

## 이동통신 망 부하 해석을 위한 대표통화량의 설정\*

이윤동<sup>1)</sup> 고승곤<sup>2)</sup> 안병진<sup>3)</sup>

### 요약

본 연구에서는 이동통신사의 네트워크 관리의 효율성 향상을 목적으로 이동통신 통화량 자료의 통계적 특성을 분석하고, 이를 통하여 월별 통화량 자료를 요약하기 위한 통계적 방법을 국내 특정 이동통신사의 사례를 중심으로 살펴보게 된다.

주요용어: 접속망, 대표통화량, 이상치

### 1. 서론

효율적인 통신 서비스를 위한 확률적 혹은 통계적 현상들에 대한 연구는 Erlang 이후 하나의 줄기를 이루면서 연구되어진 분야이다(참조: Brockmeyer, Halstrom와 Jensen, 1948). 특히 확률론과 대기행렬이론 분야에서는 매우 중요한 연구 주제 중의 하나로 많은 연구자들에 의하여 연구되어져 왔고 지금도 다양한 분야와 연계되어 활발히 연구되고 있다. 그 동안 이론적 측면에서 특히 강조되어 온 연구 주제는 확률론적 측면에서 Elarng-B, Elarng-C 분포의 해석과 확장에 관한 연구(예: Halfin와 Whitt, 1981), 그리고 운용 효율성을 도모하기 위한 통화량 구축 방법 등과 관련된 이론적 망 구성 방법론에 대한 연구 주제들을 주로 들 수 있다. 통화량 관리에서 가장 기본적인 주제는 A. K. Erlang이 활동하던 시절과 마찬가지로 시도호와 통화 처리자 수 사이의 관계 해석이다. 이전 시대에 비하여 통신 서비스 시스템은 근본적으로 매우 다른 규모로 발전하였고 이동통신으로 대표되는 통신방식 면에서의 주요한 변화가 있었음에도 불구하고 역시 통신회사들의 주관심 사항은 변화되는 통신 수요에 대비하여 통신망 처리 규모를 어느 정도로 준비하여야 하는 것일까 하는 문제에 답하는 것이라고 하겠다. 보다 구체적으로는 마케팅 측면에서 예상되는 수요를 감당하기 위한 합리적 규모의 망 설비들을 경영 효율성을 만족시킬 수 있는 적절한 시기에 준비하도록 하는 것이라 할 것이다. 이

\* 본 연구는 한국학술진흥재단의 지원에 의하여 이루어졌음(KRF-2006-C00056).

1) (143-701) 서울시 광진구 화양동 1번지, 건국대학교 응용통계학과, 조교수

E-mail: poisson@dreamwiz.com

2) (461-701) 경기도 성남시 수정구 복정동 산65번지, 경원대학교 응용통계학과, 부교수

E-mail: sgk@kyungwon.ac.kr

3) (143-701) 서울시 광진구 화양동 1번지, 건국대학교 응용통계학과, 교수

E-mail: bjahn@konkuk.ac.kr

러한 점들을 고려할 때 상대적으로 고가의 장비를 사용하여야 하는 이동통신 회사들의 입장에서, 특히 통신망의 규모와 수요에 따른 통신망 부하 해석은 경영 효율성 달성을 위한 일차 과제 중 하나가 될 것이다.

이러한 연구들에 있어서, 통화 처리자 수와 시도호에 따라 호손율(통화 서비스 미처리 비율) 계산을 가능하게 하는 Erlang의 모형들은 기본적으로 매우 중요한 역할을 하게 된다. 여기에 더하여, 대규모 통신시스템 관리에서 특별히 중요하게 고려되는 사항은 매우 다양한 망 관리 요소들 사이의 관계에 따른 통화량의 변화를 해석하는 것이다. 최근 국내 이동통신 시장은 가입자 확보에 상당한 경쟁적 노력을 경주하고 있다. 이때 예를 들어 대규모 통신시스템에서 현재 또는 장래에 예상되는 마케팅 활동의 결과로 증가되는 신규가입자군이 망부하에 어느 정도를 기여하게 될 것인가를 예측하는 것은 전체 망 운영과 물자 수급계획의 작성에 있어서 매우 중요한 역할을 하게 된다. 그러나 이러한 분야에 대하여는, 개개의 통신회사별로 통신망의 운용 방식이 다르고 그에 따라 나타나는 현상이 다른 이유 등으로 인하여 구체적인 학술적 연구 주제가 되지 못했다.

본 연구에서는 국내 주요 이동통신사 한 곳의 사례를 중심으로 통신망 부하의 해석을 위한 통계적인 작업 과정을 구체적으로 살펴보기로 한다. 다음의 제 2절에서는 해당 이동통신사의 경우를 중심으로 하여 통신망 구성과 운용 방식, 그리고 기존에 사용되는 망관리 기법들에 대하여 간략하게 소개를 하기로 한다. 다음으로 제 3절에서는 통계적 방법에 의한 통화량 모형화 방법에 대하여 설명하기로 한다. 제 4절에서는 대표통화량의 설정 방법에 대하여 설명하기로 한다.

## 2. 통신망 구성과 운영 및 관리 방식

이동통신 기술은 음성통화와 단문메시지 전송만을 대상으로 하던 IS-95A/B (이하 2G라 칭함) 중심에서 2000년 이후 데이터통신 기능이 포함된 CDMA2000 1x, EVDO (이하 1X라 칭함)로 발전하였고, 근래에는 동영상 멀티미디어 전송이 가능한 기술이 개발 사용되고 있으며, 극히 최근에는 이동통신 단말기에서 위성이나 지상파를 이용한 방송 정보 수신에 가능한 DMB 기술이 운영되고 있다.

이에 따라 현행 사용되고 있는 이동통신 단말기는 그 개발 판매 시기에 따라 성능 면에서 다양한 차이를 보이고 있다. 특히 현행 이동통신 체계에서의 통신망은 초기 2G 단말기 중심의 서비스를 대상으로 개발된 2G 망과, 1X 서비스를 중심으로 2000년 이후 신규로 설치된 1X 망의 이중구조를 가지고 있다. 통신망의 이중구조에서 발생하는 이용상의 비효율성을 극복하기 위하여 복잡한 통신망 할당계획이 운영되고 있다.

