

# 상태도에 기반한 택시 텔레매틱스 히스토리 데이터 분석<sup>†</sup>

## Analysis of the taxi telematics history data based on a state diagram

이정훈\* / Junghoon Lee, 권상철\*\* / Sangcheol Kwon

### 요약

본 논문은 제주 택시 텔레매틱스 사업에서 수집된 차량의 히스토리 데이터를 온라인 혹은 오프라인으로 효율적으로 분석하기 위한 스테이트 다이어그램을 정의하고 이에 기반하여 택시의 운행과 배차에 대한 분석 결과를 산출하는 것을 목표로 한다. 차량에서 수집된 정보는 기본적인 GPS 정보 이외에 차량의 상태를 나타내는 필드를 포함하고 있으며 지도 정보의 결합으로 맵 매칭 및 도로상에서의 위치 비율 등을 계산할 수 있다. 구축된 레코드들에 의해 택시의 승객 탑승 빈도, 탑승시 이동 거리, 탑승 시간 등을 분석할 수 있으며 콜 택시의 중요한 성능 척도가 되는 배차 후 차량의 이동 거리 및 이동 시간을 분석하여 배차 방식의 효율성을 검증할 수 있다. 이 정보는 심도있는 분석을 수반하여 향후 교통 정보 예측, 혼잡상황 회피 등의 다양한 응용의 진화를 가능하게 한다.

### Abstract

This paper presents a data analysis method for the taxi telematics system which generates a great deal of location history data. By the record consist of the basic GPS receiver-generated fields, device-added fields such as taxi operation status, and framework-attached fields such as matched link identifier and position ratio in a link, each taxi can be represented by a state diagram. The transition and the state definition enable us to efficiently extract such information as pick-up time, pick-up distance, dispatch time, and dispatch distance. The analysis result can help to verify the efficiency of a specific taxi dispatch algorithm, while the analysis framework can invite a new challenging service including future traffic estimation, trajectory clustering, and so on.

**주요어 :** 택시 텔레매틱스, 히스토리 데이터, 상태도, 배차시간 및 거리 분석, 데이터처리 프레임워크

**Keyword :** Taxi telematics system, history data, state diagram, dispatch time and distance analysis, data processing framework

† 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음  
(IITA-2008-C1090-0801-0040)

■ 논문접수 : 2008.1.30      ■ 심사완료 : 2008.3.8

\* 제주대학교 전산통계학과 교수(jhlee@cheju.ac.kr)

\*\* 교신저자 제주대학교 사범대학 지리교육전공 교수(kwonsc@cheju.ac.kr)

## 1. 서론

텔레매틱스는 컴퓨터를 기반으로 한 정보 기술과 무선 통신이 결합된 용어로서 다양한 민간 혹은 공공 서비스를 광범위의 네트워크를 통해 전달함을 그 목적으로 하고 있다[1]. 차량 텔레매틱스는 차량 내에 텔레매틱스 장치를 설치하여 차량 내에서도 운전자 혹은 사용자들로 하여금 단절없이 다양한 정보 서비스에 연결이 될 수 있도록 지원한다. 최근 마이크로소프트사의 Windows Automotive 버전이 출시될 정도로 차량 텔레매틱스 장치의 개발과 보급에 관련된 많은 관심들이 모아지고 있다[2]. 텔레매틱스 시스템에서의 서비스는 차량의 본질적인 이동성 때문에 위치기반 서비스(LBS; Location-Based Service)의 형태를 갖게 되며 이러한 서비스는 각 장치들로 하여금 필수적으로 GPS (Global Positioning System) 수신기와 같은 위치결정 장치를 포함하도록 한다. 또 통신을 위해 언제 어디서나 연결성이 있는 무선 통신 인터페이스를 포함하여야 하는데 대한민국에서는 CDMA (Code Division Multiple Access) 전화망이 무선 캐리어가 되며 이는 차량내 텔레매틱스 장치와 중앙의 텔레매틱스 서버와 제한된 대역폭이긴 하지만 데이터 통신을 지원하는 기반이 된다[3].

이상과 같은 장치들을 기반으로 차량의 위치 추적, 사고 정보 전파, 위치기반 운전정보들이 운전자나 사용자들에게 전달될 수 있다[4]. 위 서비스들 중에 차량 위치 트래킹은 가장 기본이 되는 서비스로서 차량의 현재 위치나 상태, 운전 형태들을 중앙에서 수집하여 관리자들이 쉽게 파악할 수 있도록 하고 향후 이에 따른 계획을 수립할 수 있도록 한다[5]. 대상 차량은 렌터카, 영업용 택시, 화물차 등 다양할 수 있으며 특히 많은 수의 택시를 운영하는 콜 택시 회사의 경우는 택시들의 위치를 파악함으로써 승객의 콜 요청에 대해 가장 근거리 택시의 배차, 현재 교통 상태를 고려한 최적 경로 추천 등의 서비스를 제공하여 시간 및 유류 비용의 절감 등 다양한 영업 이득을 가져올 수 있다[6].

