

# gCRM과 공간데이터마이닝

## ( gCRM and Spatial Data Mining )

황정래\*, 이기준

부산대학교 시공간데이터베이스연구실

Jung-Rae Hwang, Ki-Joune Li

Spatio-Temporal Database Lab. Pusan National University

[jrhwang@quantos.cs.pusan.ac.kr](mailto:jrhwang@quantos.cs.pusan.ac.kr), [lik@hyowon.cc.pusan.ac.kr](mailto:lik@hyowon.cc.pusan.ac.kr)

### 초록

고객관계관리(CRM)나 마케팅과 같은 경영 방식에서도 대용량의 공간 데이터베이스를 사용하는 지리정보시스템(GIS)과 같은 응용 분야를 접목하고 있다. gCRM은 지리정보시스템과 고객관계관리를 결합한 것으로, 이러한 실정을 단적으로 보여 주고 있는 경영 방식이다. gCRM은 대용량의 데이터베이스로부터 관심있는 분야를 찾아내고 분석하게 된다. 그러기 위해서는 데이터마이닝이라는 기술이 필요하다. 하지만, gCRM은 일반적인 데이터베이스뿐만 아니라 공간 데이터베이스 역시 많이 사용되어진다. 이러한 공간 데이터베이스로부터 관심있는 부분이나 관계 그리고 특성 등을 찾아내기 위해서는 공간데이터마이닝이 요구된다. 본 논문에서는 gCRM 솔루션들의 기능을 중심으로 다양한 공간데이터마이닝 기법과 어떠한 관계가 있는지를 살펴봄으로써 gCRM과 공간데이터마이닝이 접목할 수 있는 부분에 대하여 정리하였다.

### 키워드

CRM, gCRM, 공간데이터마이닝, 클러스터링, 공간적 연관 규칙, 일반화

### 1. 서론

고객관계관리(CRM)나 마케팅과 같은 경영 방식에서도 대용량의 공간 데이터베이스를 사용하는 지리정보시스템(GIS)과 같은 응용 분야를 접목하고 있다. gCRM은 지리정보시스템과 고객관계관리를 결합한 것으로, 이러한 실정을 단적으로 보여 주고 있는 경영 방식이다. 예를 들어 '부산 스포츠 센터를 중심으로 반경 1km내에 주거하는 년 소득이 3000만원 이상 되는 30대 남성의 고객분포 분석'이나 '부산 보험회사 일대의 어떤 정유회사 주유소에 대한 직장 여성들의 이용도 분석' 등과 같이 고객관리에 지리적인 요소가 추가됨으로써 고객에 대한 정보를 입체적으로 분석이 가능하게 된다. 이 때문에 gCRM은 대규모의 영업지점 관리가 필요한 금융업계나 유통업계 분야에서 주로 사용된다. 그리고, 고객위치의 속성을 파악하고 텍스트로 보여 주던 분석결과를 이해하기 쉬운 멀티미디어로 전달할 수 있다는 점에서 기존의 고객관

계관리와 차별된다고 할 수 있다.

gCRM은 대용량의 데이터베이스로부터 어떤 규칙적인 지식을 찾아내고, 분석하게 된다. 그러기 위해서는 데이터마이닝이 필요할 것이다. 현재의 gCRM은 고객관계관리와 지리정보시스템이 결합한 상태이다. 따라서, 고객관계관리에 적용된 데이터마이닝을 사용하고 있다[8]. 하지만 엄격히 말하면, gCRM은 일반적인 데이터베이스뿐만 아니라 원격탐사, 지리정보시스템, 컴퓨터 지도제작 그리고 환경 평가 및 계획 등 다양한 응용분야에서 수집되어진 공간 데이터로부터 이루어진 공간 데이터베이스 역시 많이 사용되어진다. 이러한 공간 데이터베이스로부터 관심있는 부분이나 관계 그리고 특성 등을 찾아내기 위해서는 공간데이터마이닝이 요구되어질 것이다.

그림 1은 공간데이터마이닝을 현 시점의 gCRM 구조에 접목시켜 어떻게 운용될 수 있는지 간단히 그림으로 나타낸 것이다.