통신망의 구성은, 전국적으로 약 8000여개의 기지국 (BTS)과, 그리고 다수의 교환국과 (BSC), 집중교환국 (CGS), 그리고 망 등록사용자 정보를 기록하는 사용자기록장치 (HLR), 패킷통신 서비스장치 (PDSN) 등의 부분으로 구성된다. 이들을 업무 분장 측면에서 접속망(access network), 중심망(core network), 전송망으로 구분하여 관리한다. 접속망의 주 관리 대상은 각

기지국을 포함한 단말기에서 망 접속 부분과 이에 직접적으로 연관되는 부분이다. 교환국과, 집중교환국, 사용자기록장치, 패킷통신 서비스장치 등은 중심망으로 분류하여 관리된다. 또한 이들 사이의 연결에 소요되는 광케이블 등의 전송선로들과 이에 직접적으로 관련되는 장비들을 포함하는 부분을 전송망으로 따로 구분하여 관리한다. 지역적으로는 통신망을 서울 경기 인천을 포함하는 수도권지역, 부산 경남 울산을 포함하는 부산지역 등 전국을 9개 지역으로 구분하고 이를 다시 5개 권역으로 묶어 관리하고 있다. 그림 2.1은 이동통신 통화망의 대략적 구조를 나타내고 있다.

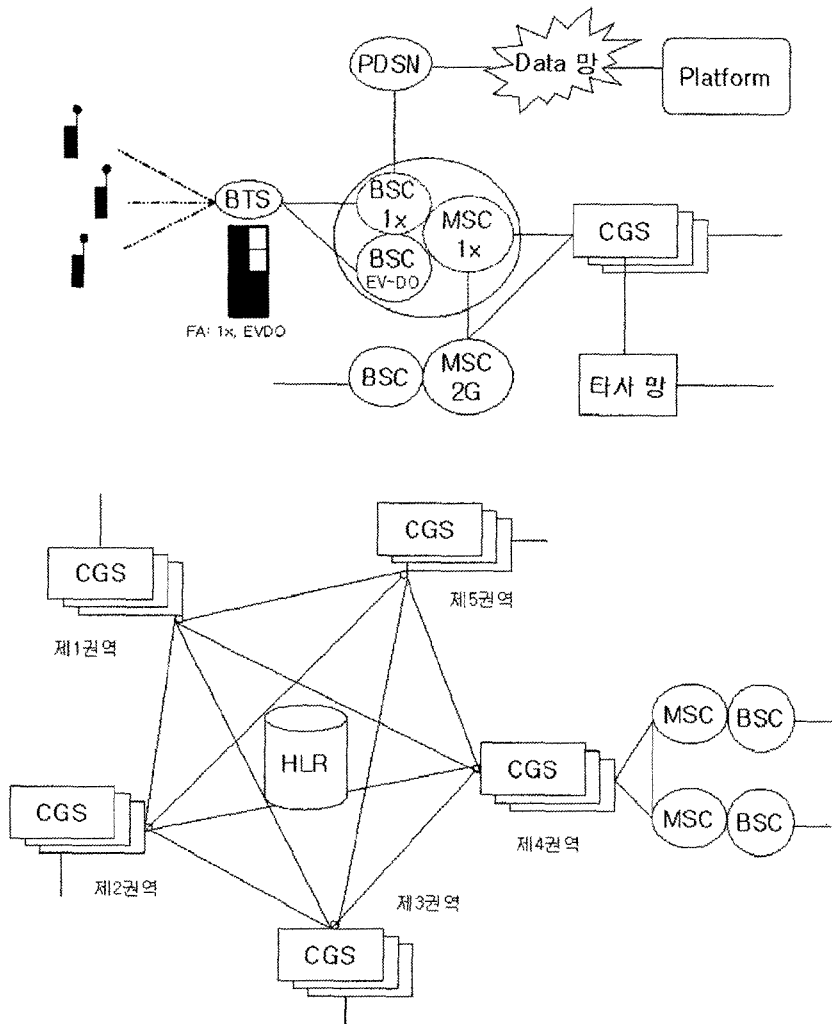


그림 2.1: 이동통신망에서 접속망과 중심망의 구조

통신망 관리를 위하여 현재 사용하고 있는 방법은 다음과 같다. 통화량 자료는 시스템에서 시간별로 집계되어 데이터베이스에 저장되고, 이 통화량 중에서 각 기지국별 일별 최대값을 추출하고, 각 기지국의 일별 최대값들로부터 각 기지국의 해당 월을 대표하는 대표통화량을 결정하게 된다. 이렇게 결정된 대표통화량은 이후 망 부하관리와 자재소요계획을 위한 해당 기지국의 월별 기준통화량으로 사용되게 된다.

이동통신에서는 이동 중인 통화자가 담당 기지국의 해당 관리영역(이하 셀)을 넘어가더라도 통화의 끊김 없는 관리를 가능하도록 하기 위하여 기지국이 담당하는 셀(cell) 들을 이웃 기지국들과 중첩시켜서 관리하게 된다. 따라서 통화자가 이러한 기지국 셀들 사이의 중첩 부분에서 통화를 하는 경우는 복수의 기지국에서 동시에 통화 관리를 하게 되어 통화망 부하가 두배로 늘게 된다. 대략적으로는 셀의 중첩 부분으로 인하여 부가되는 통화량은 전체의 15% 정도라고 알려져 있다. 이때 중첩 부분의 통화량을 단일 통화량으로 계산한 통화량을 CT(Channel Traffic) 통화량이라 하고, 중첩 부분의 통화량을 복수 통화로 고려하는 통화량을 CE(Channel Element) 통화량이라 부른다. 또한 통화량의 기준으로 통화시간 요소를 무시하고 통화수(CA; call amount)만을 집계하는 방법이 사용되기도 한다.

또한 일최대 값을 집계하는 방법에도 몇 가지의 다른 기준이 사용된다. 관리최번시, 실최번시 기준이 그것이다. 관리최번시 기준은 해당 기지국의 통화량이 가장 많아질 것으로 예상되는 시간을 미리 결정해 놓고 해당 시간에서의 통화량을 해당 기지국의 일일 최대통화량으로 간주하는 것이다. 실최번시는 해당 기지국별로 일일 통화량 변화에서 실제로 통화량이 최대가 되는 시간을 일일 최대통화량으로 간주하게 되는 기준이다. 중심망이나 전송망의 경우와 같이 개개의 기지국이 아니라 통신망을 전체적으로 관리하는 경우에는 해당 기지국별 관리최번시나 실최번시를 사용하기 보다는 관리 대상이 되는 시스템 전체에서의 관리최번시가 보다 중요한 의미를 갖게 되는데 이를 관리합최번시라고 한다.