위치 트래킹 기능을 가능하게 하려면 우선적으로

각 차량들은 주기적으로 자신의 위치를 중앙 서버에게 보고하여야 하는데 여기서 위치는 GPS 수신기로부터 전달된 경도와 위도로 나타내어진다. 보고 주기가 짧을수록 더 정밀한 위치 트래킹과 차량의 운행이력을 파악할 수 있으나 차량의 대수가 증가함에 따라 그 저장 비용과 통신 비용이 막대하게 초래된다. 일반적으로 통신 비용은 통신회사와의 협상에 의해 월정액으로 결정될 수 있으며 주어진 대역폭 내에서 데이터 교환을 극대화하여 위치를 수집한다. 이 수집된 데이터는 트래킹뿐만 아니라 택시의 운행에 대한 다양한 분석에도 사용될 수 있어서 승객들의 이동 거리, 이동 시간 패턴 분석, 배차 시간, 배차시 차량의 이동 거리 등 다양한 정보들을 제공할 수 있다. 이 정보들은 콜택시 회사들로 하여금 현재 택시의 운영에 대한 의사결정 자료를 제공하여 새로운 배차 방식의 발굴이나 장소별 시간별 탑승 거리 및 시간에 대한 자료를 제공할 수 있다[7].

본 논문은 차량 위치의 이동 보고에 대한 온라인 및 오프라인 분석을 위하여 택시들로부터 수집되는 정보를 처리하는 프레임워크를 제시하고 택시의 상태 정보에 따라 천이하는 스테이트 다이어그램을 정의한 후 이를 기반으로 택시의 운행에 관련된 정보들을 추출하고자 한다. 제주 택시 텔레매틱스 사업을 통해 실제로 수집된 택시들의 운행 정보들을 처리하였으며 다양한 분석 데이터를 산출하여 제안하는 방식의 효율성을 보인다.

본 논문의 구성은 다음과 같다. 1장에서 전반적인 주제에 대해 소개한 후 2장에서는 제주 택시 텔레매틱스 사업을 기반으로 한 배경과 관련연구들을 설명한다. 3장에서는 차량 텔레매틱스에서 보고되는 데이터를 온라인 혹은 오프라인으로 처리하는 프레임워크를 제시하고 4장에서는 운행정보를 추출하기 위한 스테이트 다이어그램을 보이고 이에 의한 분석 결과를 제시한다. 끝으로 5장에서는 본 논문을 요약하고 추후 연구 과제를 소개한다.

## 2. 배경 및 관련연구

### 2.1 제주도에서의 텔레매틱스 사업

제주텔레매틱스 시범도시사업은 제주에서 처음으로 시행된 시범사업으로, 렌터카에 텔레매틱스 단말기를 부착하여 관광객들이 텔레매틱스 서비스를 이용할 수 있도록 한다[8]. 텔레매틱스 서비스를 전국적인 대국민 서비스로 확대 발전시키기 위한 모델을 구현하고, 텔레매틱스 서비스 체험의 기회를 확대하여, 서비스 수요 창출 및 산업 대중화를 도모하며, 대국민 안전 및 복지 정보화 사회를 실현함을 목적으로 수행되었다.

이 사업은 2004년 9월부터 2006년 7월까지 1,2차 사업이 완료되었으며 서비스 센터시스템을 구축하는 「텔레매틱스 서비스 센터 구축」, 차량 내 단말기에서 서비스가 가능한 체계를 구축하는 「텔레매틱스 서비스 개발」, 그리고 「텔레매틱스 단말기 보급」 등의 세부사업으로 구성되어 있다. 또한, 제주도의 특성을 잘 반영하고, 텔레매틱스 산업 및 관련 산업 활성화를 선도할 수 있는 시스템 및 시범도시를 구성하는데 있어서 필요한 요소 기술들을 정의하고 이들간의 상호 호환성을 극대화할 수 있도록 한다. 이 사업을 통하여 텔레매틱스 산업의 조기 활성화 및 국제 경쟁력을 확보하고 관련 산업간의 시너지 효과를 창출하며 제주도 지역의 발전을 도모 등의 효과와 가시화되기를 기대하고 있다.

