지능정보연구 제15권 제4호 2009년 12월(pp.37~51)

SOM을 이용한 제품수명주기 기반 서비스 수요예측

장남식
서울시립대학교 경상대학 경영학부 교수
(nchang@uos.ac.kr)

서비스 인적자원 운용의 효율성 제고와 부품 또는 시설 할당의 적정성 향상을 위해 서비스센터를 통해 접수되는 서비스 요청 건수를 보다 정확하게 예측하고자 하는 필요성이 제조업을 중심으로 증가하고 있다. 본 연구에서는 제품의 특성을 반영하여 제품수명주기 별로 제품들을 군집화하고 군집 별로 적절한 예측모형을 구축한 후 예측 값을 통합하는 개별예측방식을 LCD 모니터 제조사의 사례를 통해 제시한다. 또한 예측 결과를 총량방식 및 기존에 기업이 사용하고 있는 방식과 비교· 평가하여 우수성을 증명함으로써 제품이나 산업의 특성을 반영한 맞춤형 수요예측 기법 도입의 필요성을 부각하고, 그에 따른 이론적, 실무적 가이드라인을 제공한다.

논문접수일 : 2009년 09월 22일 논문수정일 : 2009년 10월 04일 게재확정일 : 2009년 10월 15일 교신저자 : 장남식

1. 서론

고객중심의 경영이 기업전반에 중요한 이슈로 부각되면서 많은 기업들이 고객의 불만과 요구사항을 신속하고 친절하게 접수하고 처리하여 고객만족도를 높이고자 하는 고객 지향적 경영체제로의 변환에 많은 노력과 투자를 하고 있다. 고객만족도 제고를 위해서뿐만 아니라 고객들로부터 접수되는 불만과 요구, 특히 제조산업에 있어 서비스 요청 건수와 내용 등은 자원배분의 효율성 증진과 서비스 품질 향상을 위해서도 매우 중요한 정보임에 틀림없다. 따라서 기업들은 향후 발생하는 서비스에 대해 정확한 예측하고 그에 따라 적절히 대처하는 능력을 키우는 것이 치열한 경쟁 체제에서 생존하기 위한 필수 요건임을 인지하고 있다. 그러나 실제로 이들 기업들이 활용하고 있는 서비스 수요예측 방법은 대부분 과거의 경험이나 관념 등에 의존하거나, 자신들이 생산하고 판매하는 제품의 특성을 전혀 고려하지 않은 채 단순한 기초통계 분석에 그치는 수준에 불과한 것이 국내의 현실이다.

본 연구의 사례 대상인 A사는 지난 몇 년간 200개 이상의 LCD 모니터 모델을 생산하는 기업으로 연간 제품 판매량이 꾸준히 증가하는 호황을 누려왔다. 이와 더불어 서비스 요청 건수 또한 지속적으로 증가하는 추세이며, 서비스 수요예측의 정확성 제고의 필요성을 어느 때보다도 절실히 실감하고 있다. 그러나 일부 모델은 이미 단종되었고, 신

* 이 논문은 2008년도 서울시립대학교 학술연구조성비 지원에 의하여 연구되었음.

규 모델은 계속 출시되고 있으며, 모델의 종류 또한 매우 다양하다는 제품의 특성으로 인해 정교한 예측 방식의 도입이 요구됨에도 불구하고 모델에 관계없이 전체 LCD 모니터를 대상으로 과거의 서비스 건수만을 이용하여 수요를 예측하는 변형된 이동평균 방식을 사용하고 있을 뿐이다.

이에 본 연구에서는 '시장진입', '성장', '성숙', '쇠퇴', '철수' 등으로 설명되는 제품수명주기(PLC : Product Life Cycle)에 기반하여 제품들을 군집화하고, 군집 별로 적합한 예측 기법을 적용한 후 예측 값을 통합하는 개별예측(disaggregate forecasting) 방식을 제안한다. 또한 결과를 총량예측(aggregate forecasting) 방식 및 기존에 A사가 사용하는 방식과 비교·평가 함으로써 제품이나 산업의 특성을 반영한 맞춤형 수요예측 기법 도입의 필요성을 부각하고, 그에 따른 이론적, 실무적 가이드라인을 제공하고자 한다. 특히 PLC별 제품 세분화를 위해 인공지능 기반의 SOM(Self-Organizing Map)의 활용을 제안한다.

연구의 구성은 다음과 같다. 제 2장에서는 SOM을 이용한 군집화, 수요예측 방법론, 제품수명주기 등에 관한 관련 연구들을 조사하고, 제 3장에서는 SOM의 이론적 배경을 살펴본다. 제 4장에서는 실증분석을 위해 수집한 데이터의 특성과 평가기준을 정리하고, 제 5장에서는 기초 통계분석, SOM을 이용한 군집화, PLC별 예측모형 구축, 모형평가 등의 4단계로 구성한 실증분석의 전 과정을 다룬다. 제 6장에서는 연구의 결과를 정리하고 시사점을 논한다.

