

# 이동통신서비스 해지자수 예측모형: 셀룰라폰 서비스의 경우

전덕빈\* · 임철주\* · 주영진\*\* · 박윤서\* · 박명환\*\*\* · 박영선\*\*\*\*

## Forecasting Modeling for the Number of Mobile Telecommunication Service Cancelers: Application to Cellular Phone Service

Duk Bin Jun · Churl Zu Lim · Young Jin Joo · Yoon Seo Park  
Myoung Hwan Park · Young Seon Park

### 〈요 약〉

이동통신서비스 가입자수에 대한 예측은 이제까지 주로 확산모형[2][3][4](diffusion model)이나 로지스틱 모형에 의한 신규가입자수의 예측을 위주로 발전되어 왔다. 그러나, 가입자들의 통신요구의 변화 및 개인적, 사회적 환경의 변화에 따라 이동통신서비스 가입자들 중 다수가 서비스를 해지하고 있다. 이 경우 특정 이동통신서비스의 기존 가입자 중에서 발생하는 해지자수를 정확히 예측하는 것은 해당 이동통신서비스 가입자수를 합리적이고 정확히 예측하기 위해 매우 중요한 문제이나, 기존에는 거의 연구가 이루어지지 않거나 설문조사에 따른 방대한 자료의 분석에 의존하여 왔다. 본 논문에서는 이동통신서비스의 기존 가입자가 지니는 해지 특성을 가입시기에 따른 해지성향과 특정 시점의 가입자가 나타내는 해지까지의 가입수명분포로 정의하고, 기존의 해지자수 자료에서 해지성향 및 가입수명분포를 추정함으로써 해지자수를 예측할 수 있는 모형을 개발하여, 이를 기존의 셀룰라폰 서비스 해지자수 예측에 적용하였다.

## 1. 서론

이동통신서비스 시장의 환경변화에 대한 능동적 대응 및 적절한 전략수립을 위해서는 서비스 수요에 대한 정확한 예측이 필수적이다. 시장의 확대에 따른 시설의 확충 및 관리, 고객관리 및 마케팅 전략수립 등과 관련된 경영의사결정을 지원하고 기업경영의 합리화를 도모하기 위한 목적 뿐 아니라 이동통신서비스 산업에 대한 과소 및 과다 투자를 방지하여 산업구조

를 건실화하는데도 필수적이다. 이동통신서비스 시장 분석에 있어서 주된 관심인 서비스가입자수는 크게 누적가입자수, 신규가입자수 그리고 가입자순중의 3가지로 구분된다. 여기서 누적가입자수는 특정 시점에 가입되어 있는 가입자의 총수, 신규가입자수는 특정 기간에 새로 가입한 사람의 수, 가입자순중은 그 기간말의 누적가입자수에서 기간초의 누적가입자수를 뺀 수를 지칭한다. 이상의 수치들은 이동통신서비스기업의 전략적 의사결정에 있어서 매우 중요한

\* 한국과학기술원 테크노경영대학원 (KAIST, Graduate School of Management)

\*\* 한국전자통신연구소 통신경영연구실 (Electronics and Telecommunications Research Institute)

\*\*\* 한성대학교 산업공학과 (Han Seong Univ.)

\*\*\*\* 서경대학교 산업공학과 (Seo Kyoung Univ.)

변수로 작용하게 된다. 서비스가입자수들의 관계는 식 (1)과 같다.

$$\begin{aligned} \text{누적가입자수} &= \text{전기의 누적가입자수} + \text{가입자순증} \\ \text{가입자순증} &= \text{신규가입자수} - \text{해지자수} \end{aligned} \quad (1)$$

어떤 시점의 누적가입자수는 바로 전 시점의 누적 가입자수와 그 기간동안 늘어난 가입자순증의 합으로 구해진다. 가입자순증은 그 기간 동안의 신규가입자 수와 해지자수의 차로써 구한다. 따라서 서비스가입자수의 예측치를 구하기 위해서는 신규가입자수와 해지자수의 예측이 필요하다. 그런데 기존의 이동통신 서비스 가입자수 예측에서는 대부분 신규가입자수 예측만을 시도하였거나 가입자순증만을 예측하였다. 따라서 각각의 중요성을 가지고 있는 누적가입자수, 신규가입자수, 가입자순증을 구별하여 예측할 수 없는 단점이 있었다. 그러므로 신규가입자수와 함께 해지자수의 예측도 동일한 비중으로 다루어져야 한다.

