

GPS 데이터를 이용한 이동객체의 이동패턴 분석

조재희^a 서일정^b 이덕규^c 하병국^d

^a 광운대학교 경영대학 경영정보학과 교수
서울특별시 노원구 월계동 447-1 광운대학교
Tel: +82-2-940-5433, Fax: +82-2-940-5430, E-mail: misl@kw.ac.kr

광운대학교 경영대학 경영정보학과 박사과정
서울특별시 노원구 월계동 447-1 광운대학교
Tel: +82-2-940-5433, E-mail: ^beily@kw.ac.kr, ^cjerry@kw.ac.kr, ^ddjkoogy@kw.ac.kr

Abstract

GPS 수신기의 지속적인 가격 하락과 GPS 기반의 다양한 위치기반서비스 개발로 인하여 개인 휴대용 GPS 수신기의 보급이 확대되고 있다. 이동객체의 위치 및 시간 정보를 포함하고 있는 GPS 데이터를 분석하면 이전에는 불가능했던 이동패턴을 파악하고 이해하는 것이 가능해진다. 이동객체 데이터의 저장과 분석에 관한 연구들이 진행되고 있지만, 이동객체의 속성에 따른 다차원적 이동패턴 분석에 관한 연구는 찾아보기 힘들다. 본 연구는 개인 휴대용 GPS 수신기를 통해 수집된 이동 데이터와 이동객체의 속성 데이터를 통합하여 이동객체의 시공간적 특성을 다차원적으로 분석할 수 있는 데이터마트를 구현하고 시각적으로 표현하였다. 이러한 과정을 통해 GPS 데이터를 이용한 이동패턴 분석의 유용성과 문제점을 탐색적으로 살펴보았다.

Keywords:

GPS, 이동객체, 데이터마트, 다차원 데이터 모델링, 데이터 시각화, OLAP

1. 서론

GPS(Global Positioning System)는 1970년대 초반 미국방성에서 자국의 군사 목적을 위하여 개발한 것으로 총 24개 위성이 고도 약 20,200km 상공에 배치되어 24시간 전 세계 어디에서나 위치와 시간 측정이 가능한 시스템이다[14]. GPS 수신기의 지속적인 가격 하락과 GPS 기반의 다양한 위치기반서비스 개발로 인하여 개인 휴대용 GPS 수신기의 보급이 확대되고 있다[19]. 이는 GPS 데이터 분석의 필요성을 제기한다.

시간의 흐름에 따라 위치 및 모양이 연속적으로 변화하는 특징을 가지는 객체를 이동객체라고 하며[8], 이동객체의 위치 및 시간 정보를 포함하고 있는 GPS 데이터를 분석하면 이전에는 불가능했던

이동패턴을 파악하고 이해하는 것이 가능해진다. 위치기반서비스 또는 여행/관광, 테마공원, 레저 등 지리적 공간을 서비스 기반으로 하는 분야에 GPS 데이터 분석을 활용한다면, 고객의 이동패턴에 대한 이해를 바탕으로 다각적인 서비스 개선을 기대할 수 있다.

GPS 데이터는 시간에 따른 이동객체의 위치 변화를 기록한 시공간 데이터(spatiotemporal data)로 이를 효과적으로 분석하기 위해서는 기존과는 다른 방식을 요구한다. 이러한 요구에 따라 시공간 데이터를 저장하고 분석하기 위한 시공간 데이터 모델링[13][15][17], 이동객체 데이터 모델링 [2][6][9][16], 지리적 시각화[1][10][11]에 대한 선행 연구들이 진행되고 있다. 이상의 연구들은 개별 이동객체의 추적과 이동패턴 표현에 집중되어 있고, 이동객체의 속성(성별, 연령, 직업 등)에 따른 다차원적 이동패턴 분석에 관한 연구는 찾아보기 힘들다.

본 연구는 일련의 분석 프로세스를 통해 GPS 데이터를 시각적으로 분석함으로써, GPS 데이터를 이용한 이동패턴 분석의 유용성과 문제점을 탐색적으로 살펴보려고 하였다. 이를 위하여 ‘축구경기의 선수이동’, ‘K대학교 학생의 캠퍼스이동’, ‘등산객의 산행’을 분석 대상으로, 개인 휴대용 GPS 수신기를 통해 수집된 데이터와 이동객체의 속성 데이터를 통합하여 이동객체의 시공간적 특성을 다차원적으로 분석할 수 있는 데이터마트를 구현하고 시각적으로 표현하였다.