2. 관련 연구

SOM은 군집화를 위한 데이터마이닝 기법 중에서 가장 보편적인 것으로서 고객 및 시장 분석, 상품 분석 등에 널리 사용되고 있는 것으로 알려져 있다(Shaw et al., 2001). Curry et al.(2003)는 소비자 행동분석에서 군집화 분석을 통해 소비자 행동에 영향을 미치는 주요 핵심요인을 도출하는 데에 SOM을 사용하였는데 다른 군집분석 기법 등에 비해 군집 해석력이 가장 뛰어나다는 것을 강조하였다. 이와 더불어 SOM을 통해 군집의 수를 결정한 후, K-means 알고리즘을 이용해 추가 군집화를 수행하는 것이 바람직하다는 의견이 제시되기도 하였다(Kuo et al., 2002). Hung과 Tsai(2008)는 군집 결과의 시각화에 보다 효과적인 계층적 이차원 시각화 지도(HSOM : Hierarchical Self-Organizing Map)를 제안하였고 실증분석을 통해 검증하였다. 이외에도 B2B 마켓플레이스, 자동차회사, 주식시장, 인터넷 쇼핑, 여행사 등 다양한 산업분야에서 CRM을 위한 고객 세분화에 SOM을 적용한 연구가 꾸준히 진행되어 왔다(김신영 & 김수영, 2003; Velido et al., 1999; Bloom, 2004; Shin and Sohn, 2004; Lee and Park, 2005; Bang, et al., 2007).

수요예측 방법론과 관련해서는 상향식(bottom-up)과 하향식(top-down) 접근법에 대한 연구가 주목할 만한데 대체적으로 상향식 접근법의 우수성을 입증하는 연구 결과들이 주를 이룬 반면(Dunne et al., 1971; Gross and Sohl, 1990; Dangerfield and Morris, 1992), 어느 하나가 특별히 우수하다고 입증할 수 없다라는 사실을 설문조사를 통해 제시한 연구도 발견할 수 있었다(Wealtherby, 1985). Schwarzkopf et al.(1988)은 수요예측 시 두 가지 접근법을 조합해서 사용하는 방안을 제시하였으며, Kahn(1998)은 이를 위한 구체적인 가이드라인을 발표하였다. Lo et al.(2008)은 5개의 단계로 구성된 계층적 수요예측(Hierarchical Forecasting) 방법론을 제안하였고 LCD 모니터 시장 데이터를 이용해

검증하였다.

한편, 제품수명주기에 따른 수요나 가격 예측에 관련된 연구들도 활발히 진행되어 왔다. 이 중 예측모형 구축을 위한 데이터의 양이 충분하지 않을 때, 특히 과거 이력 정보나 수요예측에 영향을 미칠 수 있는 변수들이 상대적으로 부족한 짧은 수명주기를 지닌 제품들을 위한 예측 기법으로 SVM (Support Vector Machine)을 제안한 흥미로운 연구가 진행되었다(Xu and Zhang, 2008). Huang과 Tzeng(2008)은 하이테크 제품의 생산 계획 및 전략 수립에 있어서 제품수명주기 예측의 중요성을 부각시키며 예측방법으로 퍼지 회귀분석(Fuzzy piecewise regression)을 16메가 DRAM 수명주기 예측 사례를 통해 설명하였다.

3. SOM 알고리즘

SOM은 다차원 개체공간에 있는 개체벡터를 저차원 그리드(grid) 공간에 표현하기 위해 코호넨에 의해서 개발된 자율 학습 알고리즘이다. SOM의 학습은 개체벡터에 가장 가까운 중량벡터를 찾아 개체벡터 방향으로 이동시키는 과정을 반복하는데, 이 과정에서 가장 가까운 중량벡터의 주변 중량벡터도 함께 개체벡터 방향으로 이동시킨다. 여기서 가장 가까운 중량벡터를 모든 중량벡터와의 경쟁에서 승리하였다는 의미에서 '승자(winner)'라고 부른다. 학습의 반복은 모든 중량벡터 값의 변화가 거의 소멸하거나 미리 지정된 최대 한계에 도달할 때까지 계속 진행되며, 이러한 학습과정을 통해 얻은 최종적인 중량벡터는 개체공간상에서 개체들이 지니는 위상을 반영하게 된다. 본 논문에서 사용할 코호넨의 SOM 알고리즘을 정리하면 다음과 같다(Kohonen, 1995).

단계 0 : 분석 자료가 n개의 p차원 개체벡터 $x_1, \cdots, x_n$ 으로 구성되어 있다고 하자. 저차원 그리드 상의 m개의 노드는 p차원 개체공간에 있는 중량벡터 $w_1, \cdots, w_m \in R^p$와 연결되어 있다. 노드와 중량벡터를 구분하기 위하여 노드를 $z_1, \cdots, z_m \in R^q$로 표기하기로 하자. 여기서 $q(\leq p)$는 그리드 공간의 차원수이다.

단계 1 : 시점 t의 첫 값을 1로 하고, m개의 p차원 중량벡터 $w_1, \cdots, w_m$에 초기값을 부여한다. 또한 초기 학습률 α_0과 최종 학습률 α_1을 지정한다($0 < \alpha_1 < \alpha_0 < 1$).