접속망 관리의 기준이 되는 통화량은 실최번시 CE 통화량이고, 중심망은 그 관리 기준으로 관리합최번시(BH)에서의 통화수(CA)를 기준으로 사용한다. 보다 구체적으로는 관리합최번시에서의 통화수(BHCA)를 망등록사용자수(VLR)로 나눈 값 BHCA/VLR을 기준으로 대표값을 계산하고 이를 다시 VLR을 대신할 것으로 예상되는 다른 값을 곱하여 월 대표통화량으로 사용한다. VLR은 HLR에 등록된 사용자를 기준으로 계산된다. VLR값이 통화량을 설명하는데 매우 설명력이 높은 변수임에도 불구하고 VLR값 역시 통화량과 마찬가지로 궁극적으로 그 변화를 해석 예측해야 하는 값이기 때문에 VLR 값을 대신해서 예측 또는 계획이 가능한 다른 변수를 대신 사용하게 된다. 이를 위해 사용되는 변수는 1X 단말기와 2G 단말기의 판매 대수를 말하는 Sub1X, Sub2G의 변수 그리고 통신망 할당 계획의 결과로 나오는 Trans라는 변수이다.

이와 같이 얻어진 일일 통화량을 이용하여 다음에서 간략하게 서술된 바와 같은 방법에 따라 월별 대표통화량을 설정하게 된다.

1. 해당 월에서 대표성이 높은 주간을 기준으로 선정한다. 이 때 연휴나 명절이 기준 주에

포함되지 않도록 한다.

2. 해당 월에 해당 기지국에서 일별 최대 통화량의 값이 주중에 높은 지, 주말에 높은 지를 판단한다. 주말에 통화량이 높은 기지국을 주말 국소, 주중에 높은 국소를 주중 국소라 한다.
3. 주중 국소는 주중값만을 이용하고, 주말 국소는 주말값만을 이용하여 먼저 평균( $m_0$ )을 구한다. 일별 최대값중 그 값이 매우 크거나 작은 값들을 제거하고, 다시 평균( $M$ )과 표준편차( $S$ )를 구한다. 여기서 평균  $M$ 에 관리 여유를 고려하기 위하여 표준편차  $S$ 에 상수를 곱한 값을 더하여  $M + cS$ 의 형태로 대표 통화량을 설정한다. 이때 상수  $c$ 는 접속망이나 중심망이나에 따라 해당망 관리의 특성을 고려하여 적당한 값을 이용한다.

이와 같은 대표통화량 설정 방법에는 다음과 같은 단점이 있다. 먼저 어떤 주를 기준 주로 선택할 지에 따라, 또 주중값을 사용할 것인지 주말값을 사용할 것인지에 따라 대표통화량의 값은 매우 크게 달라질 수 있음에도 이에 대한 명확한 기준이 없어, 경험있는 업무 담당자의 정황판단에 의존하게 된다. 또한 이렇게 구한 대표통화량의 통계적 의미가 명확하지 않은 이유로 이를 이용한 통계적 해석과 추론에 어려움을 겪게 된다.

### 3. 통화량의 특성 분석

그림 3.1은 1X망 전국 CT 통화량의 2004년 11월에서의 시간별 통화량의 변화 모습을 보여준다(종축의 통화량 숫자는 절대 단위가 아님). 아래쪽 그림은 특히 세 번째 주의 변화를 자세히 보기 위하여 따로 그린 그림이다. 통화량의 전반적 변화 형태 중 가장 큰 요인은 요일 효과이다. 주중에 통화량이 많고 주말에 통화량이 적은 형태를 보여준다. 11월 세 번째 주 수요일의 최대 통화량이 다른 주의 수요일과 달리 약간 큰 값을 갖는 것을 볼 수 있는데, 11월 17일 이 날은 수학능력고사 일로 전국적 통화량 패턴이 다른 날과 차이를 보이고 있다. 일일 변화 형태에서 가장 눈에 띄는 사항은 통화량 패턴이 하루에 두 번 고점을 형성한다는 것이다. 먼저 나타나는 작은 고점들은 보통 오전 11시-정오 사이의 통화량이고, 뒤에 나타나는 큰 고점들은 오후 5시에서 6시 사이의 통화량이다. 또 오후 8시에서 9시 사이에 변곡점이 형성되어 있다. 이는 점심이나 저녁식사 약속 또는 퇴근후의 만남 약속 등이 통화량의 고점 형성의 주요인임을 말하고 있다. 요일별로 형성되는 봉오리 면적의 비율은 (월, 화, 수, 목, 금, 토, 일)=(19, 19, 15, 15, 15, 13, 10)의 형태를 보이고 있다. 월간 통화량의 시간별합을 구하고 이를 해당 월의 (평균)가입자 수로 나눈 것을 MOU라 한다. 이를 통해서 볼 때 월별로 정리된 MOU는 월별 일수에 영향을 받고, 또한 해당 월에 주말이 몇 번 들어있는가에 상당한 영향을 받게 됨을 미루어 알 수 있다. 가입자 1인당 사용하는 평균 통화량, 즉 통화성향을 말하는 호습성의 계절적 변화를 알기 위해서 MOU에 들어 있는 이러한 변동 요인들을 제거하고 나면 그림 3.2와 같이 나타나는 MOU 계절성을 얻게 된다.

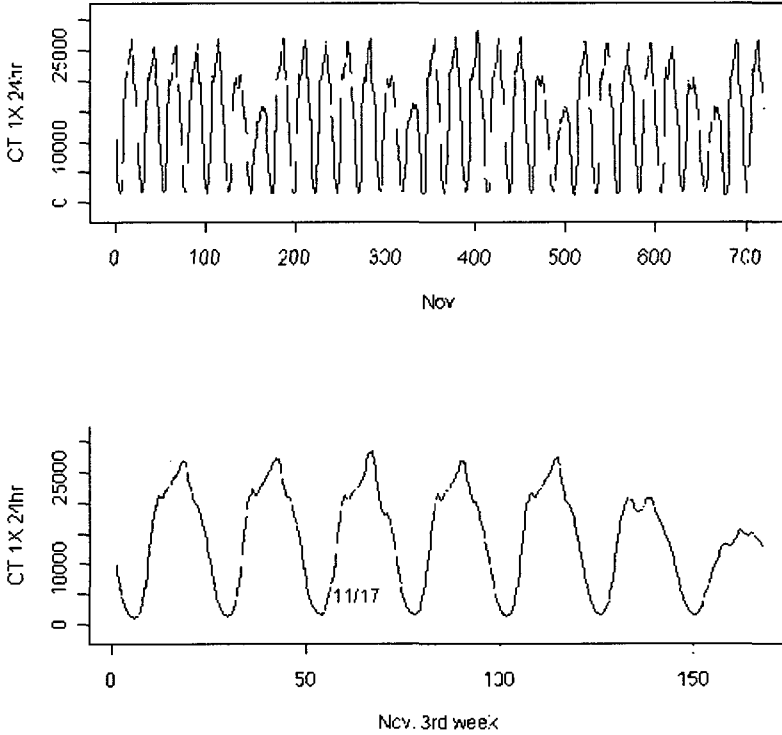


그림 3.1: 1X망 전국 CT 통화량의 2004년 11월 1개월(상)과 1주일(하)동안의 시간별 변화

앞서 살펴본 2004년 11월 동안의 시간별 CT 통화량 자료에서와 마찬가지로 CE통화량도 거의 동일한 변화 형태를 보이게 된다. 이때 일별로 그 최대값만을 취한 것이 일일 실적변시 자료에 해당한다. 일반적으로 분포 모수의 변화가 없는 것을 가정한 경우에서 얻어지게 되는

