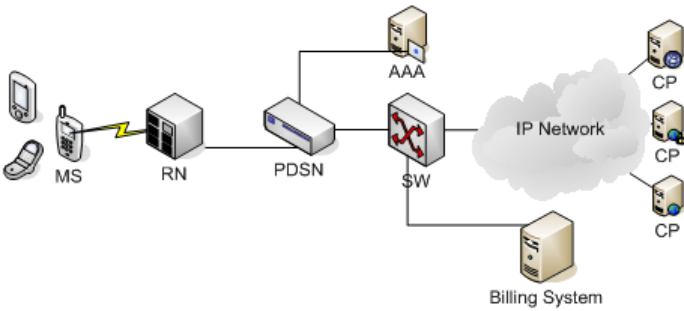


그림 1) CDMA 네트워크 과금 시스템의 구조

본 논문의 2장은 관련 연구로 이루어져 있고, 3장은 분석 방법론에 대해 설명하고 있다. 4장은 분석한 결과를 설명하고, 5장은 앞으로 기대되는 연구 방향에 대해서 설명한다. 마지막으로 6장에서 논문의 결론을 포함한다.

## 2. 관련 연구

유선 네트워크에 대한 특성화 분석을 오랜 기간 동안 연구가 되었고 많은 부분이 확인되었다. 그러나 무선 네트워크에 대한 특성화 분석은 최근 연구되고 있는 주제이다. 현재까지 진행된 연구의 대부분은 IEEE 802.11을 사용한 무선랜을 대상으로 하고 있다. 실험 환경 구성이 간단하고 서비스 형태가 개방적이기 때문이다. 반면에 이동통신 네트워크에 대한 연구의 경우는 제한적인 단말간의 트래픽 모델링과 소규모 시뮬레이션에 의한 연구가 대부분이다.

Tamas Varga는 GPRS (General Packet Radio Service) 네트워크에서 WAP 서비스에 대한 트래픽 모델을 연구했다. PCN이 시작된 초기에 이루어진 연구이고 PCN의 중요한 서비스인 WAP의 특징을 밝혔다.

Luc A. A는 CDPD (Cellular Digital Packet Data) 네트워크에서 본 논문과 같은 방법으로 과금 정보를 이용해서 사용자의 행동 패턴을 분석했다. 네트워크의 특징을 체계적으로 정리하고 상관분석을 이용한 결과를 제안한다. K-mean clustering 알고리즘으로 각 사용자를 그룹화한 것이 특징이다[2].

Anand Balachandran은 IEEE 802.11 무선랜 네트워크에 대해서 특성을 분석했다. 작은 규모의 연구이지만 특성화 기법 및 방법론에 대해서 명확한 정의를 하고 있다[3].

Young J. Won는 CDMA 네트워크의 트래픽을 수집한 후 유선랜과 비교해서 분석했다. 무선 네트워크에서 일반적으로 나타나는 특징으로 업로드와 다운로드 트래픽이 대비가 되며, 전체적으로 네트워크 상의 패킷 사이즈가 작고, 개별적인 세션 시간이 짧은 특징을 보였다[4].

지금까지 이동통신망의 PCN에 대해서는 다른 분야의 연구결과를 바탕으로 예측할 뿐이었다. 우리는 이

논문에서 PCN의 특징에 대해서 실질적인 정보를 제공한다. 자세한 내용은 3장과 4장에서 설명한다.

## 3. 분석 기법

우리가 분석하는 대상은 이동통신망의 패킷 데이터에 대한 과금 데이터이다. PCN은 과금 처리를 위해 독립된 시스템을 사용한다. 모든 패킷 데이터의 정보가 과금 시스템을 경유하며 처리되어 기록된다. 따라서 과금 데이터는 전체 네트워크의 특성을 포함한다.