신규가입자수에 대한 예측은 주로 확산모형을 이용하여 수행되어왔다. 순증가입자수는 신규가입자수의 예측치에서 그 기간의 해지자수의 예측치를 빼줌으로써 구해진다. 따라서 해지자수의 예측이 필요해지는데 이제까지는 단순히 해지율을 정의하여 해지율의 추세를 분석하거나 신규가입자수 예측에 사용되는 확산모형에 해지자를 포함시켜 선택모형[4](choice model)을 이용하여 예측하였다. 기업에서 실무적으로 사용되는 해지율은 다음의 두가지가 있다.

$$R_{1t} = \frac{H_t}{M_t}, \quad t=1,2,3,\dots \quad (2)$$

$$R_{2t} = \frac{H_t}{N_t}, \quad t=1,2,3,\dots \quad (3)$$

$R_{1t}$  =  $t$  기의 해지율

$H_t$  =  $t$  기의 해지자수

$M_t$  =  $t$  기의 누적가입자수

$N_t$  =  $t$  기의 신규가입자수

위와같이 해지율을 정의한 경우 직관적 이해가 부족하다. (2)와 같이 정의하면 한 시점의 총가입자수의

일정비율로 나타나는데 그 시점의 가입자들의 가입기간이 모두 다르므로 적절치 못하다. 즉, 서비스제공 초기시점에는 서비스가입기간이 짧기 때문에 누적가입자중 해지하는 비율이 상대적으로 작으며 시간이 흐름에 따라 가입기간이 긴 가입자들의 비율이 늘어나 해지하는 비율이 상대적으로 커질 것이다. 그리고 확산모형에 의하면 누적가입자수는 S자 형태를 나타내게 되어 포화수준에 다다랐을 때 가입기간의 분포는 종모양의 형태가 될 것이므로 가입기간에 대한 편倚(bias)가 발생한다. (3)과 같은 경우도 신규가입자수의 일정비율로 나타나는데 신규가입자수가 많을 때는 해지자수도 많아야 하는 결과가 나타나 (2)와 비슷한 문제가 있다. 확산모형에 해지자수를 포함시켜 선택모형을 사용한 경우에는 방대한 설문자료를 다루어야 하기 때문에 비용이 커진다.

본 논문에서는 통신서비스의 기존 가입자가 지니는 해지특성을 가입시기에 따른 해지성향과 가입시점으로부터 해지시점까지의 수명분포로 정의하여 직관적 이해는 물론 자료의 분석에 따른 비용이 적은 합리적인 예측모형을 제시하고 이를 응용하여 셀룰라폰 서비스의 기존 해지자수 자료로부터 해지성향과 수명분포를 추정하여 해지자수의 예측에 적용하였다.

## 2. 해지모형

본 연구에서 제시되는 모형은 다음과 같은 자료가 있음을 전제로 한다.

① 2기 이상의 해지자수와 해지자의 수명

② 서비스 시작시점부터의 신규가입자수

모형에서 기본적으로 가입자의 특성이 둘로 나뉘는데 가입시기에 따른 해지성향과 가입시기와는 무관한 수명분포이다. 해지성향은 서비스가입자가 가입한 시기의 서비스품질이나 사회적상황 등으로부터 가입자의 서비스이용행태에 받는 영향을 의미한다. 수명분포는 가입시기와는 관계없이 서비스가입자가 공통적으로 가지는 가입기간의 분포이다.

본 연구에서는  $t$ 기의 해지자수를 다음과 같이 정의한다.

$$H_t = \sum_{i=0}^n N_{t-i} \times a_{t,i} \times b_i, \quad \sum_{i=0}^n b_i = 1 \quad (4)$$