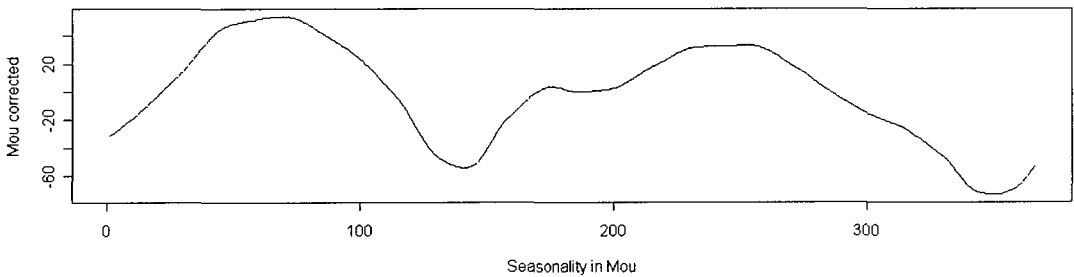


그림 3.2: MOU로부터 얻은 일년 365일의 일별 계절성 지수

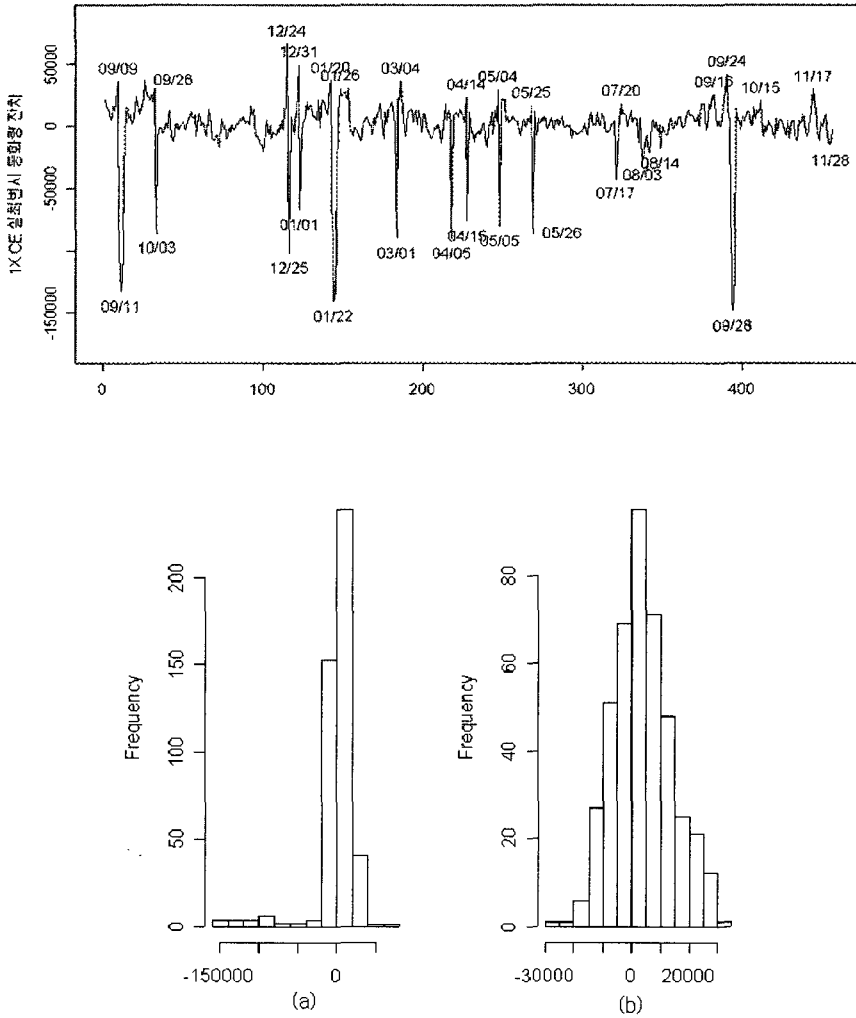


그림 3.3: 1X망 전국 CE 통화량에서 요일효과를 제거한 잔차의 시계열 그림과 히스토그램

최번 자료의 해석에는 극단값 이론을 고려해야 하는 등의 수리적 어려움이 있을 수 있다. 그러나 위에서 살펴본 바와 같이 최번시의 양상이 매일 동일한 형태의 변화 과정을 통하여 나타나고 있어서 일일 동안의 시간별 변화의 주원인이 확률적 현상에 의하여 일어난 것이라고 보기 보다는 분포의 평균이 시간에 따라 변화 하여 일어나는 것이라고 해석하는 것이 보다 합리적이다. 따라서 일일 실적변시 자료를 이용하는데 있어서 특별히 별도의 수리적 고려가 필요한 것으로 보이지는 않는다.