2. GPS 데이터

GPS는 크게 우주부문, 제어부문, 사용자부문에 구성된다. 우주부문은 총 24대의 GPS 위성들로 구성되며 위성의 시간, 위치, 작동상태, 시간보정 등의 정보를 사용자에게 전송하는 역할을 수행한다. 제어부문은 우주부문의 위성들을 관리하기 위해 지상에 설치된 관제소로 위성을 추적하여 위성상태 및 위성정보를 수정하는 업무를 수행한다. 사용자 부문은 GPS 수신기로 위성으로부터 전송된 정보를 수신하고 처리한다. GPS 위성은 자신의 고유 ID,

궤도위치, 시간정보를 24시간 송출하게 되고 GPS 수신기는 위성에서 발사하는 이들 정보를 수신하여 자신의 위치를 계산한다.

GPS 데이터 표준 포맷에는 RINEX, NGS-SP3, RTCM SC-104, NMEA 0183 등이 있으며[7], 본 연구에 사용한 <그림 1>의 Sony GPS-CS1 수신기는 NMEA 0183을 데이터 표준 포맷으로 사용한다. NMEA 0183은 1983년도 NMEA(National Marine Electronic Association)에 의해 정의된 표준 포맷으로 <그림 2>와 같이 ASCII 형태로 구성되어 있다.

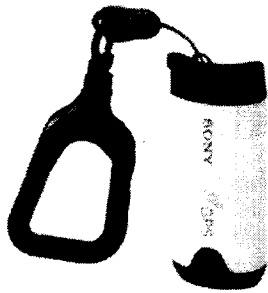


그림 1 - Sony GPS-CS1

```
#Sonygps/ver1.0/wgs-84
$GPGGA,06.1005,3737.1777,N,12703.5473,E,1,03,11.2,00062.6,M,018.5,M,,0000
$GPGSA,A,2,24,26,29,,,,,,,,,11.3,11.2,01.0=00
$GPGSV,2,1,06,24,65,216,50,26,65,021,49,29,54,036,46,09,,,44=44
$GPGSU,2,2,06,10,,,36,07,,,00,,,,,,,,,7C
$GPRMC,06.1005,A,3737.1777,N,12703.5473,E,001.2,000.0,210407,,,A=75
$GPVTG,000.0,T,,M,001.2,N,002.2,K,A=0E
$GPGGA,06.1020,3737.1835,N,12703.5492,E,1,03,11.2,00062.6,M,018.5,M,,0000
$GPGSA,A,2,24,26,29,,,,,,,,,11.3,11.2,01.0=00
$GPGSV,2,1,05,24,65,215,40,26,64,021,46,29,54,037,45,09,,,45=48
$GPGSU,2,2,05,10,,,46,,,,,,,,,,,,,7F
```

그림 2 - GPS-CS1의 GPS 데이터

GPS-CS1에 기록된 GPS 데이터는 크게 5개의 구문(sentence)으로 구성되어있으며, 각 구문은 <표 1>의 정보를 포함하고 있다.

표 1 - GPS-CS1에 기록된 GPS 데이터 정보

문장	설명
\$GPGGA	시간, 경도, 위도, 위치오차, 사용된 위성 수, 고도
\$GPGSA	위치를 결정하는데 사용된 위성 수, 측정의 정확도
\$GPGSV	측정에 사용된 위성 수, 위성의 고유번호, 고도, 방위각, 신호대비 잡음비율
\$GPRMC	시간, 위도, 경도, 고도, 시스템장애, 속도, 경로, 날짜
\$GPVTG	진행방향, 속도

GPS 데이터의 좌표는 WGS84(World Geodetic System 1984) 경위도 좌표로, GPS 데이터를 수치지도 상에 표현하기 위해서는 <그림 3>의 과정을 거쳐 Bessel 타원체에 준거한 TM(Transverse Mercator) 좌표로 변환해야 한다.