이동통신망을 이용하는 모든 패킷 데이터는 두 종류로 구분된다. 이동통신 서비스 제공자의 코어망을 이용하는 내부 서비스 패킷과 코어망을 제외한 외부 인터넷을 이용하는 외부 서비스 패킷이다. 그림 1)을 보면, 단말이 사용하는 두 종류의 패킷은 반드시 PDSN (Packet Data Serving Node)을 경유해야 한다. 패킷은 PDSN에 연결된 스위치에서 복사되고, 복사된 패킷 중 한쪽 패킷이 과금 시스템으로 전송된다. 전송된 패킷은 과금 시스템에 의해서 과금 처리 데이터로 생성된다.

과금 처리 데이터는 PDSN이 생성하는 유저 세션 정보인 UDR (Usage Data Record)과 과금 시스템이 생성하는 개별 서비스 기반의 세션 정보인 CDR (Contents Data Record)로 구성된다. **오류! 참조 원본을 찾을 수 없습니다.** UDR은 CDMA 표준에 정의되어 있지만, CDR의 경우는 개별 서비스 제공자가 독립적으로 규정해서 사용하는 기준이다. UDR은 서비스 접속/종료시간, 사용자 환경 정보 (Network Access Identifier, 단말 IP, 위치 정보 등)과 In/Outbound 패킷 사용량 등의 정보를 담고 있으며, CDR은 사용자의 서비스 세션에 대한 상세 정보 (서비스 종류, 서버 IP, Port 정보, Up/Down Packet Size, Pack Count, 재전송 패킷 수와 크기)을 담고 있다.

과금 처리 데이터는 시간 당 대략 1GBytes의 분량이다. 따라서 장기간 데이터 수집이 불가능 하다. 그래서 우리는 전체적인 변화를 파악하기 위해서 통계정보를 활용했다. 과금 시스템을 통해 처리되는 모든 처리 결과 중 일부가 기록되어 통계정보로 제공된다. 이 정보를 활용하면 장기간의 특징과 변동을 파악할 수 있다.

우리가 사용한 트래픽은 두 종류이다. 첫째, 일주일 동안 (2008년 3월 13일~2008년 3월 20일)의 과금 처리 데이터이다. PCN을 이용하는 사용자의 세션에 대한 정보를 UDR을 이용해서 파악하고, 개별적인 서비스에 대한 정보를 CDR을 이용해서 파악하는 방법을 사용했다. 둘째, 14개월 동안 (2006년 12월~2008년 1월)의 과금 통계 데이터이다. 통계 데이터를 이용해서 서비스 타입의 특징과 추세를 분석할 수 있었다.

현재 PCN에서 제공되는 서비스는 대표적인 서비스를 포함해서 대략 40가지이다. 초기의 소수의 서비스에서

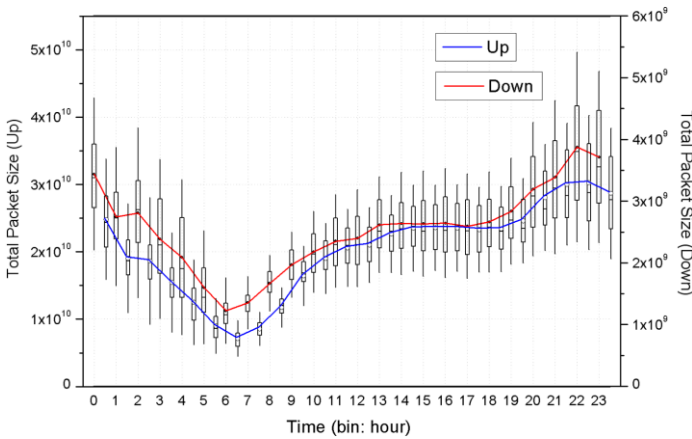


그림 2) 사용자 트래픽 패턴의 시계열 분석

시간이 흐르면서 다수의 서비스가 추가되어 확대되었다. 그러나 우리의 분석 목적은 전체적인 네트워크의 특징이기 때문에 개별적인 서비스를 분석하는 것은 의미가 없다. 그렇기 때문에 우리는 군집 분석을 시도했고, 모든 서비스 타입에 대해서 가용한 모든 정보에 활용해서 분류를 정의했다. 분류 기준과 분류 결과에 대해서는 4장에서 결과 분석과 함께 설명한다.