제주 택시 텔레매틱스 사업은 택시에 텔레매틱스 단말기를 부착, 다양한 결제 방식과 자동배차 기능 제공 및 텔레매틱스 서비스를 제공하는 관광·생활 융합형 텔레매틱스 사업이다[9]. 전국적인 대국민 서비스로의 확대·발전을 위한 텔레매틱스 서비스 체험의 기회를 확대하고 수요증대를 통한 텔레매틱스 산업의 대중화 기틀을 마련하며 텔레매틱스 서비스를 택시에 접목하여 텔레매틱스 산업에 대한 인지도 향상 및 지능형 콜 서비스와 위치기반 광고 서비스 등 다양한 택시 수익모델을 창출 등을 목적으로 하고 있다.

이 사업은 2006년 5월부터 2007년 12월까지

1,2차 사업을 추진하였고 택시 텔레매틱스 서비스 콜센터 운영 시스템 및 택시 텔레매틱스 서비스 개발, 택시 단말기 보급 등의 세부사업으로 구성되어 있다. 또한 택시 텔레매틱스를 통하여 지역 내 생활형 텔레매틱스 기반을 구축하고 택시의 특수성에 적합한 텔레매틱스 프로토타입의 개발과 아울러 타코미터 기반 요금체계 및 택시 서비스기반 모델의 개발에 의한 지역 정보통신 및 소프트웨어 기반 조성 및 글로벌 시장 확대를 사업목표로 하고 있다. 이 사업을 통하여 택시사업자의 수익증대와 체계적인 도심 내 교통흐름 관리, 수익증대를 통한 신규 고용창출 및 브랜드화를 통한 대국민 서비스로의 확대 등의 효과를 기대하고 있다.

### 2.2 텔레매틱스 서비스 및 데이터 분석 시스템

제주 택시 텔레매틱스 서비스를 고도화하기 위한 연구가 수행된 바 있으며 이 기존 연구에서는 네트워크 거리에 기반한 택시의 배차, 차량의 방향을 고려한 경로 배정, 다중목적지 플래닝 등의 기법 등이 제시되었다[3]. 네트워크 거리에 기반한 택시의 배차는 현재 택시들의 위치가 파악이 되어 있는 경우, 승객의 호출지점으로 가장 빨리 도착할 수 있는 택시를 결정하여 배차하는 기법이다. 이 과정에서 승객의 호출지점에 가장 가까운 노드를 시작점으로 설정하고 이 지점부터 순차적으로 다익스트라 알고리즘을 수행한다[10]. 너비 우선 탐색 방식인 다익스트라 알고리즘을 변경하여 각 스펠마다 택시의 위치와 바운드된 노드에 도달했는지 확인하도록 한다. 도달한 경우는 빈 택시를 찾게 된 것으로서 이 택시를 배차용 택시로 설정한다. 차량의 방향을 고려한 경로 배정 방식은 먼저 GPS에 의해 결정된 현재 차량의 진행방향과 각 도로 링크의 방향성을 고려하여 차량의 링크 내에서의 진행 방향을 결정한다. 이 진행방향에 따라 차량이 도달할 다음 노드를 결정하게 되며 이 노드가 유턴이 가능한 경우는 이 노드로부터 A\* 알고리즘을 수행하여 경로를 결정한다[11]. 만약 유턴이 불가능한 경우라면 방금 차량이 지나온 링크의 비용을 무한대로 설정하여

A\* 알고리즘을 수행한다.

또 제주도 지역의 택시 영업의 특이한 형태로 일일대차가 있어서 관광객이 정해진 시간 동안 여러 목적지를 방문할 수 있도록 하는데 이 경우 다중 목적지에 대한 효과적인 방문 순서 결정은 영업의 효율화에 기여할 수 있다. 이 서비스는 리눅스 클러스터 상에서 A\* 경로설정 알고리즘, Lin-Kernighan 과 같은 TSP (Traveling Salesman Problem) 솔버를 적용하여 구현되었다. 결국 택시 텔레매틱스 시스템은 새로운 텔레매틱스 서비스를 구현하고 시험하고 적용하는데 있어서 유용한 환경을 제공한다고 할 수 있다.

### 2.3 관련연구

본 관련연구에서는 데이터 수집에 관련된 연구, 택시 배차에 관련된 연구, 히스토리 데이터 분석에 대한 연구 등을 소개한다.