$N_t$  :  $t$ 기의 신규가입자수

$a_t$  :  $t$ 기의 신규가입자중  $(t+n)$ 기안에 해지하는 비율

$b_i$  : 수명이  $n$ 기이내인 가입자의 수명분포

이때  $n$ 은 서비스가입자중 해당서비스에 대한 선호도가 강하여 장기간 가입하는 가입자만 남아있게 되는 시점을 의미한다. 즉,  $t$ 기에 서비스에 가입한 사람들중 사망, 단말기대체, 서비스종료 등의 자연발생적 해지 이외의 의도적 해지가 발생하게 되는 기간을 뜻한다.  $a_t$ 는  $t$ 기에 가입한 사람이  $t+1$ 까지  $n$ 기동안 서비스에 대한 불만족이나 경제적이유 등으로 의도적인 해지를 하게되는 비율을 의미한다.  $b_t$ 는 의도적 해지가 진행되는  $n$ 기간동안의 해지분포를 의미한다. 즉,  $b_t$ 는 가입한 후  $t$ 시간후에 해지할 확률이다.  $b_t$ 는  $t$ 에 존하지 않는다. 위의 수식을 도표화하면 <표 1>과 같다. <표 1>에서 4기의 해지자수는 해당 열의 합이다. <표 1>은 다시 해지자가 가입한 시기의 신규가입자수로 나눈 <표 2>와 같이 나타낼 수 있다. 이제 <표 2>로부터  $a_t$ 와  $b_t$ 를 추정한다.

<표 1> 신규가입자수로 나타낸 해지자수

$t$	의도적 해지자	1	2	3	4
$N_1$	$N_1 a_1$	$b_1(N_1 a_1)$	$b_2(N_1 a_1)$	$b_3(N_1 a_1)$	$b_4(N_1 a_1)$
$N_2$	$N_2 a_2$		$b_1(N_2 a_2)$	$b_2(N_2 a_2)$	$b_3(N_2 a_2)$
$N_3$	$N_3 a_3$			$b_1(N_3 a_3)$	$b_2(N_3 a_3)$
$N_4$	$N_4 a_4$				$b_1(N_4 a_4)$

<표 2> 해지자수의 분포

의도적 해지자	1	2	3	4
$N_1 a_1$	$b_1 a_1$	$b_2 a_1$	$b_3 a_1$	$b_4 a_1$
$N_2 a_2$		$b_1 a_2$	$b_2 a_2$	$b_3 a_2$
$N_3 a_3$			$b_1 a_3$	$b_2 a_3$
$N_4 a_4$				$b_1 a_4$

### 3. 추정 및 예측

추정방법은 크게 두가지를 고려할 수 있다. 하나는 최소화승법을 이용하여 추정하는 것이고 다른 하나는 일관된 셀들로 구성된 연립방정식을 풀어 구한 값들의 평균을 이용하는 방법이다. <표 2>의 자료가 주어졌을 경우 두 방법을 비교하면 다음과 같다.

#### 1) 최소화승법

$$\min \sum_{i=1}^4 \sum_{j=1}^{5-i} (\widehat{b}_i a_j - c_{ij})^2 \quad (5)$$

$c_{ij}$  : 주어진자료의셀값

식(5)를 만족시키는 모수를 구한다.

#### 2) 연립방정식을 이용한 방법

미지수가 8개이므로 일정한 기준에 의해 10개의 셀중 8개의 셀을 선택하여 연립방정식을 세운다. 여러개의 조합을 선택하여 각각 방정식을 풀어 구한 해들의 평균값을 사용한다.

$n$ 이 작은 경우에는 모수의 개수가 적기 때문에 최소화승법을 사용하면 최적화과정에서 최소값으로 수렴하는 해를 찾을 수 있으나,  $n$ 이 큰 경우에는 모수의 개수가 많기 때문에 최소값으로의 수렴을 보장할 수 없다. 따라서 이경우에는 연립방정식을 이용한 방법이 효과적이다.

앞에서 구한  $b_t$ 는 가입자의 수명분포로서 사용되며  $a_t$ 는 가입자의 가입시기에 따른 해지특성치가 된다. 시계열 분석방법중 하나인 ARIMA(Autoregressive Integrated Moving Average) 모형을 이용하여 원하는 시점의 가입자의 해지특성치를 예측할 수 있다. 따라서 해지자수의  $i$ -step-ahead forecast는 다음과 같이 구해진다.

$$\widehat{H}_{i,t} = \sum_{j=0}^{i-1} \widehat{b}_j \times \widehat{a}_{t+i-j} \times \widehat{N}_{t+i-j} + \sum_{j=i}^n \widehat{b}_j \times \widehat{a}_{t+i-j} \times \widehat{N}_{t+i-j} \quad (6)$$

, where  $\widehat{N}_t$  =  $t$ 기의 신규가입자수 예측

$\widehat{a}_t$  =  $t$ 기의 가입자의 해지특성 예측치

신규가입자수 예측치는 이미 확산모형 등을 이용해 구해진 신규가입자수 예측치를 사용한다.