#### 4. 분석 결과

우리는 두 가지 정보를 차례로 분석하는 방법을 사용했다. 우선 전체적인 트래픽 분산 패턴을 파악하기 위해서 통계 정보를 분석했다. 그림 2)는 하루 동안의 시간별 트래픽 사용량을 보여준다. 전체 데이터에 대해서 일자별로 동일한 결과를 보여준다. 가장 트래픽이 많은 시간이 21시부터 23시까지 2시간이고 가장 트래픽이 적은 시간이 5시부터 8시까지 3시간이다. 이러한 특징은 사용자의 생활 패턴을 반영하고 있다. 우리는 이 특징을 바탕으로 실제 트래픽을 수집하는 시간의 기준을 결정했다.

자세한 분석을 위해서 과금 처리 정보를 분석했다. 우선, 데이터의 일반성을 확인하기 위해서 UDR에 대한 시계열 분석을 한 결과 UDR 세션 수가 시간 별로 일정하게 분포됨을 파악했다. 사용자의 사용 시간에 대한 패턴을 알아 보기 위해 세션 시간 별 사용자 세션 분포를 살펴 본 결과 우리는 200초 이내 서비스

완료되는 비중이 전체 70%이상을 차지 하고 있으며, 전체 90%가 400초 내에서 서비스를 사용하고 있음을 알 수 있다. 전체 세션의 평균 시간이 300초 이내로 무선 네트워크의 특징을 보여준다. 그리고 사용자 세션간 Inter Arrival Time의 분포를 분석한 결과 전체 세션의 80%가 100초 이내에 분포하고 있었다. UDR의 경우 서비스의 전환이나 에러로 인해서 하나의 사용자에게 의해서 연속적으로 사용되는 경우가 많기 때문이다. 사용자 세션 별 사용 패킷 사이즈에 대한 분포를 살펴 본 결과 1Kb 미만의 세션이 90% 비중으로 전체 대부분을 차지하고 있다. 이상치를 구분하기 위해 1%의 보정을 적용했다. 1%의 보정치에 속한 데이터를 분석해본 결과 전체 패킷 양의 대부분을 이상치가 점유한다는 특징을 찾았다. 소수의 사용자가 전체 네트워크의 대부분을 점유하는 있기 때문이다.

이어서 우리는 과금 처리 정보에서 CDR 정보를 분석했다. CDR에는 UDR과 달리 개별 서비스에 대한 세션 정보가 담겨 있다. UDR과 연동해서 처리한 결과로 우리는 단위 시간당 사용자 세션 수와 세션 시간과 사용자 세션간의 Inter Arrival Time, 사용 패킷 사이즈 별 유저 세션 수를 분석할 수 있었다.

우리가 정의한 분류는 모두 4가지 이다. 분류1은 WAP이다. WAP을 바탕으로 텍스트 기반 서비스를 분류했다. 분류2는 VM (Virtual Machine)을 다운받거나 1MByte이하의 정보를 주고받는 서비스이다. 분류3은 VOD와 DMB (Digital multimedia Broadcasting)와 같은 스트리밍 서비스를 포함하는 멀티미디어 서비스이다. 넷째는 코어망을 제외한 외부 인터넷을 이용하는 서비스이다.

서비스 세션 별 서비스 분류별 분포를 살펴 본 결과 우리는 분류1의 세션 수가 전체 대비 50%를 차지하고 있음을 알 수 있었다. 이러한 결과는 단말 브라우저를 통해서 코어망을 경유하는 서비스가 주로 이용되고 있음을 파악할 수 있다. 단말에서 패킷 기반 서비스를 이용하는 대부분의 경우가 WAP으로 초기접속 과정을 거친 후에 이루어지기 때문이다. 그리고, 외부 인터넷을 이용하기 위한 세션, 즉 분류4가 나머지 약 50% 비중을 차지하고 있으며, 이러한 세션의 비중이 꾸준히 증가하고 있다. 이러한 이유는 사용자들이 서비스

