

eCRM(electronic Customer Relationship Management) 시스템에서의 데이터 분석

안정용¹⁾, 김석기²⁾, 한경수³⁾

요약

CRM(Customer Relationship Management)은 데이터를 수집하고, 적절한 분석 단계를 통하여 정보를 제공할 수 있는 데이터 통합(수집, 관리, 분석, 서비스) 시스템 구축을 통한 고객관계관리를 의미한다. 본 연구에서는 eCRM 시스템의 구조 및 기능 등에 대해 살펴보고, 간단한 예제를 통하여 웹 데이터의 유용한 활용 방법을 제시해 보고자 한다.

주요용어 : eCRM, Web Data Mining, 데이터 통합 시스템

1. 서론

정보통신 산업과 디지털 기술의 발전은 엄청난 양의 데이터를 양산해 내고 있으며, 이러한 데이터의 효과적 활용에 많은 관심이 모아지고 있다. 데이터 웨어하우징(Data Warehousing), 데이터 마이닝(Data Mining), OLAP(On-Line Analytical Process), 데이터베이스 마케팅(Database Marketing) 등은 모두 데이터를 효과적으로 활용하기 위한 기술들이라 할 수 있으며, 최근에는 eCRM에 관한 연구가 활발히 진행되고 있다.

CRM은 고객을 관리하기 위하여 기업이 보유하고 있는 다양한 채널을 통하여 데이터를 수집하고, 적절한 분석 단계를 통하여 정보를 제공할 수 있는 데이터 통합 시스템을 구축하는 것이다. 일반적으로 CRM은 온라인과 오프라인 상의 모든 프로세스를 포함하며, CRM의 개념을 웹(World Wide Web)이라는 새로운 환경에서 실행하는 것을 eCRM 또는 웹 데이터 마이닝(Web Data Mining)이라 한다.

eCRM은 웹 상에서 발생하는 모든 데이터를 분석 대상으로 삼으며, Madria 등(1999)에서 제시하는 바와 같이 서버 로그 데이터(server access log data), 사용자 등록 정보(user registration or profiles), 사용자 세션 또는 트랜잭션(user sessions or transactions), ERP 데이터(enterprise resource planing data) 등을 통합하여 분석한다. 특히, 로그 데이터의 이용에 관

1) 서남대학교 컴퓨터정보통신학부 조교수, 전북 남원시 광치동 720

2) 군산간호대학 간호정보공학센터 연구원, 전북 군산시 개정동 413

3) 전북대학교 수확통계정보과학부 교수, 전북 전주시 덕진구 덕진동 664-14

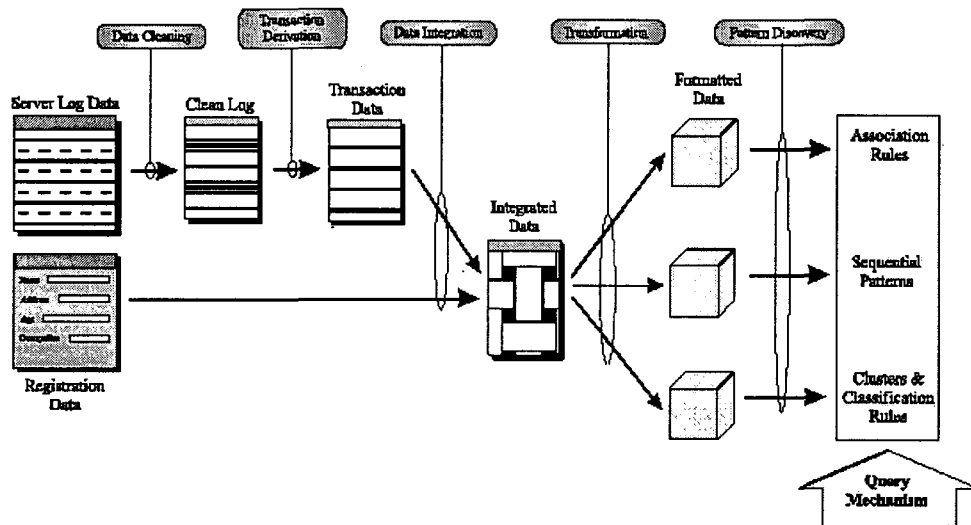
한 연구가 많이 진행되었으며, 몇 가지만 살펴보면 Borges와 Levene(1999), Buchner 등(1999), Pei 등(2000), Srivastava 등(2000)을 들 수 있다.

본 연구에서는 eCRM 시스템의 구조 및 기능 등에 대해 살펴보고, 간단한 예제를 통하여 통계적인 관점에서 웹 데이터의 유용한 활용 방법을 제시해 보고자 한다.