#### 4. 셀룰라폰 서비스 해지자수예의 응용

여기에 주어진 자료는 1994년 1분기부터 1995년 3분기까지 7분기동안의 해지자수와 해지자의 수명 그리고 서비스실시 시작분기부터의 신규가입자수이다. 7분기동안의 해지자들을 대상으로 각 분기별 수명분포를 plotting한 결과가 <그림 1>과 같으며 이를 다시 <표 2>와 같은 형태로 나타낸 것이 <표 3>에 나와있다. <그림 1>로부터 가입후 12분기 이후에는 거의 해지자가 없는 것으로 미루어 볼 수 있으므로 아날로그 셀룰라폰 해지자수 예측에서는  $n=14$ 를 사용하였다. 즉 가입한지 3년 6개월이상 지속된 가입자는 아날로그 셀룰라폰 서비스에 대한 선호도가 높다고 가정하였다.

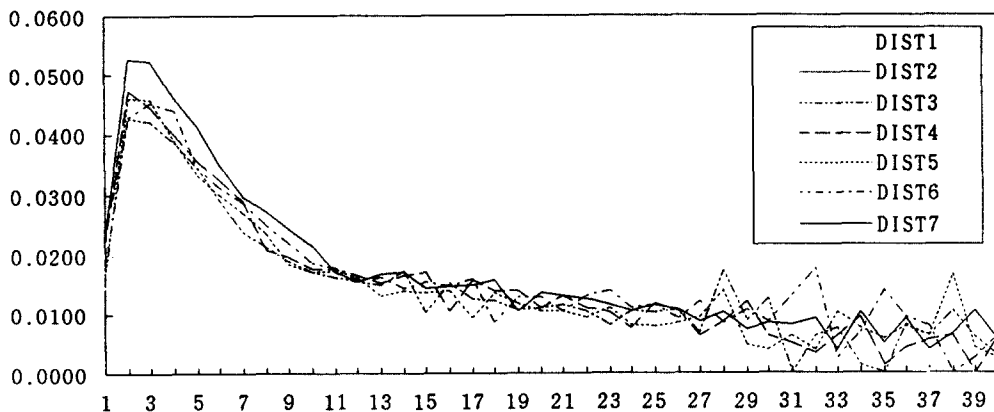
이상의 자료에서 추정해야하는 모수는  $a_1, \dots, a_n, b_0, \dots, b_n$ 이다. 추정해야할 모수가 많으므로 여기서는 연립방정식을 이용한 방법을 사용하였다. 셀을 고르는 기준은 다음과 같다.

- 1) 가장 최근(마지막 분기)의 자료를 이용한다.
- 2) 나머지 분기를 한번씩 이용한다.
- 3) 사용되지 않은 분기중에서 가장 최근의 셀을 하나씩 선택한다.