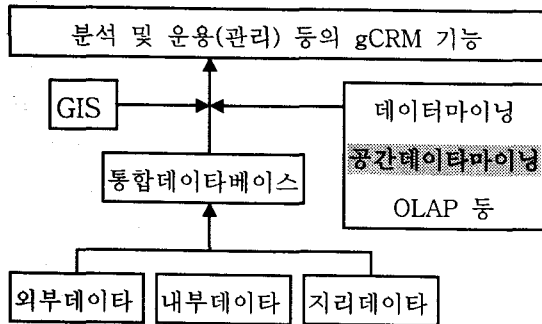


그림 1. 공간데이터마이닝을 포함한 gCRM의 구조

본 논문에서는 gCRM의 기능을 중심으로 하여 공간데이터마이닝이 gCRM에 유용하게 접목될 수 있는 부분을 정리하였다. 본 논문의 이러한 정리는 공간데이터마이닝을 gCRM에 접목시킴으로써 현재의 gCRM

을 한층 더 발전시킬 수 있을 것으로 예상된다.

본 논문의 구성은 다음과 같다. 2장에서는 현재 국내외에서 개발된 gCRM 솔루션들에 대하여 소개한다. 3장에서는 기존에 연구되어진 공간데이터마이닝 기법들에 대하여 살펴보고, 4장에서는 gCRM 솔루션들의 기능을 중심으로 공간데이터마이닝 기법과 접목될 수 있는 부분에 대하여 살펴본다. 마지막으로 5장에서는 결론 및 향후과제에 대하여 언급한다.

## 2. gCRM

gCRM은 국내뿐만 아니라 국외에서도 점차 급부상하고 분야로, 현재 국내외 몇몇 기업에서 gCRM 솔루션을 개발 및 구축하여 금융업계나 유통업계에서 사용하고 있다. 본 장에서는 국내외에서의 gCRM에 대한 연구 및 개발에 대하여 살펴보기로 한다.

표 1은 국내외에서 대표적인 gCRM 솔루션들을 표로 간단히 나타낸 것이다. 표 1에 있는 몇 가지 솔루션들을 예를 들어 간단히 살펴보기로 한다.

[9]에서 개발된 Marzer 솔루션은 gCRM을 위한 GIS 마케팅 분석 패키지이다. 그리고, GEHO(Geo Coding & House Holding)은 고객정보의 필수 요소인 주소에 대하여 올바른 주소 체제로 정제하는 솔루션이다. [9]와 더불어 국내 gCRM 시장에 크게 참여하고 있는 [10]에서 개발된 Terrician 솔루션은 GIS를 활용한 영업 마케팅 솔루션이다. 그리고, CAA(Commercial Area Analyzer) 솔루션은 상권분석 및 매출 예측 솔루션으로, [9]에서 개발된 GIS 상권분석 솔루션과 같은 시스템이다. [11]은 최근에 gCRM 시장에 참여한 업체로 비즈니스

표 1. 국내의 대표적인 gCRM 솔루션 분류

분류	솔루션	설명	비고
국내	Marzer	GIS 마케팅 분석 패키지	니즈아이
	GIS 목표배정	목표배정 및 성과관리 솔루션	니즈아이
	GEHO	주소정제 솔루션	니즈아이
	GIS 점포구조조정	영업지점을 조정(신설, 이전, 폐쇄, 확대 및 축소)	니즈아이
	GIS 상권분석	GIS를 통한 상권분석 솔루션	니즈아이
	Terrician	GIS를 활용한 영업 마케팅 솔루션	타스테크
	CAA	상권분석 및 매출예측 솔루션	타스테크
	시장분석 시스템	상권분석 활동을 지원하는 솔루션	오픈메이트
국외	Ameritech	고객의 위치를 파악해서 고객서비스를 증대	미국
	meineke	이익발생 가능시장 침투하여 타겟으로 결정	미국
	DSS Corp	GIS를 통해 핵심 타겟 지역에 대한 마케팅 집중	미국
	Microgis Corp	인구수 기준으로 표적집단 거주 지역을 포착하여 지능적인 마케팅 전개	스위스

지리정보시스템을 기업의 경영으로 삼고 있다. [11]에서 개발한 시장분석 시스템 솔루션은 지리 및 지역적 위치와 관련이 있는 고객이나 영업지점 정보를 체계적으로 관리 및 분석, 표현한다. 이것을 바탕으로 시장, 고객 그리고 점포망의 지역적인 분포 현황 파악을 기초로 합리적인 영업영역 분할과 시장변화 예측 및 대책 수립 등의 상권 분석 활동을 공간적, 시각적 그리고 과학적으로 수행할 수 있도록 지원하는 시스템이다.