2004년 일최번시 통화량(1X는 2003년 9월에서 2004년 12월 사이의 실적변 CE 통화량, 2G

는 2003년 1월에서 2004년 12월 사이의 BHCA) 자료에 대하여 탐색적 방법을 통하여 자료를 검토하였다. 그림 3.3는 실적변 CE 통화량의 전국합에서 회귀분석을 통하여 요일효과 등 주요 효과들을 제거하고 남은 잔차에 대한 시계열 그림과 히스토그램들이다. 이 값들을 살펴보면 대부분의 잔차값들이 전체적으로 일정 범위 이내의 값을 갖는데 반하여, 추석이나 설 같은 명절과 크리스마스 또는 공휴일 전후와 같은 일부 특수 일들에 매우 큰 변동치를 보이는 것을 알 수 있다. 아래쪽 히스토그램 (a)는 이 잔차를 직접 이용하여 그린 것이다. 이 히스토그램에서 알 수 있듯이 통화량 자료의 특징은 꼬리가 두꺼운(heavy tail) 형태를 보이는 것을 알 수 있다. 이것은 일반적인 통화량 자료가 갖는 현상 중 하나이다. 그러나 이러한 잔차들에 대하여 몇 가지 시험을 하게 되면 좀 더 구체적인 분포적 특징을 살펴볼 수 있게 된다. 히스토그램 (b)는 동일한 잔차에 대하여 다음에서 설명하게 될 이상치(outlier) 제거 방법에 따라 이상치를 제거하고 남은 값들을 이용하여 그린 것이다. 일부 이상치를 제거한 후의 잔차들은 정규분포를 가정하기에 무리가 없는 특징을 보이고 있다. 이러한 특징은 전국이나 지역별로 집계된 통화량에서 뿐만 아니라 전체 기지국 각각에 대하여도 동일한 형태를 띄게 된다. 이상치를 제거한 후에 나타나는 정규성의 특징은 이론적 측면과도 부합된다. 보통 통화량 자료의 정태적 분포는, 통화자의 되먹임 가정을 하지 않는 경우 Erlang-B 분포를 갖는 것으로 알려져 있다. Erlang-B 분포는, 쉽게 말해서 최대 통화 처리자가 유한하다는 유한조건 하에서의 조건부 포아송 분포로 해석될 수 있는데, 이때 유한조건 값이 매우 크고 포아송 모수도 매우 큰 경우라면 당연히 정규분포와 매우 흡사한 형태의 분포를 갖게 된다. 통화량 분포의 경우 꼬리가 두꺼운 형태의 분포를 보이게 되는데 그 원인은 통화량 분포의 모수가 시간에 따라 계속적으로 변화하게 되기 때문이다. 만약 통화량 분포를 결정하는 모수의 변화 원인을 규명하고 그로 인한 변동의 대부분을 제거할 수 있게 된다면 당연히 그 주변 분포는 정규분포에 매우 가까운 형태를 갖게 될 것이다. 이에 근거하여 통화량 자료에 대하여 다음과 같은 모형화가 가능하다.

해당 시계열 번호를  $s$ 로 나타내고 관측일  $i$ 에서의 통화량을  $Y(s, i)$ 라고 할 때

$$\begin{aligned} Y(s, i) &= \mu_s(w_i, m_i, q_i, \dots) + e_s(i) + d_s(i) \cdot I_s(i) \\ &= \mu_s(w_i, m_i, q_i, \dots) + e_s(i) \cdot \bar{I}_s(i) + v_s(i) \cdot I_s(i) \end{aligned}$$

와 같은 성질을 갖는다. 여기서  $s$ 는 기지국을 의미할 수도 있고, 만약  $Y(s, i)$ 가 전국통화량의 합을 의미한다면  $s$ 는 전국을 의미한다고 해석된다.  $\mu_s$ 는 통화량 변화를 설명하는 요인들이 통화량 변화에 미치는 영향을 의미하는 비 확률적 값으로 정의된다.  $w_i$ 와  $m_i$ 는 각각  $i$ 일이 해당되는 요일과 월을 의미하는 명목변수이다.  $q_i$ 는  $i$ 일에 해당하는 가입자(Sub), 통신망 할당계획에서 나오는 변수 Trans, 그리고 MOU 계절성 등의 변수를 통칭적으로 의미한다. 또한 위에서  $v_s = e_s + d_s$  이고  $\bar{I}_s = 1 - I_s$ 라고 정의되고,  $e_s(i)$ ,  $I_s(i)$ 와  $v_s(i)$ 는 다음과 같이 두 가



지 방법으로 모형화하는 것이 가능하다.

$$\begin{aligned} \text{모형 I:} \quad e_s(i) &\sim \text{Normal}(0, \sigma_s^2) \\ I_s(i) &\sim \text{Bernoulli}(p_s) \\ v_s(i) &\sim F(0, \delta_s^2) \end{aligned}$$

$$\begin{aligned} \text{모형 II:} \quad e_s(i) &\sim \text{Normal}(0, \sigma_{s,m}^2) \\ I_s(i) &\sim \text{Bernoulli}(p_{s,m}) \\ v_s(i) &\sim F(0, \delta_s^2) \end{aligned}$$

위 모형들의 의미는 보통의 잔차들의 경우 평균이 0인 정규분포를 따른다는 것을 무리없이 가정할 수 있는데 비하여 일부 특수 일들의 경우는 평균이 0이고 매우 큰 산포를 갖는 어떤 분포  $F$ 를 갖는다고 보는 것이다.  $I_s(i)$ 는 0 또는 1인 값을 갖는 확률변수로 모형화 되었는데,  $I_s(i) = 1$ 인 경우는 명절이나 연휴와 같은 특수일을 의미하는 것으로 전국단위 통화량 예측의 경우만을 고려하는 경우 특정일  $i$ 가  $I_s(i) = 1$ 인 경우에 해당할지  $I_s(i) = 0$ 인 경우에 해당할지에 대하여는 상당한 정확성으로 예측이 가능하므로  $I_s(i)$ 를 꼭 확률변수로 모형화 해야 할 필요는 없으나, 명절효과나 연휴효과가 생각 외로 복잡한 특성을 갖는 측면도 있고, 이의 효과에 대한 객관적 자료의 축적이 미비하기도 할 뿐만 아니라, 동일한 모형을 개별 기지국 단위까지 적용한다고 할 때  $I_s(i)$ 의 값은 해당 지역의 특수 사정에 영향을 받는다는 점 등을 고려하면 확률적으로 모형화하는 것이 보다 더 일반적일 뿐만 아니라 매우 타당한 방법이라고 하겠다. 연휴나 명절 등과 같은 특수일의 비율이나 정상일에서의 오차항  $e_s(i)$ 의 산포는 해당 월  $m$ 에 따라 각기 달라질 수 있는데, 이를 고려하여 모형 I을 일반화 한 것이 모형 II가 된다. 다음에서는 기본적으로 모형 I을 중심으로 논의를 전개하기로 한다. 필요한 경우 특별한 언급없이 모수  $\sigma_{s,m}$ 이나  $p_{s,m}$ 를 언급하게 되는 경우는 모형 II를 가정하는 것이다.

위와 같은 모형하에서 추론 대상이 되는 것은  $\mu_s$ ,  $\sigma_s$ ,  $p_s$  와  $\delta_s$  이다. 특히  $\delta_s \gg \sigma_s$ 라는 내제적 가정을 엄두에 두고 볼 때 위의 모형들에 대한 추론 과정의 핵심 사항은 회귀모형하에서 이상치를 찾는 방법이다. 이에 대하여는 이미 많은 연구들이 이루어져 있고, 통계적으로 매우 우수한 성질을 갖는 알고리즘들이 알려져 있다(참고: Kianifard와 Swallow, 1996). 그러나 본 연구에서는 효율성 측면보다는 현장 사용자들이 쉽게 이해할 수 있어야 한다는 점을 감안하여 다음과 같이 단순화된 알고리즘을 사용한다.