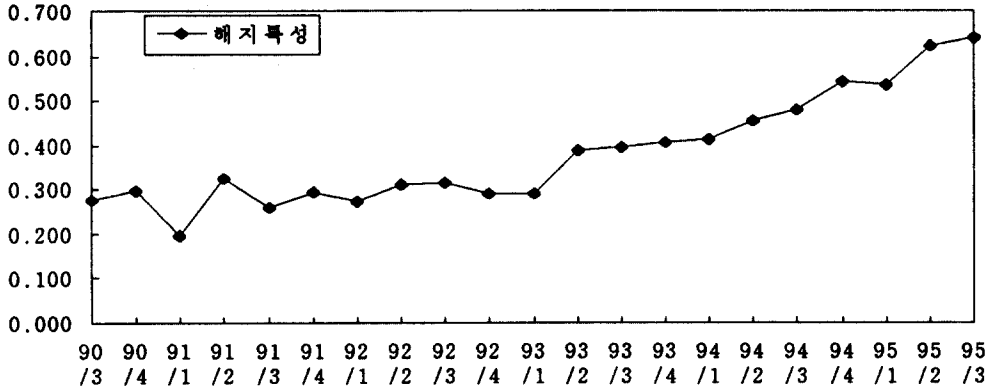
<표 3> 아날로그 셀룰라폰 해지자수 분포

94/1	94/2	94/3	94/4	95/1	95/2	95/3
0.0144	-	-	-	-	-	-
0.0168	0.0156	-	-	-	-	-
0.0108	0.0104	0.0102	-	-	-	-
0.0187	0.0183	0.0174	0.0171	-	-	-
0.0165	0.0141	0.0149	0.0166	0.0135	-	-
0.0164	0.0162	0.0154	0.0152	0.0138	0.0154	-
0.0169	0.0169	0.0162	0.0166	0.0131	0.0144	0.0143
0.0198	0.0165	0.0170	0.0175	0.0163	0.0163	0.0171
0.0223	0.0196	0.0188	0.0175	0.0172	0.0156	0.0168
0.0224	0.0209	0.0213	0.0197	0.0170	0.0178	0.0154
0.0263	0.0239	0.0239	0.0208	0.0183	0.0187	0.0169
0.0365	0.0305	0.0291	0.0288	0.0237	0.0221	0.0214
0.0425	0.0361	0.0347	0.0326	0.0271	0.0250	0.0243
0.0395	0.0387	0.0387	0.0357	0.0300	0.0286	0.0272
0.0145	0.0369	0.0423	0.0402	0.0334	0.0311	0.0296
	0.0169	0.0426	0.0445	0.0391	0.0347	0.0351
		0.0233	0.0474	0.0458	0.0439	0.0415
			0.0219	0.0461	0.0452	0.0461
				0.0166	0.0429	0.0523
					0.0179	0.0527
						0.0234

이러한 기준은 두 분기의 조합으로 이루어진 방정식 가운데 가장최근의 자료가 미래의 상황을 가장 잘 반영하고 있다는 가정하에 마지막 분기의 자료를 반



<그림 1> 해지자의 수명분포



〈그림 2〉 추정된 해지특성

드시 이용하였으며, 나머지 필요한 방정식 중에서도 가장 최근의 자료를 사용한 것이다.이상의 기준으로 조합된 6개의 방정식의 해를 다시 평균하여 얻은 최종 추정치는 다음과 같다.

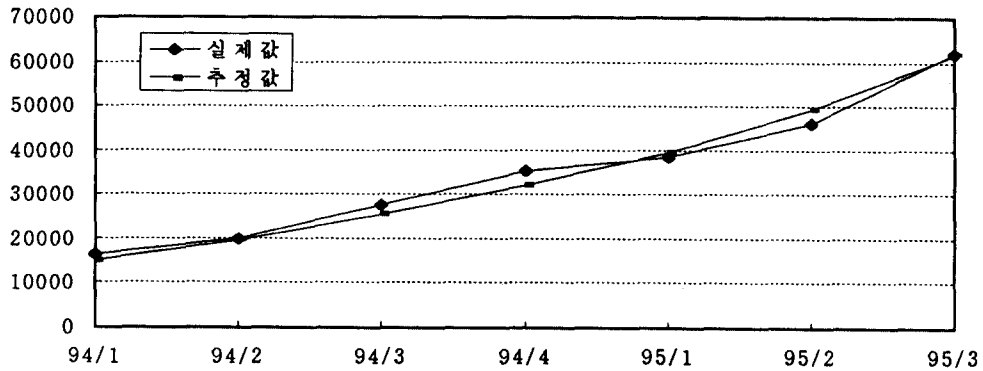
- |                  |                  |
|------------------|------------------|
| $a_1 = 0.275$    | $b_0 = 0.038$    |
| $a_2 = 0.298$    | $b_1 = 0.090$    |
| $a_3 = 0.195$    | $b_2 = 0.102$    |
| $a_4 = 0.327$    | $b_3 = 0.086$    |
| $a_5 = 0.258$    | $b_4 = 0.088$    |
| $a_6 = 0.294$    | $b_5 = 0.077$    |
| $a_7 = 0.273$    | $b_6 = 0.071$    |
| $a_8 = 0.313$    | $b_7 = 0.066$    |
| $a_9 = 0.316$    | $b_8 = 0.060$    |
| $a_{10} = 0.292$ | $b_9 = 0.054$    |
| $a_{11} = 0.291$ | $b_{10} = 0.057$ |
| $a_{12} = 0.387$ | $b_{11} = 0.053$ |
| $a_{13} = 0.396$ | $b_{12} = 0.053$ |
| $a_{14} = 0.406$ | $b_{13} = 0.054$ |
| $a_{15} = 0.413$ | $b_{14} = 0.052$ |
| $a_{16} = 0.455$ |                  |
| $a_{17} = 0.480$ |                  |
| $a_{18} = 0.543$ |                  |
| $a_{19} = 0.537$ |                  |
| $a_{20} = 0.623$ |                  |
| $a_{21} = 0.640$ |                  |