유럽과 미국 등지에서는 국가적인 일환으로 인구통계자료, 정보화 자료, 부동산, 지형정보 등이 통합적으로 운영되고 있어 데이터의 확보가 용이하며 이러한 자료를 기업의 내부 데이터베이스와 융합하여 점포 전략, 고객성향분석, 타겟 마케팅 등의 gCRM 사업에 활용되고 있다[7]. 예를 들면, 표 1에서 meineke 솔루션은 이익 발생 가능 시장으로 침투하여 가장 유리한 시

장과 소비자 위치를 파악하고 이를 타겟으로 정하는 시스템이다.

이외에도 고객과 업체간의 유대감을 형성하는 전체고객관리 솔루션이나 고객 성향을 분석하여 차별화된 고객을 관리하는 특정 지역고객관리 솔루션 등 많은 솔루션들이 있지만, 본 논문에서는 몇몇 gCRM 솔루션들에 대하여 살펴보았다. 앞으로 gCRM이 빠르게 확산되기 위해서는 공간 데이터베이스와 고객정보에 대한 지속적인 상호보완 작업과 전문가의 양성이 무엇보다 시급하지만, 데이터 마이닝과 같은 gCRM의 기반적인 기술이 다른 나라에 뒤지지 않기 때문에 충분히 경쟁이 가능할 것으로 예상된다.

### 3. 공간데이터마이닝

공간데이터마이닝에 대한 연구는 이전의 데이터마이닝 연구에 이어 국외에서는 계속적

으로 연구되고 있으나 국내에서는 아직 미숙한 단계이다. 본 장에서는 국내외에서의 공간데이터마이닝 대한 연구가 어느 정도 이루어지고 있는지 살펴보기로 한다.

공간데이터마이닝과 관련하여 국내의 몇몇 기업에서는 공간데이터마이닝과 유사한 개념의 'gMining'이라는 기술을 연구 및 개발하였다[11]. 이 기술은 데이터마이닝과 지리정보시스템을 결합한 것으로, 지리정보시스템의 연산 기능을 새로운 분석 기능이나 모델로 확장시켜 데이터마이닝 전반에 걸쳐 수행 능력을 향상시키고자 하였다. 이 기술의 장점은 첫째, 분석에 꼭 필요한 데이터 처리를 시각적 방식으로 수행할 수 있어 효율성이 향상된다. 둘째, GIS를 이용한 연산을 통해 거리나 공간연관규칙 등 기존의 수치데이터 위주의 분석으로 변수화 하지 못했던 항목을 추가할 수 있어 분석의 예측력을 크게 향상시켜줄 수 있다. 셋째, GIS의 연산기능을 이용한 새로운 분석기법 개발 및 모델 구축이 가능하여 데이터마이닝의 영역을 한단계 확장 시킬 수 있는 효과가 있어 데이터마이닝 전반에 걸친 수행 능력 향상 및 활용도 제고가 가능하게 된다. 이 기술은 위와 같은 장점들을 이용하여 데이터마이닝 툴과 지리정보시스템의 자유로운 결합을 이끌어낼 수 있다.

공간데이터마이닝에 관한 연구는 국내보다 국외에 많이 연구되어 알려져 있다. 표 2는 기존에 연구된 대표적인 공간데이터마이닝 기법들에 대하여 간단히 표로 나타낸 것이다. 표 2에서 보듯이 가장 대표적인 공간데이터마이닝 기법은 공간 클러스터링이다. 공간 클러스터링 기법은 개념적인 계층구조와 같은 어떤 배경 지식을 사용하지 않고도 데이터로부터 바로 관심있는 구조나 클러스터를 찾는 방법이다.