Step 1: 전체 자료에 대하여 회귀모형을 적합하고, 잔차  $r_i$ 들에 대하여 표준편차  $s_0$ 를 구한다.

Step 2:  $|r_i/s_0| > 3$ 이 되는  $i$ 번째 관측치를 자료에서 제거한다.

Step 3: 새로이 제거된 자료가 없으면 종료한다. 새로이 제거 변경된 자료를 이용하여 Step 1, Step 2를 반복한다.

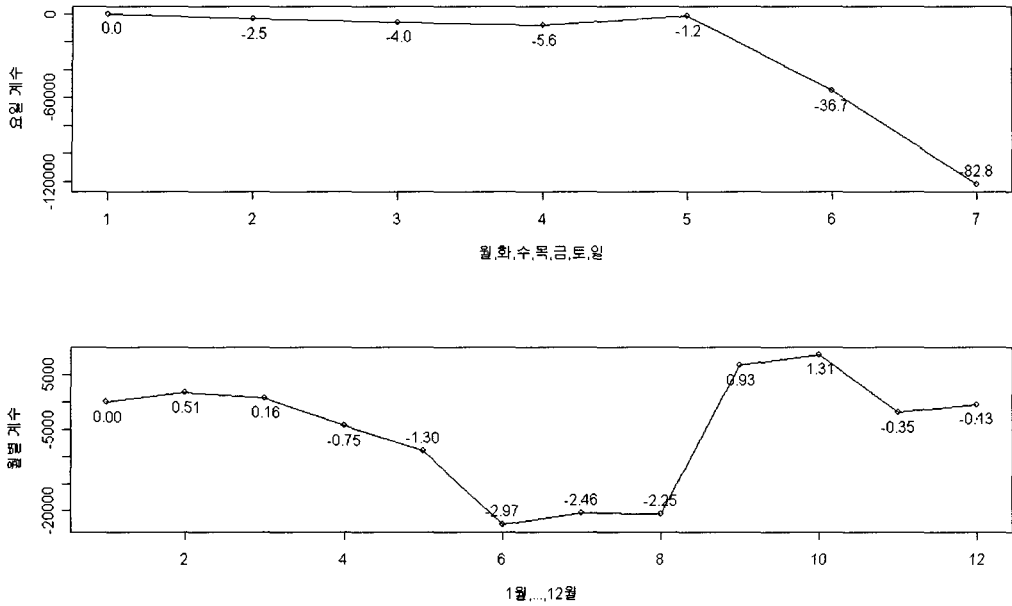


그림 3.4: 1X망 전국 CE 통화량에서 요일효과(위쪽)와 월별효과(아래쪽)

이 과정을 통하여 최종적으로 얻어지는 표준편차  $s_0$ 는 실제 추정하려는  $\sigma_s$ 를 약간(98.5%) 과소 추정하는 경향이 있으므로 이후의 사용에 있어서 이를 감안하여야 한다. 그림 3.4는 전국의 각 기지국별 1X CE 통화량의 일별 최대통화량의 합에 대한 추론에서 평균함수  $\mu_s$ 를 적합했을 때 얻어지는 요일별 효과(위쪽 그림)와 월별 효과(아랫쪽 그림)를 그림으로 나타낸 것이다. 그림에서 해당 점들 옆에 나타난 숫자는 각 값의 t-value이다. 절편의 추정치와 함께 설명변수로 사용된 Trans, 1X 단말기 보유가입자 수(Sub, 단위 천명), 그리고 MOU 계절성으로부터 얻는 계절 변동지수에 대한 추정치는 각기, -32496.5, -46.6, 39.7, 55.1이고 이들 추정치들에 대한 t-value는 각기 -0.41, -7.6, 6.0, 1.29이다. 이 때 t-value 들의 자유도는 319이다.

표 3.1: 1X망 각 기지국별 실최변 CE 통화량의 전국합에 대한 평균함수

지역	부산	수도권	충북	충남	대구	강원	전북	제주	전남	전국
Trans	-66.8	-56.3	-81.3	-55.4	-51.7	-54.6	-77.0	-36.7	-62.4	-46.6
Sub	48.3	43.7	57.3	45.8	41.3	36.1	37.8	31.6	32.8	39.7
MOU 계절성	0	0	5.7	7.7	9.7	0.7	1.8	0	0	55.1

동일한 모형을 이용하여 1X CE 통화량에 대한 전국 9개 지역별 모형을 적합했을 때, Trans, 가입자 수, MOU 계절성이 갖는 모형에서의 계수는 표 3.1에 요약되어 있다. MOU 계절성이 설명력이 없다고 나타난 경우들에 대하여는 모형에서 변수를 제거하였다. 그 경우 계수가 0으로 나타났다. 또한 전국 9개 각 지역별 통화량 자료를 앞서의 모형 I을 적용했을 때

표 3.2: 각 지역별 1X CE 실최번 통화량 합에 대한 모형 모수

	$\delta_s$	$\sigma_s$	$p_s$	Max $\sigma_{s,m}$	Min $\sigma_{s,m}$	Max $p_{s,m}$
부산	10244.2	1080.56	0.06284	1412.7	830.10	0.1613
수도권	41053.7	3906.22	0.08743	4827.0	2214.93	0.1935
충북	1477.7	172.49	0.06831	245.8	100.69	0.2258
충남	3486.6	530.43	0.05738	683.6	369.27	0.1935
대구	5556.9	686.95	0.06831	920.0	513.92	0.2258
강원	1105.7	183.41	0.07650	274.1	114.22	0.2258
전북	1831.3	254.50	0.05464	347.4	174.86	0.1613
제주	409.5	80.02	0.06011	96.8	63.48	0.1613
전남	3311.6	441.86	0.07377	614.9	304.34	0.2333
전국	68083.6	7346.38	0.06831	9745.0	5636.33	0.1935

얻어지는  $\delta_s$ ,  $\sigma_s$ ,  $p_s$  그리고 모형 II를 적용했을 때 얻어지는  $\sigma_{s,m}$ ,  $p_{s,m}$ 들에 대한 월별 최대 값  $\text{Max}_m\{\sigma_{s,m}\}$ ,  $\text{Max}_m\{p_{s,m}\}$ 와 최소값  $\text{Min}_m\{\sigma_{s,m}\}$ 은 표 3.2에 나타나 있다. 각 경우에서  $\text{Min}_m\{p_{s,m}\}$ 은 모두 0으로 나타나 별도로 나타내지 않았다.