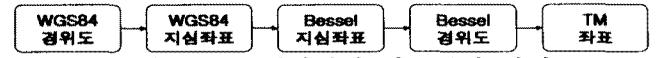


그림 3 - GPS데이터의 좌표변환 과정

3. GPS 데이터의 시각적 분석

GPS 데이터를 시각적으로 분석하기 위하여 <그림 4>의 분석 프로세스를 수행하였다. 데이터 수집 단계에서는 GPS 수신기를 휴대한 개별 이동객체의 위치 및 시간정보가 실시간(GPS-CS1은 매 15초)으로 기록된 GPS 수신기를 컴퓨터에 연결하여 ASCII 형태의 로그파일을 추출한다. 데이터 전처리 단계에서는 분석의 정확성을 위하여 위치오차가 큰 데이터를 정제한다. GPS 데이터에서 제공하는 위치오차(DOP; Dilution of Precision)는 수평(HDOP), 수직(VDOP), 3차원(PDOP), 기하학(GDOP)위치오차가 있다. 평면의 경우 통상적으로 수평 위치오차가 6이상이면 오차가 큰 것으로 간주하고 정제한다. 그 다음 WGS84 경위도 좌표를 TM 좌표로 변환한다. 다차원 데이터 모델링 단계에서는 GPS 데이터, 이동객체의 속성 데이터, 분석요구 등을 고려하여 차원, 차원항목, 측정지표를 설정하고 차원항목의 계층구조 및 측정지표의 관계를 정의한다. 데이터마트 구현 단계에서는 다차원 데이터 모델에 준거하여 GPS 데이터와 이동객체의 속성 데이터를 통합하여 물리적인 데이터베이스를 구축한다. 마지막으로 시각적 분석 단계에서는 분석도구를 사용하여 이동패턴을 시각적으로 표현하고 탐색한다. 본 연구에 사용한 분석도구는 Tableau 3.0[20]이다.

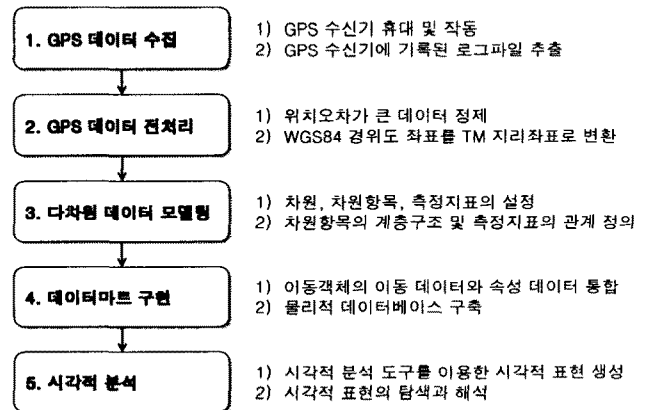


그림 4 - GPS 데이터의 분석 프로세스

축구경기의 선수이동, K대학교 학생의 캠퍼스이동, 등산객의 산행을 분석 대상으로 이상에서 설명한 프로세스에 의해 분석을 실시하였다.

3.1 분석1: 축구경기의 선수이동

2007년 4월 14일 실시한 축구경기에서 한 팀에 소속된 9명의 선수를 대상으로 GPS 데이터를 수집하였으며, 분석을 위하여 4개의 차원과 4개의 측정지표로 구성된 <그림 5>의 다차원 데이터 모델을 도출하였다. 선수번호, 포지션, 전후반은 이동객체의 속성데이터이며 시간, X좌표, Y좌표, 속도, 이동거리는 이동데이터이다. 시간, X좌표, Y좌표는 GPS로부터 수집된 원시 데이터를 축구선수의 이동패턴 분석 목적에 맞도록 가공하였다. 또한, 이동거리는 좌표를 이용하여 파생시킨 측정지표이다.