표 2. 대표적인 공간데이터마이닝 기법

기법	예(참고문헌)
공간 클러스터링	DBSCAN[2], PAM[3,6], CLARANS[1,3,6]
공간특성 묘사	Neighbors[1,5]
공간적분류	Spatial decision tree[1,2]
공간방향 탐사	Pseudo code 제시[2], global trend, local trend[1,5]
일반화	개념계층구조[3]
공간적연관	<g_close_to, highway17>[2], X → Y(80%)[3]

공간 클러스터링 기법에는 CLARANS 같은 k-medoid 클러스터링[6], 기본적으로 데이터베이스에서 모든 객체의 이웃을 검색하는 Single scan 클러스터링 그리고, 밀도 기반의 클러스터링 등이 있다[1]. 공간특성묘사 기법(Spatial Characterization)은 데이터베이스의 선택된 서브셋에 대해 간결한 묘사나 서술 등을 찾기 위한 방법이다. 공간적 분류(Spatial Classification) 기법은 객체의 속성 값을 기반으로 한 주어진 클래스의 집합으로부터 하나의 객체를 하나의 클래스에 할당하는 방법이다. 공간 방향 탐사(Spatial Trend Detection) 기법은 주어진 시작 객체로부터 움직일 때 하나 또는 그 이상의 비공간 속성의 규칙적인 변화로 정의되는 방법이다.

[3]에서는 기존에 연구되어진 대표적인 공간데이터마이닝 기법으로 일반화 기반의 지식발견, 클러스터링 그리고 공간적 연관을 조사하는 기법 등에 대하여 소개하고 있다. 공간적 연관을 조사하는 기법은 공간특성묘사 기법과 많은 관련이 있다. [1]. 그리고, 공간데이터마이닝에 대하여 연구되어야 하는 배경을 다양한 실례를 들어 설명하고 있

다[3].

[4]에서는 공간 관계의 기본적인 세 가지 타입으로 위상관계, 거리관계 그리고 방향관계를 제시하고 있다. 그리고, 이 세 가지 타입을 사용하여 여섯 가지의 확장된 정의를 내리고 있다. [2]와 [4]에서는 연관된 이웃 관계를 유형화하고, 공간 객체 자신들을 접근하는 것을 피하고자 neighborhood index를 정의하여 제시하고 있으며, neighborhood index의 유무에 따른 neighbors 연산이 수행되는 비용을 예측하기 위한 cost model을 제시하고 있다. 자세한 내용은 참고문헌을 참조하기 바란다.

#### 4. gCRM과 공간데이터마이닝의 관계

본 논문의 2장과 3장에서 gCRM과 공간데이터마이닝에 대하여 기존의 연구를 살펴보았지만, gCRM을 목적으로 한 공간데이터마이닝 연구는 거의 없었다. 그리고, 현재 몇몇 기업에서 개발되고 있는 gCRM 역시 지리정보시스템과 고객관계관리를 통합한 것으로, 공간데이터마이닝이 아닌 일반적인 데이터마이닝 기법을 고객관계관리에 적용하고 있는 실정이다. 이렇게 될 경우에는 gCRM의 목적에 맞지 않을 수도 있다. 즉, 지리정보시스템의 기능이 단순히 시각화 기능으로 머물 가능성이 높다는 것이다. 진정한 gCRM이라면 데이터마이닝만을 적용할 것이 아니라 공간데이터마이닝 기법과 병행하여 적용하는 것이 타당할 것이다. 이러한 의미에서, 공간데이터마이닝 연구는 지리정보시스템뿐만 아니라, 앞으로 개발의 가능성이 높은 gCRM에도 많은 영향을 미칠 것으로 예상된다.

2장과 3장에서 gCRM 솔루션과 공간데

이터마이닝 기법에 대하여 간단히 살펴 보았다. 본 장에서는 gCRM의 기능들 중에서 공간데이터마이닝 기법들과 연관되는 기능에는 어떠한 것들이 있을 수 있는지 살펴보기로 한다.

gCRM은 크게 분석 측면의 gCRM과 운용 측면의 gCRM으로 나눌 수 있다. 따라서 gCRM의 기능도 고객 분석이나 시장 분석 같은 분석 측면의 기능과 마케팅 전략 등과 같은 운용 측면의 기능으로 나눌 수 있다.