단계 2 : 입력개체 i(i = 1, …, n)가 제시되면 m개의 중량벡터 $w_1, \cdots, w_m$ 중 가장 가까운(유클리드 거리) 것을 찾아내고 그 중량벡터가 속한 노드를 K(i)이라 한다.

$$\| x_i - w_{k(i)} \| \leq \| x_i - x_j \|,$$
여기서 $j = 1, \cdots, m$

단계 3 : 승자노드 K(i)와 그 주변노드의 중량벡터들이 다음과 같이 업데이트한다.

$$w_j \leftarrow w_j + \alpha_t h_t(K(i), j)(x_i - x_j)$$

학습률 α_t는 초기값 α_0에서 최종값 α_1까지 t에 따라 감소하며, 주변함수 $h_t(k, j)$는 국소가중치(local weight)로서 흔히 다음과 같은 패턴의 함수가 적용된다.

$$\begin{cases} \exp[-|z_j - z_k|^2/(2\sigma_t^2)], & \text{if } d(z_j, z_k) \leq \sigma_t \\ 0, & \text{otherwise} \end{cases}$$

여기서, z_j는 노드 j의 그리드 위치점이고 $d(z_k, z$

$_j$) = || z_j − z_k ||$_\infty$이다. 주변 거리의 크기 α_t는 초기값 α_0에서 최종값 α_1까지 t가 감소하는 계단함수이다.

단계 4 : 한 개체가 단계 2와 3에 따라 처리되면 시점 t를 1만큼 증가시키고, 마지막 개체가 처리되면 첫 개체로 되돌아간다. 단, 모든 중량벡터 값의 변화가 거의 소멸하거나 시점 t가 미리 지정된 최대 한계에 도달하게 되면 진행을 멈추고 각 개체의 승자노드 K(1), …, K(n) 및 중량벡터 w_1, …, w_m을 출력한다.

본 연구에서는 자바언어를 사용하여 개발한 SOM Analysis Tool을 이용하여 총 200개의 모델을 대상으로 4×4 메트릭스 형태의 군집화를 실행하였다.

4. 데이터 수집 및 평가기준

4.1 데이터 수집 및 구성

LCD 모니터를 제조·판매하는 A사의 2005년 1월부터 2008년 12월까지의 월별 판매 DB와 서비스센터를 통해 접수되는 서비스 DB를 이용하여 월별 서비스 건수를 예측하는 모형을 구축하고 평가하였다. 데이터의 개요는 다음과 같다.

월별 판매 DB는 '모델명', '판매연월, '생산경과월', '판매건수' 등 4개의 속성과 4,420개의 레코드로 구성되어 있으며, 총 판매 건수는 8,490,339개이다. 월별 서비스 DB는 '모델명', '서비스 발생월', '판매경과월', '생산경과월', '서비스 건수' 등 5개의 속성과 80,735개의 레코드로 구성되어 있으며, 총 서비스 건수는 133,955개이다. 연도별 판매 및 서비스 통계는 <표 1>과 같으며, 이 중 2005년부터 2007년까지의 데이터를 예측모형 구축용으로, 2008년 데이터를 모형 평가용으로 사용하였다.

4.2 평가기준

모형 간의 예측력 비교·평가를 위해서는 오차의 범위 및 편차가 큰 점을 감안해 오차제곱평균(MSE : Mean Square Error) 대신 평균오차제곱근(RMSE : Root Mean Square Error)과 예측 오차의 비교가 용이하고 상대적으로 신뢰도가 높은 절대평균오차비율(MAPE : Mean Absolute Percentage Error)를 사용하였다.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{X_t - F_t}{X_t}\right)^2}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{X_t - F_t}{X_t}\right| \times 100$$

여기서, X = 실측 값, F = 예측 값, n = 실측 값의 기간 수

<표 1> 연도별 판매 및 서비스 통계

		2005	2006	2007	2008	총계
판매	레코드수	524	857	1,220	1,819	4,420
	건수	535,152	1,268,164	2,642,593	4,044,430	8,490,339
서비스	레코드수	2,539	13,244	29,053	35,899	80,735
	건수	4,530	24,800	48,922	55,703	133,955

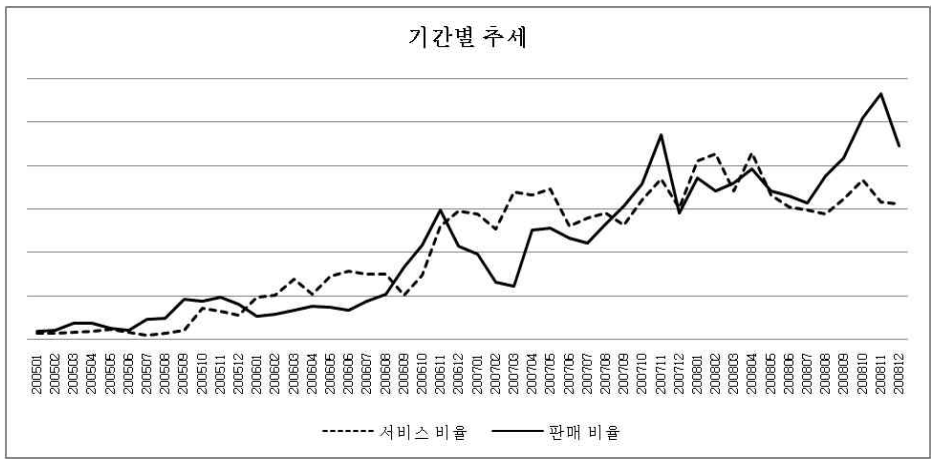
5. 실증분석

실증분석은 <그림 1>과 같이 4단계로 구성하였다.