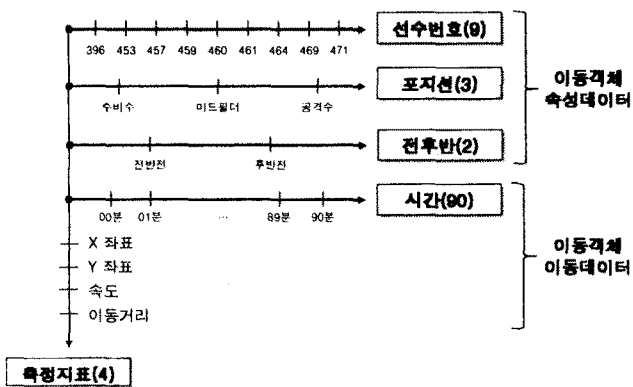


그림 5 - 축구선수 이동패턴 분석을 위한 다차원 데이터 모델

축구선수의 이동패턴을 시각적으로 분석하기 위하여 배경이미지로 축구장을 사용하였으며 이미지의 각 지점은 X, Y 좌표와 연계되어있다. <그림 6>은 특정 시점에서 선수들의 위치를 원으로 표현한 것으로, 원의 색은 각 선수를 표현하며 원의 크기는 해당 지점에서의 선수의 속도를 표현한다. 각 원에 선수의 포지션을 표기하였다. 경기시간 동안의 모든 시간에 대해 이와 같은 시각적 표현이 가능하며 지면에는 표현할 수 없지만 Table 3.3에서는 애니메이션 기능을 사용하여 시간의 흐름에 따라 이동객체의 위치를 추적할 수 있다.

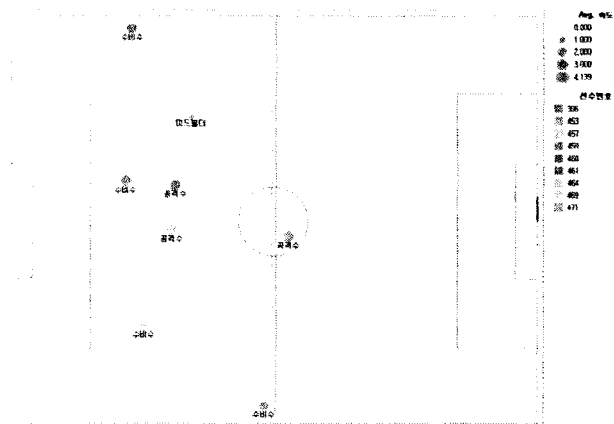


그림 6 - 선수위치와 속도

<그림 7>은 3분 동안 선수의 이동궤적을 표현한 것으로 선의 굵기는 속도를 표현하고 각 지점에 시간(분)을 표기하여 시간의 흐름에 따른 선수의 위치변화를 파악할 수 있다.

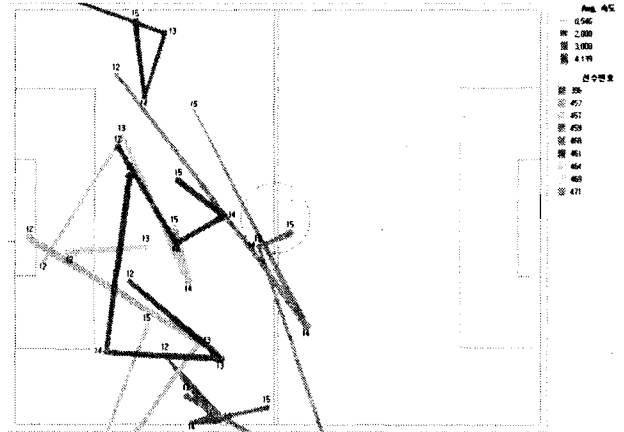


그림 7 - 선수 이동궤적

<그림 8>은 경기시간 동안 포지션별 속도변화를 표현한 것이다. 추세선을 추가하여 속도변화를 직관적으로 파악할 수 있도록 하였다.

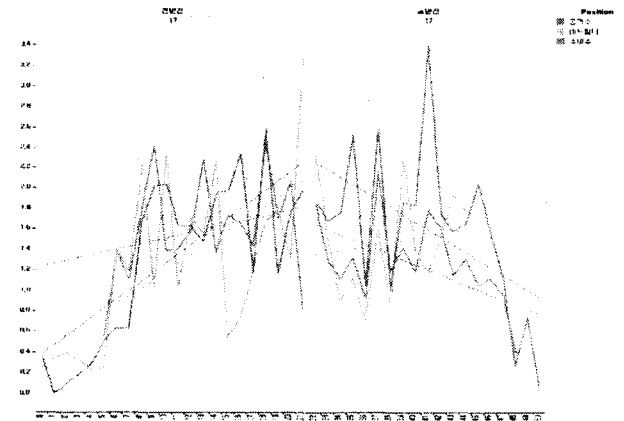


그림 8 - 포지션별 속도

3.2 분석2: K대학교 학생의 캠퍼스이동

캠퍼스이동을 분석하기 위하여 K대학교 학생 4명을 대상으로 2007년 4월 23일부터 4월 27일까지 5일 동안 오전 9시부터 오후 6시까지 GPS 데이터를 수집하였으며, 6개의 차원과 4개의 측정지표를 갖는 <그림 9>의 다차원 데이터 모델을 생성하였다. 번호, 학과, 학번 차원은 이동객체의 속성데이터이며 일, 시, 분, X좌표, Y좌표, 속도, 이동거리는 이동객체의 이동데이터이다.