표 3은 gCRM의 두 가지 기능과 공간데이터마이닝 기법간의 관계를 간단히 표로 나타낸 것이다. 물론, 항상 표 3과 같은 관계가 나타나는 것은 아니다.

표 3. 공간데이터마이닝 기법을 활용한 gCRM의 기능 요약

공간데이터마이닝 기법	gCRM의 기능
공간 클러스터링	분석 기능
공간특성 묘사	분석 및 운용(관리) 기능
공간적 분류	운용(관리) 기능
공간방향 탐사	분석 기능
일반화	분석 및 운용(관리) 기능
공간적 연관	분석 기능

표 3을 참고로 각각의 공간데이터마이닝 기법들이 어떠한 기능을 담당하며 관련이 있는지 간단한 예를 들어 설명하였다.

공간 클러스터링 기법은 데이터베이스의 객체들을 클러스터화 함으로써 그 클러스터된 객체들로부터 유사성을 찾게 된다[6]. 이러한 기법은 영업지점별 고객분석, 지역별 고객분석 그리고 지역별 시장분석 등 유사한 부분들을 그룹화 시켜 사용자들이 원하는 결

과를 얻어내는데 유용한 방법이다. 그리고, 각 영업지점의 거리 변화에 따라 고객의 점유율을 분석하는 곳에도 적용될 수 있다. 물론, 다양한 클러스터링 기법이 존재하기 때문에 더 많은 기능들에 적용될 수 있을 것이다.

공간적 특성묘사 기법은 어떤 선택된 객체에 대하여 공간과 비공간 속성을 묘사하는 것으로, 공간적으로 확장 분석이 가능하다. 따라서, 이 기법은 객체의 묘사와 분석이 가능하다는 점에서 분석과 운용 기능에 적용될 수 있을 것이다.

공간적 분류 기법은 고객을 성향이나 등급을 감안해서 지역적으로 분류한 후 차별화된 고객을 유치하고자 할 경우에 적용될 수 있을 것이다.

공간 방향탐사 기법은 어떤 지역에서 시작하여 다른 지역으로 방향성을 가지면서 시장성이나 고객 평균 소득 등이 증가 또는 감소할 수 있을 것이다. 이와 같이 시장성이나 고객 평균 소득 등을 방향적으로 분석하고자 경우 적용될 수 있을 것이다.

공간적 연관 규칙 기법은 하나 또는 그 이상의 객체가 다른 객체들과의 연관 관계를 규칙화 시키는 방법이다. 예를 들면, '많은 회사들이 모여있는 곳이나 상업 지역에는 하나 이상의 은행이 존재할 가능성이 높다'거나 '대규모 아파트 단지를 구성한 곳에는 스포츠 센터가 근교에 있을 가능성이 높다' 등의 관계가 성립할 수 있다. 이렇게 공간적 연관 규칙 기법은 gCRM의 공간적으로 연관이 있는 객체를 조사하거나 그런 관계를 분석하는 등의 기능에 적용되어질 수 있다.

운용 측면의 기능은 분석을 통하여 영업의 목표나 성과를 측정하는 분석 측면의 기능과는 달리, 일반적인 영업지점 관리나

고객 관리 등의 기능을 가진다. 다시 말해서, 운용 측면의 기능은 관리 측면의 기능이라고 해도 과언이 아니다. 예를 들어, 행정동별로 고객의 실적을 분류하여 관리한다고 하자. 각 동별로 고객을 관리할 경우, 자연스럽게 각 구별로도 관리가 가능할 것이며, 더 나아가 각 시별로도 관리가 가능하게 된다. 이러한 관리는 공간데이터마이닝 기법 중 하위 단계로부터 상위 단계로 계층화 시키는 일반화 기법에 적용되어질 수 있다. 일반화 기법은 영업지점을 합병할 경우에도 적용될 수 있다. 어떤 영업지점의 활동 영역이 상실된다면 다른 영업지점과 결합이 불가피하게 된다. 이와 같이 영업지점의 합병에 따른 상승효과를 일반화 기법에 적용하면 업무 영역의 범위를 분석하는데 유용할 것이다.