<그림 1> 실증분석 프로세스

5.1 기초 통계분석

<그림 2>는 2005년 1월부터 기록한 월별 판매 및 서비스 건수 비율을 기간별 추세로 나타낸 도표로 판매 및 서비스 공히 월별 변동량이 크고 불규칙적이며, 지속적으로 증가하고 있음을 알 수 있다. 특히 서비스 추세(점선)는 판매 추세(실선)를 기준으로 우측으로 이동(shift)되고 있는 시간격차(time lag)가 존재하는데 판매가 서비스에 영향을



<그림 2> 기간별 서비스 추세

<표 2> 파생 변수

분 류	파생 변수	분 류	파생 변수
서비스 건수 관련 (8개)	• 직전 2개월 서비스 건수 평균 • 직전 3개월 서비스 건수 평균 • 직전 4개월 서비스 건수 평균 • 1개월 전 서비스 건수 • 2개월 전 서비스 건수 • 3개월 전 서비스 건수 • 4개월 전 서비스 건수 • 5개월 전 서비스 건수	판매 건수 관련 (15개)	• 직전 1~2개월 판매 건수 평균 • 직전 1~3개월 판매 건수 평균 • 직전 1~4개월 판매 건수 평균 • 직전 1~5개월 판매 건수 평균 • 직전 2~3개월 판매 건수 평균 • 직전 2~4개월 판매 건수 평균 • 직전 2~5개월 판매 건수 평균 • 직전 3~4개월 판매 건수 평균 • 직전 3~5개월 판매 건수 평균 • 직전 4~5개월 판매 건수 평균 • 1개월 전 판매 건수 • 2개월 전 판매 건수 • 3개월 전 판매 건수 • 4개월 전 판매 건수 • 5개월 전 판매 건수

미치는 데에는 몇 개월이 소요된다는 것을 의미한다. 이를 근거로 서비스 건수 예측을 위해 최대 5개월 직전까지의 서비스 건수 및 판매 건수가 영향을 미칠 수 있다고 판단하여 총 18개의 파생변수들을 <표 2>와 같이 생성하였다. 또한 서비스 건수가 계절요인에 영향을 받는지의 여부를 확인하기 위해 <그림 3>과 같이 월별 서비스 건수 분포를 검토한 결과 계절요인이 거의 없다고 판단하였다.

5.2 SOM을 이용한 군집화

입력 데이터는 2005년 1월부터 2008년 12월까지의 48개월 간의 서비스 건수 데이터를 이용하여 서비스가 한 건이라도 발생할 경우에는 '1', 서비스가 발생하지 않은 경우에는 '0'으로 변환하여 구성하였다. 이는 서비스 발생 유무를 주요 속성으로 하여 서비스 발생시점과 종료시점간의 기간에 대한 유사도를 측정함으로써 각 제품수명주기 별로 모형을 구축하기 위함이다.

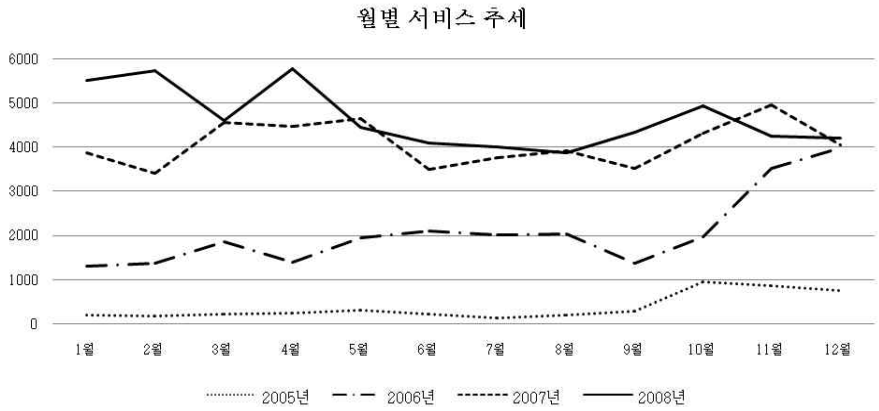
<그림 4>는 군집화 결과로 각 셀간의 거리 또는 셀의 명암 차이가 클수록 유사도가 적음을 의미하며, 괄호 안의 수는 해당 셀에 속하는 모델의 개수이다. 각 셀의 모델들을 그룹화하여 프로파일링(profiling)한 결과 'A : 진입→성숙', 'B : 성숙→쇠퇴', 'C : 쇠퇴→철수', 'D : 철수' 등의 4개의 모델군으로 분류되었다.