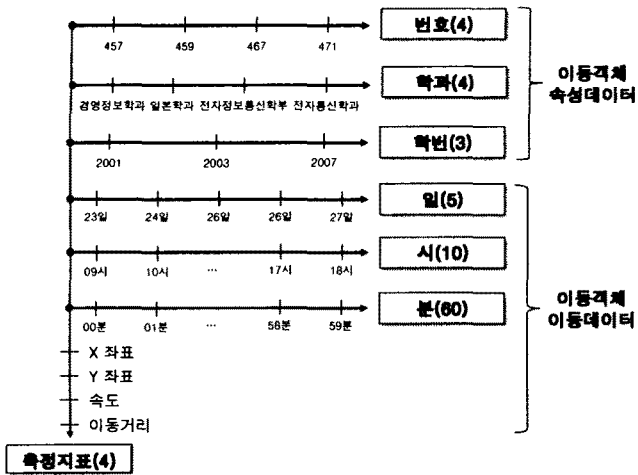


그림 9- 캠퍼스 이동패턴 분석을 위한 다차원 데이터 모델

참가학생들의 캠퍼스 이동패턴을 시각적으로 분석하기 위하여 인터넷 전자 지도 제공업체인 콩나물[18]의 위성지도서비스를 이용하여 K대학 주변 지도를 획득하고, 이를 캠퍼스 건물 중심으로 단순화하여 제작한 캠퍼스 지도를 배경이미지로 사용하였다. 지도상의 회색 부분은 K대학 건물이다. <그림 10>은 4월 25일 참가학생들의 캠퍼스 이동궤적을 표현한 것으로 선의 색상으로 개별 학생들을 구분하였다.

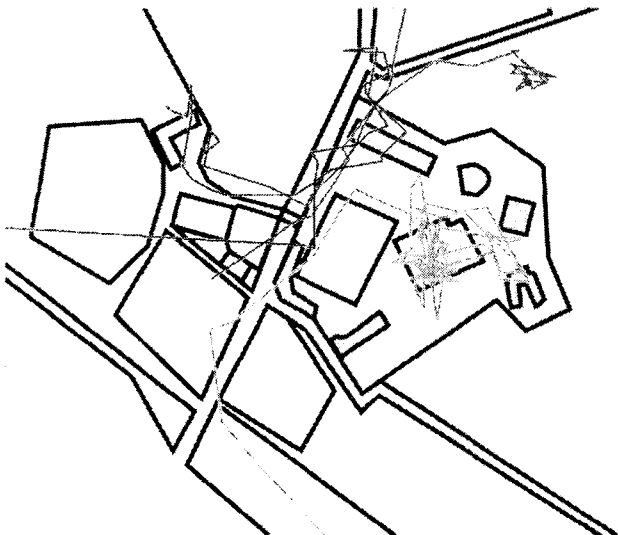


그림 10- 캠퍼스 이동궤적

<그림 10>의 이동궤적을 이동객체의 속성과 시간에 따라 구분하면 <그림 11>과 같이 표현할 수 있다. 행은 시간대(14시~18시)이고 열은 참가학생의 학과(경영정보학과, 전자정보통신학부, 전자통신학과)이다. 이와 같은 분석은 모든 시간차원(월, 일, 시) 및 속성차원(번호, 학과, 학번)에서 가능하며 이는 <그림 11>처럼 차원항목간 비교 분석을 용이하게 한다.