〈표 4〉 해지자수의 추정값과 실제값

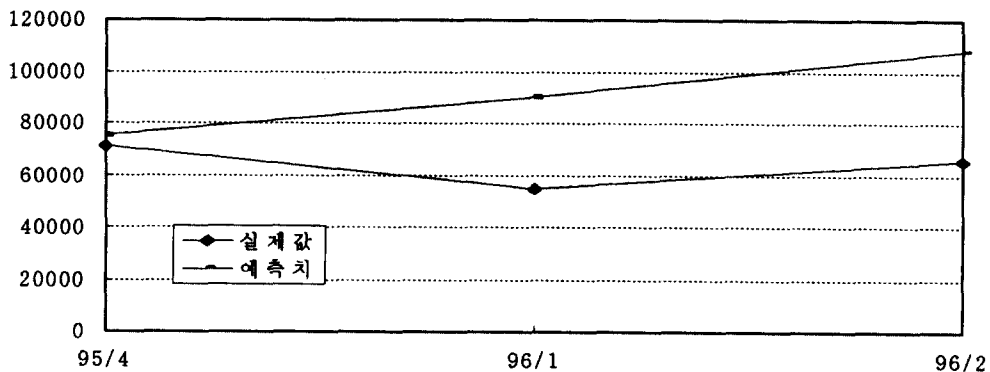
분 기	해지자수의 추정값	해지자수의 실제값
94/1	15177	16385
94/2	19553	19981
94/3	25490	27784
94/4	32276	35302
95/1	39452	38503
95/2	49288	46371
95/3	61495	61995

것으로 나타났다. 즉, 1-11분기까지 안정적인 형태를 보이다가 레벨상승이 발생하여 12-15분기까지 안정적이고 다시 그 이후에 레벨상승이 일어났다. 이는 초기 11분기 가입자들은 서비스 선호도가 높은 반면 12-15분기에 가입한 가입자들은 서비스 선호도가 상대적으로 낮고 그 이후에 가입한 사람들은 더욱 낮음을 알 수 있다. 이상의 결과를 이용하여  $a_t$ 에 대한 예측치를 구할 수 있으며 아날로그 셀룰라폰 신규가입자수의 예측치가 있다면 바로 해지자수의 예측치를 식(6)에 의해 구할 수 있다. 7분기동안의 fitting결과가 〈표 4〉와 〈그림 3〉에 나타나 있고, 1995년 4분기부터 1996년 2분기까지의 예측치와 실측치가 〈표 5〉와 〈그림 4〉에 나타나 있다. 이때 95년 4분기부터 96년 2분기까지의 신규가입자 예측치는 확산모형을 통해 예측

위와같이 구해진  $a_t$ 를 plotting하면 〈그림 2〉와 같다.  $a_t$ 를 분석한 결과 두번의 structural change가 발생한



〈그림 3〉 해지자수의 추정값과 실제값의 비교



〈그림 4〉 해지자수의 예측값과 실제값

〈표 5〉 해지자의 예측값과 실제값

분기	예측값	실제값
95/4	75266	71099
96/1	90127	54870
96/2	107814	65588

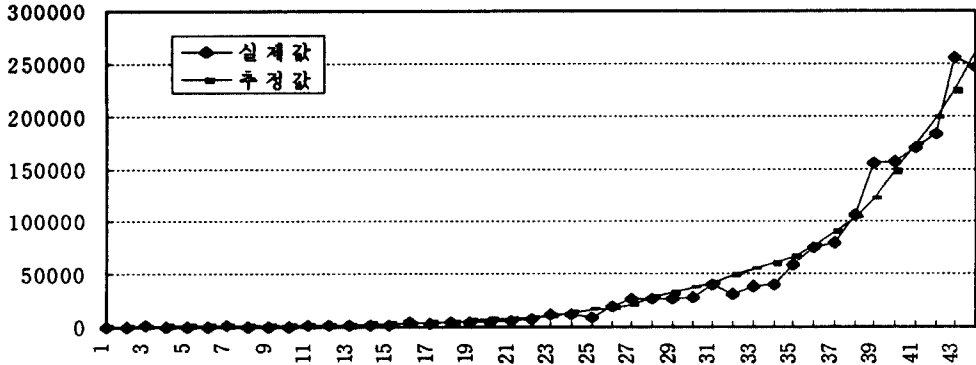
한 값(〈그림 5〉)을 사용하였다. 1996년도 이후로 디지털 셀룰라폰 서비스가 실시됨에 따라 셀룰라폰의 범주를 아날로그 셀룰라폰과 디지털 셀룰라폰을 묶어서 고려하였다.