<표 3> 모델군 개요

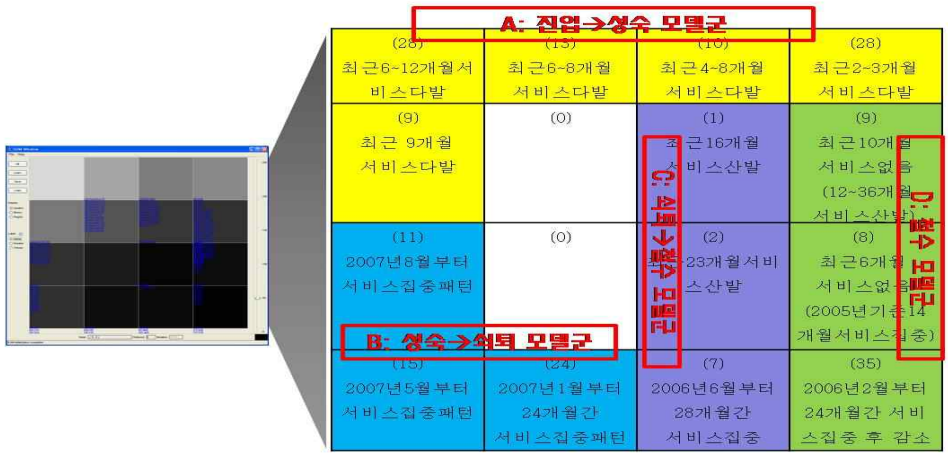
모델군	PLC 단계 (2007년 12월 기준)	모델 수	2008년 실제 총 서비스 건수 비율
A	진입→성숙	88개	21%
B	성숙→쇠퇴	50개	76%
C	쇠퇴→철수	10개	2%
D	철수	52개	1%

<표 3>은 4개 모델군 대한 통계이며, <그림 5>는 각 모델군별・기간별 서비스(점선) 및 판매(실선) 건수 비율을 보여주고 있다. 각 모델군에 대한 특징은 다음과 같다.

- 모델군 A : 2007년 12월 기준에서 진입 → 성숙 단계로 총 88개의 모델들을 포함하며, 2008년에 발생한 실제 총 서비스 건수의 비중은 21%이다. 2008년 신규 진입 제품이기 때문에 2007년 12월 현재 2008년 서비스 건수 예측을



<그림 3> 월별 서비스 추세



<그림 4> SOM을 이용한 군집화 결과

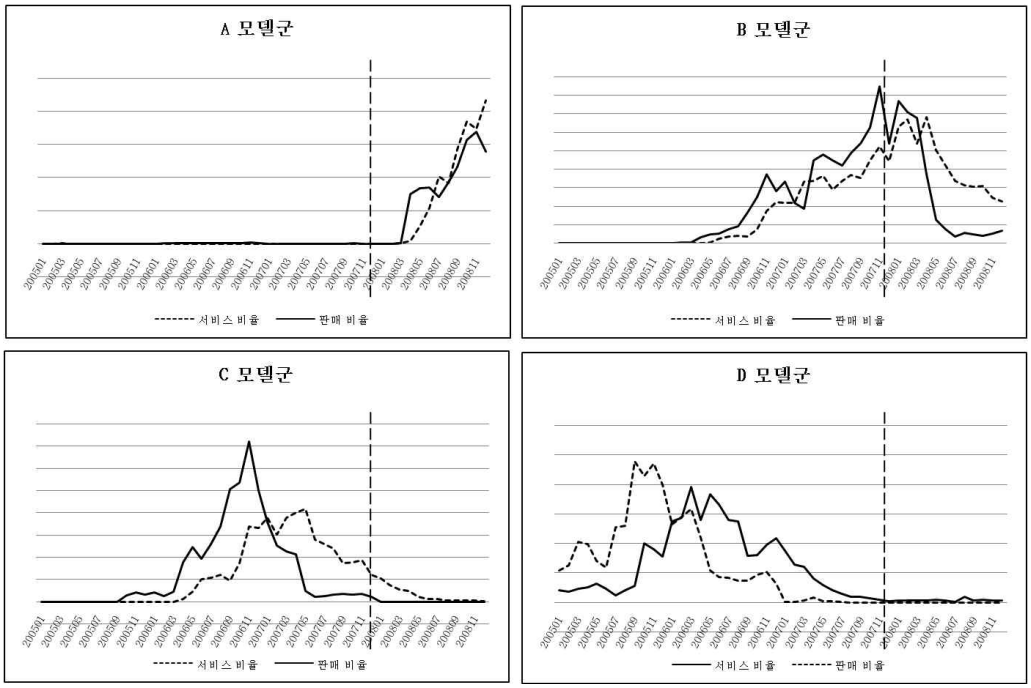
위한 모형 구축용 데이터가 존재하지 않는다. 따라서 모형 구축용 데이터 선정에 각별한 주의를 요하는 특징을 지닌 모델군이다.

- 모델군 B : 2007년 12월 기준에서 성숙→쇠퇴 단계로 총 50개의 모델들을 포함하며, 2008년 실제 총 서비스 건수의 비중은 76%이다. 2007년 12월 현재 서비스 건수가 정점을 지나고 있는 제품들로 이듬해 하반기부터 서비스 건수가 현저히 줄어드는 특징을 가진다.
- 모델군 C : 2007년 12월 기준에서 쇠퇴→철수 단계로 총 10개의 모델들을 포함하며, 2008년 실제 총 서비스 건수의 비중은 2%에 불과하다. 이미 철수가 진행 중인 제품들이기 때문에 이듬해 초반에 서비스가 존재하다가 사라지는 특징이 있다.
- 모델군 D : 2007년 12월 기준에서 철수가 완료된 단계로 총 52개의 모델들을 포함하며, 2008년 실제 총 서비스 건수의 비중은 1% 미만이다. 철수가 완료된 점을 고려하여 2008년 서비스 수요예측에서 제외해도 무방하다는 특징을 지닌다.