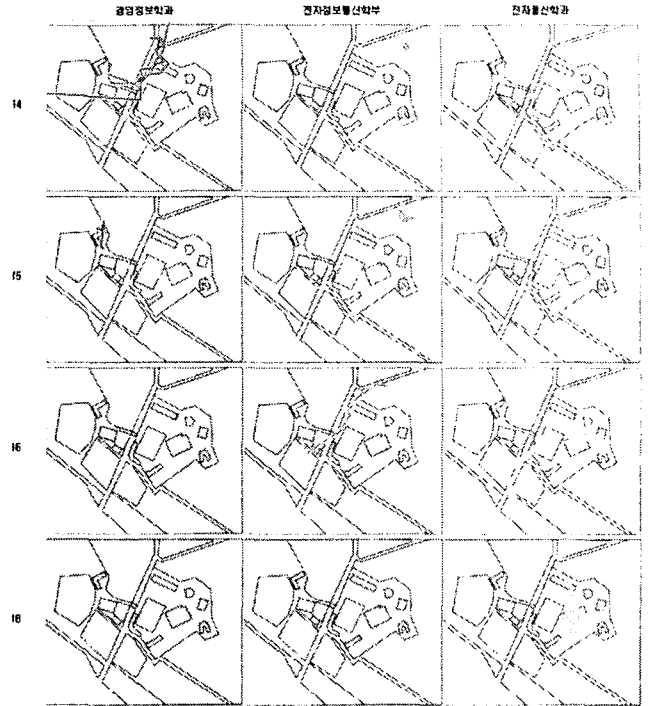


그림 11- 학과 및 시간대별 캠퍼스 이동궤적

3.3 분석3: 등산객의 산행

GPS 데이터는 경도와 위도 좌표뿐만 아니라 고도 정보를 포함하고 있다. 등산객의 산행분석은 GPS 데이터의 고도 정보를 시각적으로 표현하기 위하여 시행되었으며 2007년 4월 28일 서울 근교의 수락산을 등산한 데이터를 사용하였다.

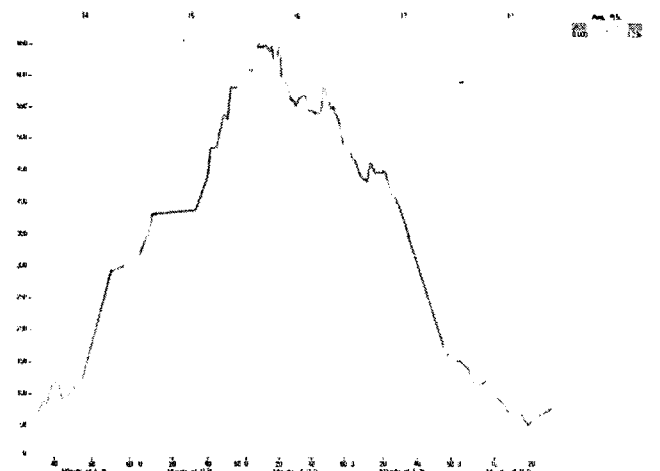


그림 12- 등산객의 시간대별 고도변화

<그림 12>는 시간이 지남에 따라 등산객의 고도 변화를 보여주고 있다. 선의 색은 등산객의 속도로 빨간색은 속도가 느린 것이고 파란색은 빠른 것이다.

4. 결론

개인 휴대용 GPS 수신기의 보급 확대로 GPS 데이터 분석의 필요성이 제기되고 있는 시점에서, 본 연구는 다차원 데이터 모델링 방법과 시각적 분석도구를 사용하여 일련의 분석 프로세스를 통해 GPS 데이터를 분석하였다. ‘축구경기의 선수이동’, ‘K대학교 학생의 캠퍼스이동’, ‘등산객의 산행’을 분석 대상으로, 개인 휴대용 GPS 수신기를 통해 수집된 데이터와 이동객체의 속성 데이터를 통합하여 이동객체의 시공간적 특성을 다차원적으로 분석하기 위한 데이터마트를 구현하고 시각적으로 표현하였다.

분석 과정에서 탐색된 GPS 데이터 분석의 유용성 및 문제점을 정리하면 다음과 같다. 첫째, GPS 데이터 분석은 기존에는 불가능했던 이동객체의 이동패턴 파악을 가능하게 한다. 둘째, 이동객체의 이동을 지리적 이미지(축구장, 캠퍼스 지도 등)를 사용하여 시각적으로 표현하는 것은 이동패턴에 대한 사용자의 직관적인 이해를 돕는다. 셋째, 특정 제품에 국한된 문제일 수도 있지만 개인 휴대용 GPS 수신기의 데이터 정확성을 보장할 수 없다. 건물 옆, 육교 밑 등에서는 수신이 불가능하거나 위치오차가 심하였다. 따라서 개인 휴대용 GPS 수신기를 통해 이동패턴을 세밀하게 분석하기에는 부리가 있는 것으로 판단된다.