오늘날의 gCRM은 공간데이터마이닝을 적용하지 않고, 이전에 고객관계관리에 적용된 데이터마이닝을 그대로 사용하고 있다. 그래서, 현재 gCRM에 사용되고 있는 데이터마이닝은 공간적인 관계를 분석하는데 한계가 있다. 데이터마이닝은 고객관계관리에 적용된 것으로, 고객관계관리는 대용량의 공간 데이터베이스보다는 고객관리와 관련된 일반적인 데이터베이스가 대부분을 차지하기 때문이다. 따라서, 본 논문에서는 공간데이터마이닝과 gCRM의 관계를 살펴보았다. 특히, 공간데이터마이닝 기법들과 gCRM의 기능을 중심으로 살펴보았으며, 이들 관계에서 서로 관련된 부분이 존재한다는 것을 알아보았다. 그리고, gCRM의 기능에 공간데이터마이닝 기법을 적용한다면 좀더 쉽고 빠르게 공간적으로 분석이 가능할 것으로 예상된다.

## 5. 결론

지리정보시스템의 발달로 CRM업체에서는

지리정보시스템을 고객관계관리에 접목하고자 하는 연구 및 개발이 활발하게 이루어지고 있다. gCRM은 지리정보시스템과 고객관계관리를 결합한 경영방식 형태로 이러한 실정을 잘 보여주고 있다. gCRM은 분석 및 관리 그리고 활용 등의 기능을 가지고 있다. 이러한 기능들은 공간데이터마이닝 기법들을 이용하면 기존의 기능들보다 개선된 성능을 보일 수 있다.

본 논문에서는 gCRM 기능들과 공간데이터마이닝 기법들 사이에 서로 접목할 수 있는 부분들을 정리하였다. 본 논문의 이러한 정리는 공간데이터마이닝을 gCRM에 접목시킴으로써 현재의 gCRM을 한층 더 발전시킬 수 있을 것으로 예상된다.

향후 연구과제는 공간데이터마이닝 기법들을 직접 gCRM에 적용하여 이전의 gCRM보다 기능의 성능면에서 우수하다는 것을 입증할 것이다. 그리고, 이전에 불가능하였던 기능이지만, 공간데이터마이닝을 사용함으로써 가능한 gCRM의 기능들을 추가적으로 찾아낼 것이다.

## 6. 참고문헌

- [1] Ester M., Frommelt A., Kriegel H.-P., Sander J. "Spatial Data Mining: Database Primitives, Algorithms and Efficient DBMS Support", *In Data Mining and Knowledge Discovery*, 4, pp. 193-216, 2000
- [2] Ester M., Kriegel H.-P., Sander J. "Spatial Data Mining: A Database Approach", *Proc. 5th Int. Symp. on Large Spatial Databases*, pp. 47-66, 1997
- [3] Koperski K., Adhikary J., Han J. "Spatial Data Mining: Progress and Challenges", *Proc. SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, TR 96-08, University of British Columbia, Vancouver, Canada, 1996
- [4] Ester M., Gundlach S., Kriegel H.-P., Sander J. "Database Primitives for Spatial Data Mining", *Proc. Int. Conf. on Databases in Office, Engineering and Science*, (BTW '99), Freiburg, Germany, 1999
- [5] Ester M., Frommelt A., Kriegel H.-P., Sander J. "Algorithms for Characterization and Trend Detection in Spatial Databases", *Proc. 4th Int. Conf. on Knowledge Discovery and Data Mining*, pp. 44-50, 1998
- [6] Ng R. and Han J. "Efficient and effective clustering methods for spatial data mining", *Proc. 20th Int. Conf. VLDB*, pp. 144-155, 1994
- [7] 한국지리정보, *gCRM: GIS와 고객관리 CRM의 만남*, pp. 58~63, 2001.10
- [8] 홍성완., *CRM을 위한 데이터마이닝*, 대청, 2000
- [9] (주)니즈아이, [www.needsi.com](http://www.needsi.com)
- [10] (주)타스테크, [www.tastech.co.kr](http://www.tastech.co.kr)
- [11] (주)오픈메이트, [www.openmate.co.kr](http://www.openmate.co.kr)