5.3 PLC별 예측모형 구축

서비스의 수명주기가 유사한 패턴을 따른다는 가정아래 현재 신규 진입 모델군의 경우 앞으로 서비스 건수의 지속적인 증가가 예상되기 때문에 과거에 이미 시장에 진입하여 성숙 단계에 있는 모델군 데이터를, 현재 성숙 단계 모델군의 경우 앞으로 서비스 건수의 감소 또는 소멸이 예상되기 때문에 과거에 성숙 단계를 거쳐 쇠퇴 또는 철수 단계에 있는 모델군 데이터를 예측모형 구축용으로 사용하였다. 모델군 별로 선정된 모델링 기법 및 결과는 다음과 같다.

- 모델군 A : 현재 성숙 단계에 있는 모델군 B가 시장에 진입했던 2006년 1월부터 2007년 12월까지의 데이터를 이용하여 모형을 구축하였고, A 모델에 적용하여 예측을 실시하였다. 단계별 변수선정(stepwise) 방식을 사용해 아래와 같은 원점을 통과하는 다중회귀분석 모형을 도출하였는데 '1개월전 서비스 건수'와 '4개월전 판매 건수'가 유의하게 나타났다.



<그림 5> 모델군별 · 기간별 서비스 및 판매 건수 비율

$Y_t = 0.751 Y_{t-1} + 0/006 X_{t-4}$ $(R_a^2 = 0.970)$

여기서, Y_t : t월의 서비스 건수

Y_{t-1} : t−1월의 서비스 건수

X_{t-4} : t−4월의 판매 건수

- 모델군 B : 현재 철수 단계에 있는 모델군 D가 성숙 → 쇠퇴 단계였던 2005년 8월부터 2007년 6월까지의 데이터를 이용하여 모형을 구축하였고, B 모델에 적용하여 예측을 실시하였다. 유사한 단계인 모델군 C의 데이터를 모형 구축용으로 이용할 수도 있었으나 모델군 C에 포함된 모델의 수가 적어 모형을 일반화하기에 어려움이 따랐다. 다중회귀분석 결과 원점을 통과하며 '1개월 전 서비스 건수'만이 유의하게 나타났다.

$Y_t = 0.959 Y_{t-1}$ $(R_a^2 = 0.916)$

여기서, Y_t : t월의 서비스 건수

Y_{t-1} : t−1월의 서비스 건수

- 모델군 C : 서비스 건수 분포를 분석한 결과 지속적으로 감소하는 패턴을 보이며 완전 철수가 예상되는 모델군으로, 자체 모델군이 쇠퇴기에 진입하는 2007년 6월부터 12월까지의 서비스 건수 데이터를 이용하여 시계열분석을 시행하였다.

$Y_t = 1005 - 80.64 X_t$

여기서, X_t : 시계열 변수로 t = 1은 2007년 6월임

- 모델군 D : 2007년 하반기에 이미 철수가 완료되었기 때문에 2008년에 서비스 수요가 없다고 판단하여 모형 구축에서 제외하였다.

5.4 모형평가

PLC 방식은 전체 모델들을 제품수명주기에 따라 4개의 모델군으로 분류하고, 각 모델군 별로 예측모형을 구축한 후, 예측 값을 통합하는 개별예측 방식이다. 이 방식을 전체 모델들을 대상으로 단일 모형을 구축하는 총량예측 방식 및 기존에 A사에서 사용하는 예측 방식과 비교·평가하였다.

5.4.1 A사의 기존 예측방식

전월까지의 서비스 실적을 이용하여 해당 월의 서비스 건수를 예측하는 방법이다. 즉, 2008년 1년간의 월별 서비스 예측 값을 1월부터 순차적으로 계산한 후에 월별 예측 값을 통합하여 1년 간의 예측 값을 산출하는 방식으로, 직전 3년 간의 서비스 건수에 대한 계절요인(seasonal factor)과 직전 2년 대비 직전 년의 서비스 증가율이 주요 변수가 된다. 기존 방식은 아래와 같다.

해당월 서비스 예측 값 = 직전 12개월 서비스 건수 합계 × 최근 3개년 월별 평균 서비스 점유율×서비스 건수 증가율 × 품질 개선 보정 지수

- 직전 12개월 서비스 건수 합계 : 최근 1년 간의 서비스 건수의 총합
- 최근 3개년 월별 평균 서비스 점유율 : 최근 3개년의 해당 월 서비스 건수 합을 최근 36개월 서비스 건수의 총합으로 나눈 값
- 서비스 건수 증가율 : 해당월로부터 13개월 전에서 24개월 전까지의 서비스 건수의 총합 대비 최근 1년 간의 서비스 건수의 증가율
- 품질 개선 보정 지수 : 고정값 0.95

위의 방식은 계산이 용이하고 시계열 요인과 계절적 영향 요인을 함께 반영하는 장점이 있지만, 과거 서비스 건수만을 사용하여 예측하기 때문에 판매 건수와 서비스 건수간의 상관관계가 존재할 경우 판매 정보를 간과하는 단점이 있다.