#### 4. 대표통화량의 설정

앞 절에서 통화량 자료들이 갖는 통계적 성질을 모형화하였고 그 결과를 살펴 보았다. 통화망 관리의 가장 좋은 방법은 앞서 정의된 모형 모수들을 개별적으로 관리하는 것이라고 하겠다. 그러나 앞서 서술된 통계적 모형에 의한 통화망 특성 파악 방법은 아무래도 상당한 정도의 선행 통계 학습이 되어 있는 경우에 이해될 수 있다는 단점이 있고, 각각의 모형 모수의 변화를 효율적으로 검출하고 관리하는 것은 또 별도의 연구가 필요한 부분이기도 하다. 이러한 이유로 실제에서는 기지국별로 월간 통화량의 변화를 가장 잘 대표할 수 있는 값을 정하고 그 값을 중심으로 통화망 관리를 하고 있다. 앞서 2절에서 설명한 대표통화량이 바로 이런 목적으로 만들어진 것이다. 따라서, 기존 방법의 틀에서 보다 효율적인 통화망 관리가 가능하도록 합리적인 대표통화량 설정 방법을 고려할 필요가 있다. 이런 측면에서 대표통화량 설정을 위하여 다음 사항들을 그 기본 추구 목표로 삼았다.

- 기존의 방법과 가능한 일치성을 보이는 것이 좋다.
- 건실한 논리적 근거를 가져야 한다.
- 기지국별, 지역별, 망별로 다양한 통화량에 가능한 동일하게 적용할 수 있어야 한다.
- 구하는 과정이 크게 어렵지 않아야 한다.
- 현실 통화량의 변화를 잘 반영할 수 있어야 한다.

이러한 조건 하에서 앞서 3절에서 설정된 통화량 모형의 특성을 이용해서, 훈련 자료로부터 얻은  $\mu_s$  함수를 이용하여  $\mu_s(w_i, m_i, \dots)$ 를 얻고 이로부터

$$g_s(i) = e_s(i) + d_s(i)I(i) = Y(s, i) - \mu_s(w_i, m_i, \dots)$$

를 얻은 후에 다음과 같이 대표통화량  $R(s, m)$ 을 결정하는 방법을 고려하였다.

$$R(s, m) = \max_{i \in m} \mu_s(w_i, m, q_i, \dots) + \bar{g}_s(m)$$

여기서  $\bar{g}_s(m) = \text{mean}_{i \in m} g_s(i)$ 를 말한다. 평균함수  $\mu_s$ 는 예상이 가능한 통화량의 변화 또는 이유있는 통화량 변동 등의 의미로 해석할 수 있고,  $g_s$ 는 확률적 변인에 기인하는 통화량 변동으로 해석할 수 있다. 그러므로 앞서 정의된 대표통화량은 이유있는 혹은 예측이 가능한 통화량의 변동에 대하여는 최대값을 취하고, 관리할 수 없는 예측이 어려운 값에 대하여는 평균을 취하여 구한 값으로 해석할 수 있다. 따라서 확률적 변동 요소들에 대하여는 그 변인을 가장 잘 대표할 수 있게 하는 평균을 취한 이유로 이후 다양한 통계적 추론 과정에서 편리하게 사용할 수 있는 장점도 아울러 가지고 있다. 뿐만 아니라 그림 4.1에서 확인 할 수 있는 바와 같이 다양한 특성을 가진 기지국들에 대하여 대부분 기존의 대표 통화량과 크게 벗어나지 않는 모습을 보여 주고 있다.

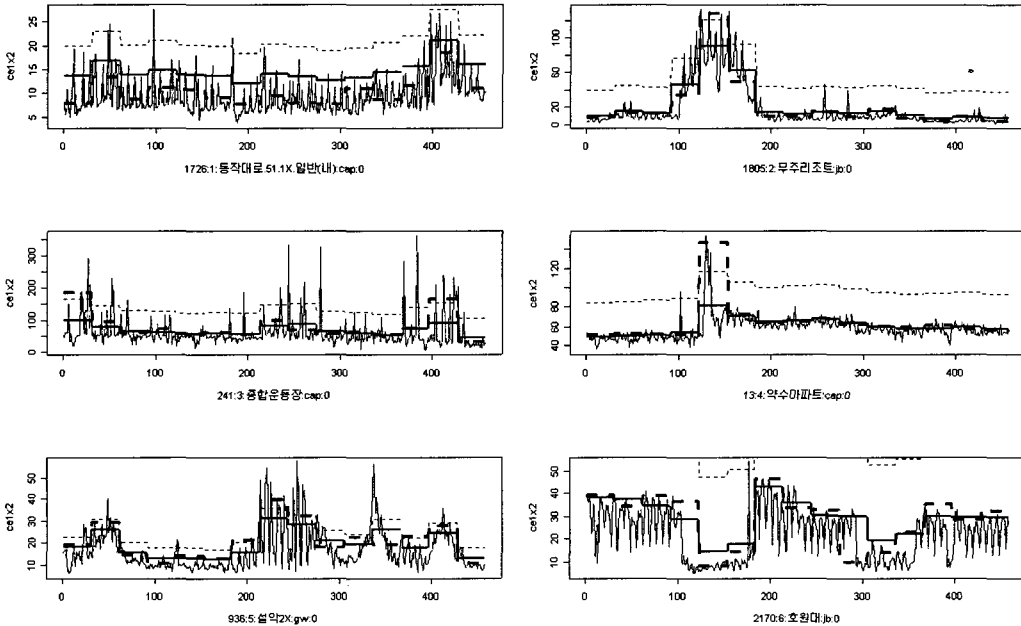


그림 4.1: 접속망에서의 대표통화량의 비교

그림 4.1에서 굵은 실선은 새로 정의된 대표통화량이고, 굵은 점선은 기존에 사용되어 오던 대표통화량이다. 그림에서 보여진 기지국들은 각기 독특한 통화량 특성을 가진 대표적인 기지국들인데 이들에 대하여 새로 정의된 대표통화량이 매우 합리적으로 움직이고 있음을 볼 수 있다. 새로 정의된 대표통화량은 통화량 변화를 잘 따라 가면서도, 약수 아파트의 경우에서 볼 수 있듯이 기존의 대표통화량에 비하여 일시적 변화에 대한 안정성을 가지고 있음을 볼 수 있다. 또한 동작대로의 경우에서 볼 수 있듯이 상시적 통화량 변동은 반대로 충실히 잘 반영하고 있음도 살펴 볼 수 있다. 그림 4.1에서의 가는 점선은  $R(s, m) + \delta_s$ 를 나타내는 선이다.

실제 자재소요계획 등에 대표통화량을 사용하고자 하는 경우는, 앞 절에서 얻어진 통화량에 변동 해석에서 얻어진 변동 요소들을 필요에 맞게 가감하여 사용하는 것도 가능하다. 중심망의 경우를 그림 4.2에서 비교하여 보면 기존의 중심망 대표통화량 산정 방법  $M + cS$ 는  $R(s, m) + \delta_s$ 와 상당한 일치성을 보여주고 있음을 알 수 있다.