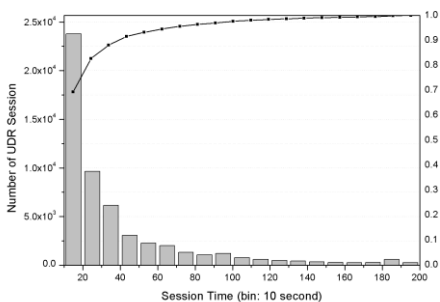


그림 3) UDR session time

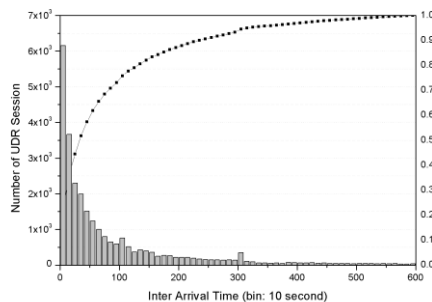


그림 4) UDR inter arrival time

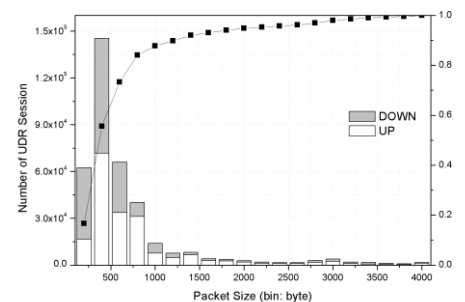


그림 5) UDR Up/Down traffic

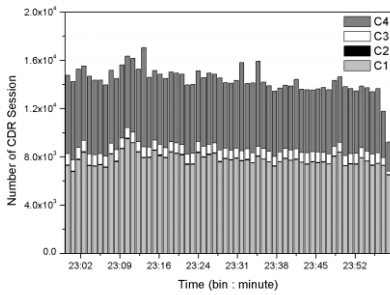


그림 6) CDR session count

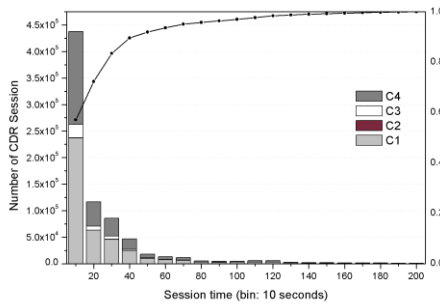


그림7) CDR session time

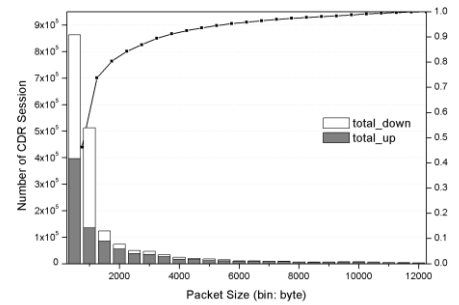


그림8) CDR Up/Down traffic

업체가 제공하는 제한적인 서비스를 이용하기 위해 PCN을 이용하는 것이 아니라 외부 인터넷을 이용하기 위한 접속 방법으로 PCN을 이용하고 있다는 점을 알 수 있다. 이러한 세션을 현재 과금 시스템 및 PCN에서는 비중 있게 다루고 있지 않지만 앞으로 변화해야 할 부분이다.