5.4.2 총량예측 방식

이 방식은 개별예측 방식과는 달리 모델들의 특성을 고려하지 않고 제품의 최상위 계층인 하나의 제품코드(LCD) 수준에서 단일예측 모형을 구축하는 것으로, 도출된 회귀모형은 아래와 같다.

$$Y_t = 676.17 + 0.009X_{t(1-5)} + 0.415\ Y_{t(1-3)} \quad (R_a^2 = 0.77)$$

여기서 $X_{t(1-5)}$: 직전 1~5개월 판매 건수 평균

$Y_{t(1-3)}$: 직전 3개월 서비스 건수 평균

5.4.3 예측모형 비교

<표 4>는 PLC, 총량예측, 기존예측 방식의 예측 정확도를 모형평가 기준인 평균오차제곱근(RMSE)과 절대평균오차비율(MAPE) 관점에서 비교한 것으로, PLC 방식의 RMSE와 MAPE가 각각 836과 14%로 다른 두 개의 모형에 비해 가장 우수하게 나타났다. 영향력있는 변수들을 파악하기 위해 3개의 모형들에 포함되는 변수들을 정리한 결과, 과거의 서비스 건수(1개월 전 서비스 건수, 직전 1~4개월 서비스 평균)와 판매량(4개월전 판매량, 직전 1~5개월 판매량 평균)이었는데 이것은 판매량의 증감에 따라 서비스 건수의 증감도 유사하게 변동한다는 것을 의미한다. 또한 제품을 수명

<표 4> 모형 평가표

예측 월	실측 값	PLC 방식	총량예측 방식	기존예측 방식
200801	5,516	3,830	4,387	6,281
200802	5,726	5,200	4,387	5,375
200803	4,592	5,410	4,383	6,794
200804	5,760	4,310	4,398	5,790
200805	4,444	5,366	4,328	5,918
200806	4,085	4,089	4,396	4,599
200807	3,992	3,723	4,422	4,479
200808	3,875	4,389	4,426	4,409
200809	4,329	4,412	4,402	3,486
200810	4,939	4,825	4,383	4,514
200811	4,255	5,237	4,402	5,399
200812	4,190	4,765	4,519	4,847
RMSE		836	1,025	964
MAPE		14%	21%	18%

주기 별로 군집화한 후 군집 별로 예측모형을 구축하고 예측 값을 통합하는 상향식 접근(bottom-up) 기반의 개별예측 방식이 제품의 특성 및 추세를 보다 적절하게 반영하였다고 할 수 있다.

6. 결론

본 연구에서는 '시장진입', '성장', '성숙', '쇠퇴', '철수' 등으로 설명되는 제품수명주기에 기반하여 제품들을 SOM을 이용해 군집화하고, 군집 별로 각각 적합한 예측 기법을 적용한 후 예측 값을 통합하는 개별예측 방식을 제안하고 총량예측 방식 및 기존에 A사가 사용하는 방식과 비교·평가하였다. 모형평가 기준인 평균오차 제곱근(RMSE)과 절대평균오차비율(MAPE) 관점에서 PLC 방식이 다른 두 개의 모형에 비해 가장 우수하게 나타났으며, 이를 통해 제품의 특성을 고려한 수요예측 기법의 도입을 강조하였다. 그러나 이것은 단지 개선된 수요예측 기법을 필요로 하는 기업들에게 단초를 제공하는 수준일 뿐이며, 보다 정교한 예측을 위해서는 다음과 같은 추가적인 많은 노력이 필요하다.

첫째, 아무리 적합한 예측 기법을 적용하여 수요를 예측한다고 하더라도 원천 데이터의 정확도가 낮을 경우 예측 결과의 신뢰도도 낮아지게 되며 궁극적으로 현업에서의 활용도 또한 떨어질 수밖에 없다. 본 연구에서도 데이터 정제 시 일부 데이터의 오류 값들과 일반 상식 선에서 이해하기 어려운 특이 값들이 다수 발견되었다. 따라서 입력단계에서부터 서비스 및 판매 데이터의 신뢰도 관리가 무엇보다도 절실히 요구된다.

둘째, 기존에는 서비스 수요 예측을 위해 과거 서비스 건수와 판매 물량만을 독립변수로 사용하여 왔다. 그러나 예측력의 보편성과 정확성을 제고하기 위해서는 지역적 특성이나 부품의 이력 등과 같이 서비스 수요에 영향을 미칠 수 있는 여러 변수들을 발굴하고 군집 별로 다양한 예측 기법을 적

용하여 개선하는 노력이 필요하다. 이와 더불어 단순히 서비스 건수 예측뿐만 아니라 서비스 비용, 비용률, 품질지표 등에 대한 추가 예측 모형의 개발에 대한 연구가 따라야 하겠다.