2. eCRM 시스템의 구조 및 기능

eCRM 시스템 구축의 목적은 일대일 서비스를 통한 효율적인 고객 관리에 있다. 고객은 각자 자신만이 원하는 서비스를 받기를 원하기 때문에 eCRM에서는 개인별 또는 같은 특성을 갖는 그룹별로 최적화된 서비스를 제공하는 것이 매우 중요하다. Spiliopoulou 등(1999), Yan 등(1996)에서 개별적인 특성에 따른 웹 서비스에 관한 연구가 이루어졌으며, 이러한 서비스를 제공하기 위하여 데이터에 대한 기본 통계량은 물론 상관(correlation)분석, 군집(clustering)분석, 그리고 판별(discriminant)분석 등이 이용된다. 군집분석은 eCRM 시스템에 가장 많이 활용되는 분석 방법이며, Hartigan과 Wong(1979)의 k-평균 알고리즘과 Kohonen(1984)의 코호넨 네트워크(Kohonen Network) 알고리즘이 주로 이용되고 있다. 그 외에도 Guha 등(1998), Zhang 등(1996)에서 여러 가지 군집분석 알고리즘에 관한 연구가 진행되었다.

2.1 eCRM 시스템의 구조



<그림 1> eCRM 시스템 구조

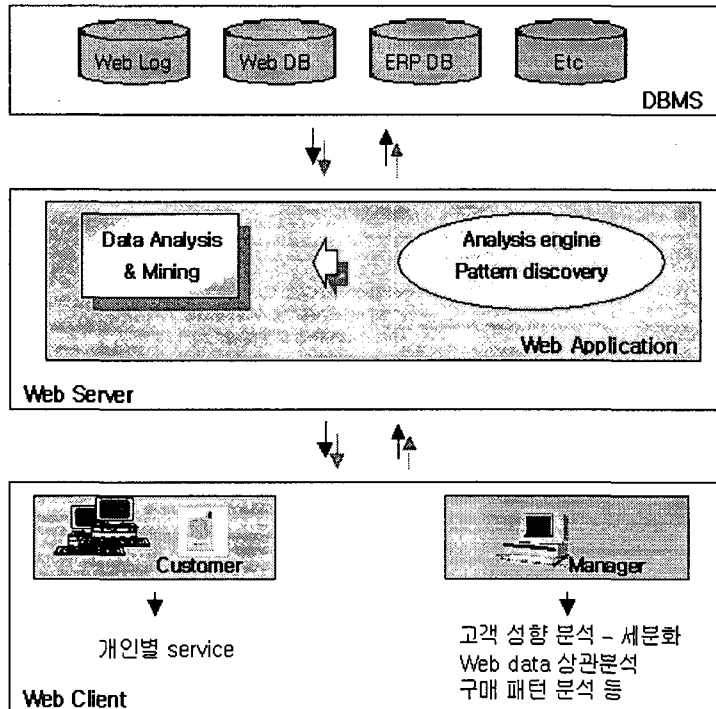
<그림 1>은 Mobasher 등(1996)에서 보여주고 있는 eCRM 시스템의 전체적인 구조이다. 데

이더 정제(cleaning)와 통합(integration), 그리고 분석과정을 통하여 사용자에게 정보를 제공하는 과정을 보여주고 있다.

2.2 eCRM 시스템의 기능

eCRM 시스템은 <그림 1>과 같은 과정을 통하여 일반 사용자(고객)에게는 개인별 웹 페이지 서비스와 상거래에 있어서 개인별 제안 등의 서비스를 제공할 수 있으며, 기업 관리자에게는 고객 성향 분석, 구매 패턴 분석, hit 추이 분석 등을 제공할 수 있는 기능을 수행한다.

3. 예제



<그림 2> 예제 시스템 구조

본 연구에서는 서버 로그 데이터, 데이터, 사용자 등록 정보, ERP 데이터 등을 통합하여 분석하는 예제를 제시한다. <그림 2>는 예제 시스템의 구조를 표현한 것이다. 데이터 수집 및 분석의 모든 과정은 온라인 상에서 실시간으로 처리되며, 사용자는 웹 브라우저를 통하여 각종

정보를 제공받을 수 있다.

예제 시스템에서는 일별, 월별에 따른 로그 정보와 웹 페이지 이용빈도, 연관 규칙 (association rule), 로그 데이터를 통한 군집분석, 기존 고객에 대한 군집 분석 및 판별분석, 그룹별 구매 패턴 분석, 고객의 이탈 가능성, 상관분석 등을 포함하고 있으며, <그림 3>은 이탈 가능성 결과를 보여 주는 그림이다.