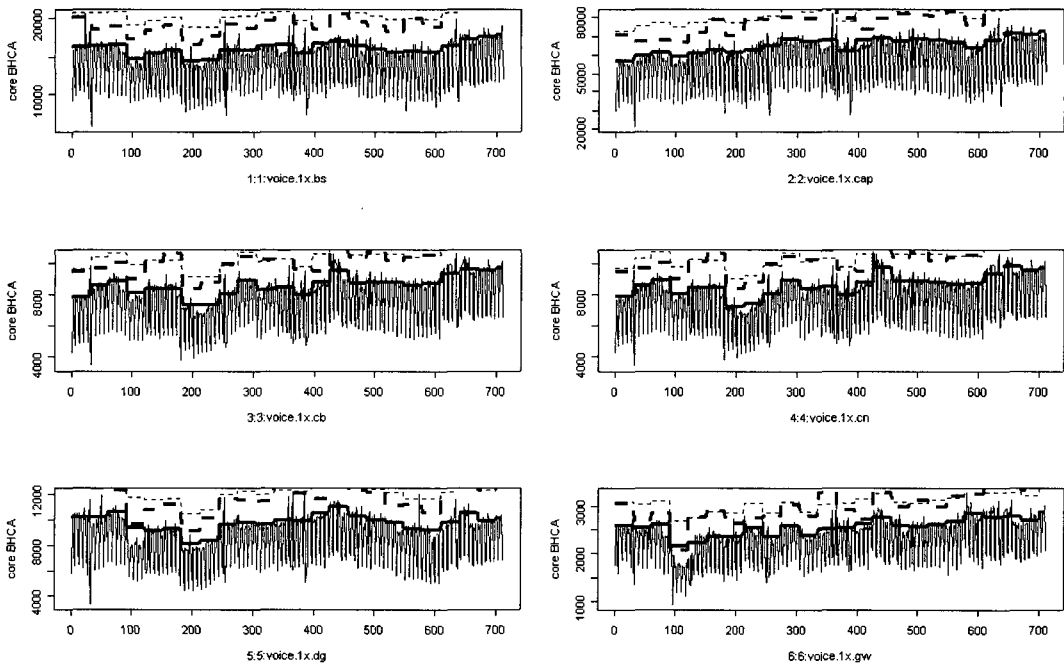


그림 4.2: 중심망에서의 대표통화량의 비교: 부산(bs), 수도권(cap), 충북(cb), 충남(cn), 대구(dg), 강원(gw)

## 5. 맺음말

앞에서 살펴본 통화량 데이터의 변동은 다양한 변인에 의하여 설명될 수 있다. 그 첫번째는 하루 내에서 매시간 변하는 일일변동이다. 다음으로 큰 변동은 명절이나 연휴등의 특수일에 나타나는 변동이다. 그 다음 순서의 변인으로 고려할 수 있는 것이 주중이나 주말이나에 따르는 혹은 요일에 따르는 요일변동이다. 이 세 가지 주요 변동이 전체 통화량 변동의 대부분을 차지하게 된다. 그 외 다른 변인에 의한 통화량 변동의 효과를 보기 위해서는 위의 세 가지 원인에 따라 발생하는 변동을 어떻게 효율적으로 제거 할 수 있는 지가 관건이다. 특히 망 증설 혹은 감설 계획을 세우고자 하는 입장에서 더 큰 관심이 되는 것은 가입자의 증감 또는 통화망 운용 방식이나 영업 정책의 변화에 어떻게 망 부하가 변화 하는 지가 주 관심 사항이 될 것이다. 이러한 관심 사항이 되는 변동은 앞서 말한 자연 주기적 변인에 의한 변동에 비하여 상대적으로 그 크기가 작아서 그 효과를 쉽게 검출하기 어려운 점이 있어 매우 조심스런 변동 해석이 요구된다.

또한 이동통신망의 경우는 단말기의 공간 위치 변동에 따르는 새로운 변동 요인이 함께 어우러져 관심 변인에 대한 효과 파악을 더욱 힘들게 만드는 점도 있다. 앞서 살펴본 변인 중 Trans의 효과는 대규모 지역 단위에서는 그 통화량 증감 효과가 잘 검출되는데 비하여 기지국과 같은 소규모 지역 단위에서는 그 증감 효과가 잘 검출되지 않는다. 그 원인으로서는 Trans의 효과가 기지국 단위에서 상대적으로 미약하다는 점 외에도 이와 같은 통화자 지역 이동의 효과도 개입이 되어 있는 것으로 판단된다. 이동통신 통화량 자료는 이러한 이유로 일반 전화의 통화량 자료에 비하여 분석이 더 어려운 점은 있으나 통화자의 공간 정보를 포함하고 있다는 점에서 더 많은 이용 가능성이 있는 결과를 줄 수 있는 자료이고 보다 다양한 분석 노력이 필요할 것으로 생각된다.

## 참고문헌

- Brockmeyer, E., Halstrom, H. L. and Jensen, A. (1948). *The life and works of A.K. Erlang*, Copenhagen: The Copenhagen telephone company.
- Halfin, S. and Whitt, W. (1981). Heavy traffic limits for queues with many exponential servers, *Operations Research*, **29**, 567-588.
- Kianifard, F. and Swallow, W. H. (1996). A review of the development and application of recursive residuals in linear models, *Journal of the American Statistical Association*, **91**, 391-400.

[ 2005년 8월 접수, 2005년 9월 채택 ]

## Traffic Summary for Analyzing Network Load in Mobile Communication System\*

Y. D. Lee<sup>1)</sup> S. G. Koh<sup>2)</sup> B. J. Ahn<sup>3)</sup>

### ABSTRACT

In this paper, we propose a method to summarize the monthly traffic amount for analyzing network load in mobile communication system. We used the traffic data obtained from a domestic telecommunication company. Based on the statistical properties of the traffic data, we devise an efficient method to summarize monthly traffic amount.

*Keywords:* Access network, Traffic summary, Outlier.

---

\* This work was supported by the Korea Research Foundation Grant.(KRF-2006-C00056)

1) Assistant Professor, Applied Statistics, Konkuk University, Seoul 143-701, Korea

E-mail: poisson@dreamwiz.com

2) Associate Professor, Applied Statistics, Kyungwon University, Sunnam, Kyonggi 461-701, Korea

E-mail: sgk@kyungwon.ac.kr

3) Professor, Applied Statistics, Konkuk University, Seoul 143-701, Korea

E-mail: bjahn@konkuk.ac.kr