참고문헌

김신영, 김수영, "군집화 기법을 이용한 B2B Marketplace 상의 최적 파트너 검색 시스템", *한국경영과학회/대한산업공학회 춘계공동학술대회*(2003), 575~579.

Bang, J., L. Hamel, and B. Loerger, "Rethinking of self-organizing maps for market segmentation in customer relationship management", *한국지능정보학회논문지*, 13권 4호(2007), 17~34.

Bloom, J. Z., "Tourist market segmentation with linear and nonlinear techniques", *Tourism Management*, Vol.25(2004), 723~733.

Curry, B., F. Davis, M. Evans, L. Moutinho, and P. Phillips, "The Kohonen self-organizing map as an alternative to cluster analysis : An application to direct marketing", *International Journal of Market Research*, Vol.45(2003). 191~211.

Dangerfield, B. J. and J. S. Morris, "Top-down or bottom-up : aggregate versus disaggregate extrapolations", *International Journal of Forecasting,* Vol.8(1992), 233~241.

Dunn, D. M., W. H. Williams, and W. A. Spiney, "Analysis and prediction of telephone demand in local geographic areas", *Bell Journal of Economics and Management Science,* Vol.2(1971), 561~576.

Gross, C. W. and J. E. Sohl, "Disaggregation methods to expedite product line forecasting", *Journal of Forecasting*, Vol.9 (1990), 233~254.

Huang, C. Y. and G. H. Tzeng, "Multiple generation product life cycle predictions using a novel two-stage fuzzy piecewise regression analysis method", *Technological Forecasting and Social Change,* Vol.75(2008), 12~31.

Hung, C. and Tsai, C., "Market segmentation based on hierarchical self-organizing map for markets of multimedia on demand", *Expert Systems with Applications*, Vol.34(2008), 780~787.

Kohonen, T., *Self-Organizing Maps*, Springer, New York, 1995.

Kuo, R. J., L. M. Ho, and C. M. Hu, "Integration of self-organizing feature map and K-means algorithm for market segmentation", *Computers and Operational Research*, Vol.29(2002), 1475~1493.

Lee, J. H. and S. C. Park, "Intelligent profitable customers segmentation system based on business intelligence tools", *Expert Systems with Applications*, Vol.29(2005), 145~152.

Lo, S., F. Wang, and J. T. Lin, "Forecasting for the LCD monitor market", *Journal of Forecasting,* Vol.27(2008), 341~356.

Schwartzkopf, A. B., R. J. Tersine, and J. S. Morris, "Top-down vs. bottom-up forecasting strategies", *International Journal of Production Research,* Vol.26(1988), 1833~1843.

Shaw, M. J., C. Subramaniam, G. W. Tan, and M. E. Welge, "Knowledge management and data mining for marketing", *Decision Support Systems*, Vol.31(2001), 127~137.

Shin, H. W. and S. Y. Sohn, "Segmentation of stock trading customers according to potential value", *Expert Systems with Applications*, Vol.27(2004), 27~33.

Vellido, A., P. G. J. Lisboa, and K. Meehan, "Segmentation of the on-line shopping market using neural networks", *Expert System with Applications*, Vol.17(1999), 303~314.

Weatherby G., *Aggregation, disaggregation, and combination of forecasts.*, PhD dissertation, Georgia Institute of Technology (1985).

Xu, X., H. Zhang, "Forecasting Demand of Short Life Cycle Products by SVM", *15th International Conference on Management Science and Engineering*(2008), 352~356.

Abstract

Product Life Cycle Based Service Demand Forecasting Using Self-Organizing Map

Namsik Chang*

One of the critical issues in the management of manufacturing companies is the efficient process of planning and operating service resources such as human, parts, and facilities, and it begins with the accurate service demand forecasting. In this research, service and sales data from the LCD monitor manufacturer is considered for an empirical study on Product Life Cycle (PLC) based service demand forecasting. The proposed PLC forecasting approach consists of four steps : understanding the basic statistics of data, clustering models using a self-organizing map, developing respective forecasting models for each segment, comparing the accuracy performance. Empirical experiments show that the PLC approach outperformed the traditional approaches in terms of root mean square error and mean absolute percentage error.

Key Words : Service Demand Forecasting, Product Life Cycle, Self-Organizing Map

* Faculty of Business Administration College of Economics and Business University of Seoul

저 자 소 개

장남식

University of Missouri에서 경영학 석사, 그리고 University of Arizona에서 경영정보시스템을 전공하여 경영학 박사학위를 취득하였으며, LG-EDS시스템에서 IT 컨설턴트로 근무한 바 있다. 현재 서울시립대학교 경상대학 경영학부에 재직 중이다. 삼성, LG, 국민, 현대 카드 및 농협, 외환은행, 부산은행 등 주로 금융분야에서 CRM 프로젝트를 수행하였으며, Journal of MIS, Decision Support Systems, Informs Journal on Computing, 경영정보학연구, Information Systems Review 등에 논문을 게재하였다. 주요 관심분야는 데이터웨어하우스 시스템 설계 및 구축, 데이터마이닝을 이용한 각종 산업별 데이터 분석 및 활용, 그리고 이들과 CRM과의 효과적인 접목 방안 연구 등이다.