본 연구는 다차원 데이터 모델링 방법을 사용하여 이동객체의 속성에 따른 이동패턴을 다차원적으로 분석하였다는 점에서 이동객체 또는 시공간 데이터에 관한 기존 연구들과 차별성을 갖는다. 또한 GPS 데이터 분석을 통해 이동객체(고객)에 대한 이해를 높일 수 있다는 가능성을 보여주었다.

GPS 수신기의 수신율을 높이고 위치오차를 줄일 수 있는 기술적 대안들을 모색하여, 테마공원의 서비스 개선을 위한 GPS 데이터 분석에 관한 연구를 계획하고 있다.

참고문헌

- [1] 박기호, 안재성, 이양원. (2005). “시공간 개인통행 자료의 지리적 시각화,” *대한지리학회지*, Vol. 40, No. 3, pp. 310-320.
- [2] 이기준, 김경숙. (2000). “이동객체의 움직임에 대한 시간적 명세,” *한국정보과학회 데이터베이스연구회지*, Vol. 16, No. 1, pp. 27-39.
- [3] 조재희, 서일정. (2005). “OLAP을 활용한 지하철 인구이동 맵 생성에 관한 연구,” *Information Systems Review*, Vol. 7, No. 1, pp. 65-80.
- [4] 조재희, 서일정. (2006). “시공간데이터 분석을 위한 다차원 모델과 시각적 표현에 관한 연구,” *Journal of Information Technology Applications & Management*, Vol. 13, No. 1, pp. 137-147.
- [5] 조재희, 서일정. (2007). “이동객체의 데이터 시각화를 통한 이동패턴 분석에 관한 연구,” *한국IT서비스학회지*, Vol. 6, No. 1, pp. 127-140.
- [6] Brakatsoulas, S., Pfoser, D., and Tryfona, N. (2004). “Modeling, Storing and Mining Moving Object Databases,” *Proceedings of the 7th International Database Engineering and Applications Symposium*, pp. 68-77.
- [7] El-Rabbany, A. (2002). *Introduction to GPS: the Global Positioning System*. Norwood, MA: Artech House.
- [8] Erwig, M., Güting, R.H., Schneider, M., and Vazirgiannis, M. (1999). “Spatio-Temporal Data Types: An Approach to Modeling and Querying Moving Objects in Databases,” *Geoinformatica*, Vol. 3, No. 3, pp. 269-296.
- [9] Forlizzi, L., Güting, R.H., Nardelli, E., and Schneider, M. (2000). “A Data Model and Data Structures for Moving Objects Databases,” *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pp. 319-330.
- [10] Kraak, M. (2003). “The Space-time Cube Revisited from a Geovisualization Perspective,” *Proceedings of the 21st International Cartographic Conference*, pp. 1988-1996.
- [11] Kwan, M. (2000). “Interactive Geovisualization of Activity-Travel Patterns Using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a Large Data Set,” *Transportation Research Part C*, Vol. 8, pp. 185-203.
- [12] MacEachren, A.M., and Kraak, M.J. (2001). “Research challenges in geovisualization,” *Cartography and Geographic Information Systems*, Vol. 28, No. 1, pp. 3-12.
- [13] Pfoser, D., and Tryfona, N. (1998). “Requirements, Definitions and Notations for Spatiotemporal Application Environments,” *Proceedings of the 6th International Symposium on Advances in Geographic Information Systems*, pp. 124-130.
- [14] Quiroga, C.A., and Bullock, D. (1998). “Travel time studies with global positioning and geographic information systems: an integrated methodology,” *Transportation Research 6C*, Vol. 1, No. 2, pp. 101-127.
- [15] Savary, L., Wan, T., and Zeitouni, K. (2004). “Spatio-Temporal Data Warehouse Design for Human Activity Pattern Analysis,” *Proceedings of the 15th International Workshop on Database and Expert Systems Applications*, pp. 814-818.
- [16] Sistla, A.P., Wolfson, O., Chamberlain, S., and Dao, S. (1997). “Modeling and Querying Moving Objects,” *Proceedings of the 13th International Conference on Data Engineering*, pp. 422-432.
- [17] Worboys, M.F. (1994). “A Unified Model of Spatial and Temporal Information,” *Computer Journal*, Vol. 37, No. 1, pp. 26-34.
- [18] 콩나물, <http://www.congnamul.com>
- [19] GPSWorld.com, <http://www.gpsworld.com>
- [20] Tableau Software, <http://www.tableausoftware.com>