예측결과(〈표 5〉)를 보면 1995년 4분기의 예측치는 비교적 잘 맞고 있으나 1996년의 예측치는 예측오차

가 크게 나타나고 있다. 이는 1996년부터 디지털 셀룰라폰 서비스가 새로 시작되고 제2사업자와의 경쟁이 도입됨으로서 소비자들의 행태에 영향을 미친 것으로 해석할 수 있다. 따라서 1996년 이후의 해지특성에 나타난 변화를 다시 모형화하는 작업이 필요할 것이다.

## 5. Conclusion

1995년까지 이동통신서비스는 아날로그 셀룰라폰과 무선호출기가 주가 되어왔다. 하지만 1996년 이후로 디지털 셀룰라폰, 발신전용전화기 등의 신규서비스가 실시됨에 따라 기존 이동통신서비스 가입자들중 다수



〈그림 5〉 확산모형에 의한 신규가입자수의 추정값과 실제값

가 그 서비스를 해지하고 새로운 경쟁 서비스로 전환하게 되었다. 이 경우 특정 이동통신서비스의 기존 가입자 중에서 발생하는 해지자수를 정확히 예측하는 것은 해당 이동통신서비스 가입자수를 합리적이고 정확히 예측하기 위해 매우 중요한 문제로 대두되었다.

본 연구에서는 이동통신서비스의 해지자수를 예측하기 위한 합리적이고 이해하기 쉬운 모형을 제시하였다. 여기서 제시된 모형은 설문조사를 통한 방대한 자료가 필요없으며 직관적으로도 이해하기 쉽기 때문에 적은 비용으로 이동통신서비스 가입자수를 예측하는데 도움이 될 것이다.

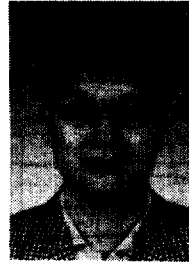
【참고문헌】

[1] Box, George E. P., Jenkins Gwilym M. and Reinsel Gregory C., Time Series Analysis, Prentice Hall, 1994

[2] Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables," *Journal of Business*, 53 (July, part 2), 51-67

[3] Mahajan, Vijay and Eitan Muller (1979), "Innovation Diffusion and New Product Growth Models in Marketing," *Journal of Marketing*, 43 (Fall), 55-68

[4] Mahajan, Vijay, Eitan Muller and Frank M. Bass (1990), "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, 54 (January), 1-26



전덕빈  
 1981년 서울대학교 산업공학과 학사  
 1985년 Univ. of California, Berkeley 산업공학 박사  
 1985년-1989년 경희대학교 조교수  
 1991년-1992년 Univ. of California, Berkeley부설 교통문제 연구소 객원연구원  
 현재 한국과학기술원 테크노경영대학원 부교수  
 관심분야: 경영 및 경제예측, 통신서비스 수요예측 및 진화전략



임철주  
 1995년 한국과학기술원 경영과학 학사  
 현재 한국과학기술원 테크노경영대학원 석사과정  
 관심분야: 통신서비스 수요예측, 시계열분석

주영진  
 1989년 연세대학교 경제학 학사  
 1990년 한국과학기술원 경영과학 석사  
 1995년 한국과학기술원 경영과학 박사  
 1995년-1996년 한국과학기술원 산업경영연구소 Post Doc.  
 현재 전자통신연구소 선임연구원  
 관심분야: 통신예측, 상태·공간 분석

**박운서**

1993년 한국과학기술원 수학 학사  
1995년 한국과학기술원 경영과학 석사  
현 재 한국과학기술원 테크노경영대  
학원 박사과정  
관심분야: 신상품 수요예측, 시계열분  
석, 통신예측

**박명환**

1983년-1986년 대우조선공업(주)  
1993년 한국과학기술원 산업공학 박사  
현 재 한성대학교 산업공학과 부교수  
관심분야: 통신서비스 진화 전략 및  
시장분석, O.R.

**박영선**

1988년 한국과학기술원 박사  
현 재 서경대학교 경영학과  
조교수  
관심분야: 생산관리 및 O.R., 통신경영