그림 6)은 한 시간 동안의 전체 CDR 세션의 분류 별 수를 나타낸 그래프이다. 분류2가 상대적으로 적다는 점과 각각의 서비스의 비율이 일정하게 나타난다는 점을 알 수 있다. CDR 세션은 개별 서비스에 대한 접속 수를 의미하기 때문에 이 비율이 곧 전체 서비스의 비율을 의미하진 않는다. 분류3의 경우 CDR 세션 수는 적지만 업로드와 다운로드의 비율을 보면 50%를 넘는 점유율을 갖는다. 다른 분류와 달리 각 CDR 세션마다 네트워크의 대역폭을 많이 차지하는 특징을 갖는다. 그림7)의 그래프는 각 분류 별 CDR 세션 시간을 나타낸다. 분류1과 분류4가 40%씩을 차지하고 분류3이 나머지를 차지한다. 하지만 이 그래프는 전체 수를 누적분포로 표시한 것이기 때문에 실제 분류3의 특징을 보여주진 않는다. 분류3의 경우 독립적인 한 세션이 이상치로 인식될 정도로 긴 CDR 세션 시간을 갖는 경우가 많기 때문이다. 그림8)의 경우는 업로드와 다운로드에 대해서 비교대상이다. CDR 세션의 90%가 6000 Bytes 이하의 트래픽 크기를 갖는다. 업로드와 다운로드의 비율은 분류 별로 크게 다르다. 분류1과 분류4의 경우는 대비가 크고, 분류2와 3은 업로드 트래픽이 거의 없다.

사용자 세션에 대해서 통계 정보를 제공한다. 사용자는 서로 다른 서비스(CDR)를 하나의 사용자 세션(UDR)을 통해서 이용할 수 있다. 서비스 세션과 사용자 세션의 관계는 분류 별로 큰 차이를 보여준다. 현재 PCN에 대한 분석은 지금까지 이루어진 연구에서 사용한 기준과 달리 패킷 기준, 또는 사용자 세션 기준으로 연구되어선 특징을 보여줄 수 없다. 서비스 세션 기준의 분석과 사용자 세션의 분석이 연동되어야만 적절한 모델이 제시될 수 있다.

5. 결 론

이 논문은 상용 이동통신 환경에서 사용자의 행동 패턴 및 네트워크 특징 분석했다. 분석 대상으로 과금 시스템에서 과금 처리 정보를 수집해서 사용했다. 논문의 분석 결과는 다음과 같다. 이동통신망의 패킷 기반 데이터 서비스는 트래픽의 변동이 크고, 부하가 시간 의존적이다. 서비스 특성에 따른 편차가 크며, 사용자 세션과 서비스 세션간의 특징이 분류에 따라 다르다. 본 논문의 결과는 현재 이동통신망의 가입자가 다른 환경으로 이동하는 상황에서도 현재 제공되는 서비스와 차별성이 있는 특별한 응용 프로그램의 등장 없이는 사용자의 패턴이 크게 달라지지 않을 것이다. 따라서, 본 논문의 결과는 서비스 제공자가 네트워크를 운영하는 기준으로 사용할 수 있으며, 이동통신 네트워크의 특징을 연구하는데 기반 자료로 활용 가능한 것이다.

6. 참고 문헌

[1] Tamas Varga, "Analysis and Modelling of WAP Traffic in GPRS Networks". Proceedings of ITCSS, September 2004  
 [2] Luc A. A, "Analysis of Use Behavior from Billing Records of a CDPD Wireless network", Proceedings of LCN, November 2002  
 [3] Anand Balachandran, "Characterizing User Behavior and Network Performance in a Public Wireless LAN", Proceedings of SIGMETRICS, June 2002  
 [4] Young J. Won, "Measurement Analysis of Mobile Data network", Proceedings of PAM, April 2007

표 1) UDR과 CDR 및 서비스 분류 별 특징

View of user session (UDR)					
Number of session	418987				
Average Session Time	37.86				
Average Inter arrival Time	113.31				
Average Packet Size (in/out)	3653.94 / 25206.55				
View of service session (CDR)					
Category	C1	C2	C3	C4	OTHER
Average Number of Session	467086	2756	50913	346597	104028
Average Session Time	37.25 (148.74)	65.54 (82.37)	298.38 (368.06)	173.47 (190.65)	13.23 (161.12)
Average Packet Size per Session (total up/down)	751.3/ 4552.3	2675.3/ 65665.9	5100.1/ 126070.9	1174.5/ 4227.0	353.7/ 1061.4
Average Packet size per Packet (total up/down)	107.9/ 713.7	55.3/ 1312.8	158.2/ 661.1	80.5/ 362.3	137.8/ 692.7