Clusters ID	Users ID
1	53 56 89 121 143 166 204 282 287 292 340 362 425 429 442
2	135 153 203 435
3	118 194 234 339 357 373
4	107 110 115 123 392
5	6 50 70 76 80 90 98 122 154 227 247 260 269 290 347 438 449
6	99 156 164 169 179 187 208 210 233 238 279 334 336 404 407
전체비율	62/98

<그림 3> 이탈가능성

4. 결론 및 연구과제

본 연구에서는 eCRM 기술에 대해 소개하고, 간단한 예제를 통하여 웹 데이터의 활용 방안을 살펴보았다.

eCRM과 같은 시스템들은 여러 가지의 데이터를 통합하여 분석하는 것을 필요로 한다. 이러한 데이터의 적절한 분석은 데이터가 대용량화 되어 가고 있는 현대 사회에서 매우 중요한 사안이라 평가되며, Garofalakis 등(1999), Shapiro(2000) 등은 이러한 분야들의 중요성을 강조하고 있다.

앞에서도 언급한 바와 같이 현재까지는 로그 데이터의 이용에 관한 연구가 대부분을 차지 하고 있으나, 앞으로 대용량 데이터 분석 방법에 관한 연구, 데이터 정제 및 통합 방법에 관한 연구, 개인별 특성을 고려한 효과적인 웹 서비스 방안, 웹 페이지 방문 패턴 등을 이용한 정보 추출 방안 등에 관한 연구가 필요할 것이라 생각된다.

참고문헌

- [1] Borges, J. and Levene, M. (1999), Data Mining of User Navigation Patterns, *Proceedings of the workshop on Web Usage Analysis and User Profiling (WebKDD'99)*, <http://www.acm.org/sigs/sigkdd/proceedings/webkdd99/toonline.htm>
- [2] Buchner, A.G., Baumgarten, M., Anand, S.S., Mulvenna, M.D. and Hughes, J.G. (1999), Navigation Pattern Discovery from Internet Data, *Proceedings of the workshop on Web Usage Analysis and User Profiling (WebKDD'99)*, <http://www.acm.org/sigs/sigkdd/proceedings/webkdd99/toonline.htm>
- [3] Garofalakis, M.N., Rastogi, R., Seshadri, S. and Shim, K.S. (1999), Data Mining and the Web: Past, Present and Future, *Proceedings of the workshop on Web Information and Data Management*, <http://cs.kaist.ac.kr/~shim/>
- [4] Guha, S., Rastogi, R. and Shim, K.S. (1998), CURE: An Efficient Clustering Algorithm for Large Databases, *Proceedings of ACM SIGMOD International Conference on Management of Data*, 73-84, <http://cs.kaist.ac.kr/~shim/>
- [5] Hartigan, J.A., Wong, M.A. (1979), Algorithm AS 136: A K-means clustering algorithm, *Applied Statistics*, Vol. 28, 100-108
- [6] Kohonen, T. (1984), *Self-Organization and Associative Memory*, Springer-Verlag
- [7] Madria, S., Bhowmick S. and Lim, E. P. (1999), Research Issues in Web Data Mining, *Proceedings of the International Conference on Data Warehousing and Knowledge Discovery*, <http://www.ntu.edu.sg/home/awkng>
- [8] Mobasher, B., Jain, N., Han, E.H. and Srivastava, J. (1996), Web Mining: Pattern Discovery from WWW Transactions, *Technical Report*, Department of Computer Science, University of Minnesota, <http://maya.cs.depaul.edu/~mobasher/pubs-subject.html>
- [9] Pei, J., Han, J., Mortazavi-asl, B. and Zhu, H. (2000), Mining Access Patterns Efficiently from Web Logs, *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, <http://db.cs.sfu.ca/sections/publication/kdd>
- [10] Shapiro, G.P.(2000), Knowledge Discovery in Databases: 10 years after, *SIGKDD Explorations*, Vol. 1, Issue 2, 59-61, <http://www.acm.org/sigs/sigkdd/explorations/>
- [11] Spiliopoulou, M., Pohle, C. and Faulstich, L.C. (1999), Improving the Effectiveness of a Web Site with Web Usage Mining, *Proceedings of the workshop on Web Usage Analysis and User Profiling (WebKDD'99)*, <http://www.acm.org/sigs/sigkdd/proceedings/webkdd99/toonline.htm>
- [12] Srivastava, J., Cooley, R., Deshpande, M. and Tan, P.N. (2000), Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, *SIGKDD Explorations*,

Vol. 1, Issue 2, <http://www.acm.org/sigs/sigkdd/explorations/>

- [13] Yan, T.W., Jacobsen, M., Molina, H.G. and Dayal, U. (1996), From User Access Patterns to Dynamic Hypertext Linking, *Proceedings of the fifth International World Wide Web Conference*, http://www5conf.inria.fr/fich_html/papers/P8/overview.html
- [14] Zhang, T., Ramarkrishnan, R. and Livny, M. (1996), BIRCH: An Efficient Data Clustering Method for Very Large Databases, *Proceedings of ACM SIGMOD International Conference on Data Management*, <http://citeseer.nj.nec.com/zhang97birch.html>