데이터의 수집에 관련된 연구로서 이와 더불어 Microsoft 사의 J. Krumm 등은 갈만 필터 방식을 기반으로 도로의 통행 속도를 추정하고 POI (Point Of Interest)를 추출하는 방식을 제시하였으며 웹 기반 브라우저에 그 결과를 포함시키도록 하였다[13]. 이 연구를 위하여 Microsoft Multi-Person Location Survey (MSMLS) 데이터가 사용되었으며 이 데이터에는 GPS 히스토리 데이터들이 포함된다[14]. 타임스탬프까지 추가된 GPS 경위도 레코드를 수집하기 위해 40개의 GPS 수신기가 사용되었으며 시애틀 지역의 자원자 차량에 설치되어 각 차량마다 2주 동안 차량 위치 수집을 수행하였다. 결국 102 개의 차량 이동 이력을 조사하였으며 운전자의 평소 운전 습관을 변경하지 않도록 하였다. 이 데이터는 수집 주기는 짧을 수 있지만 그 양이나 내용적으로 볼 때 제주 택시 텔레매틱스 시스템에서 수집된 데이터가 많은 잠재력을 갖고 있음을 보이고 있다.

AVLDS (Automatic Vehicle Location Dispatch System)은 싱가포르에서 운영되고 있는 시스템으로서 GPS, 컴퓨터, 무선 통신 기술의 결

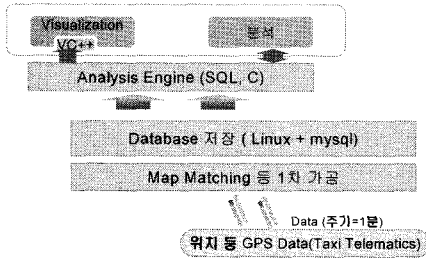
합에 의해 승객에게 가장 가까운 택시를 배차할 수 있도록 한다. 특히 승객이 택시 배차를 요청하는 방법을 다변화하여 전화를 이용하거나 인터넷을 통해 배차를 요청할 수 있음은 물론 시내 곳곳에 설치된 공중전화의 CabLink Hot Button을 이용하여 택시를 요청할 수 있다. 이 시스템은 택시 배차 시스템중 가장 모범적으로 알려져 있으며 승객의 만족도를 극대화한다[6]. 이와 유사한 트래킹 및 배차 시스템으로 인도에서 운영 중인 트래킹 기반 예약 시스템이 포함된다[7].

히스토리 분석에 관련된 연구로서 Lachesis 프로젝트에서는 차량의 이동데이터를 기반으로 차량이 stay와 destination을 추출해내는 확률적인 모델을 제시하였다[12]. stay는 차량이 머무는 곳으로 정의되어 시작시간과 끝시간으로 표현되어 시간 스케일에 따라 결정된다. 반면 destination은 stay들이 그루핑된 것으로서 유사한 반경내에 위치한 stay들을 clustering한다. 이 모델을 기반으로 특정 시간대에서 이동객체, 즉 차량의 움직임을 마코프 체인, 확률적 체인 등의 기법을 사용하여 모델링하였다.

### 3. 시스템의 구축

각 차량으로부터 위치 보고 데이터를 수집하여 이를 저장하고 처리하는 과정은 <그림 1>과 같다. 현재 최대 200 대의 택시가 1분마다 자신의 위치를 보고하고 있으며 각각의 레코드는 GPS 수신기로부터 입수된 타임스탬프, 경도, 위도, 속도, 방향 등의 기본 자료 이외에 택시 ID, 택시의 상태를 나타내는 필드 등을 포함하고 있다. 이러한 레코드를 받아 맵 매칭 프로그램은 이 좌표값에 해당하는 링크를 검색하여 link 필드에 추가한다. 또 링크 상에서 어느 정도 비율에 있는지를 계산하여 pos 필드에 저장한다. pos 필드값에 의해 링크의 양끝 쪽에 가까우면 맵 매칭에서 오류가 발생할 가능성이 높으며 또 양 끝에 가까운 위치에서 정지하고 있을 때에는 신호대기일 가능성이 높다. 결국 각 레코드가 가지는 정보는 <표 1>에서 보는 바와 같이

요약될 수 있으며 각 레코드는 택시 ID와 타임 스탬프를 기본 키(primary key)로 갖는다.

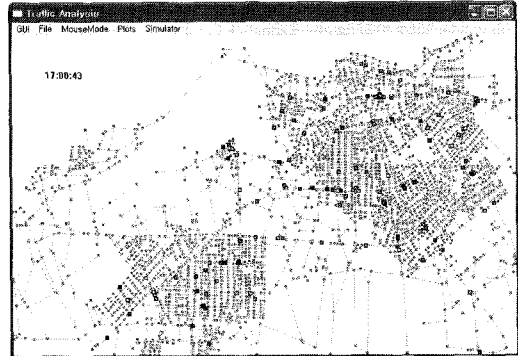


<그림 1> 레코드 처리 프레임워크의 개관

<표 2> 레코드들의 필드 구성

필드	타입	설명
timestamp	datetime	시간
tid	char(6)	택시 Id
x	double(11,10)	경도
y	double(11,10)	위도
dir	double(3,2)	방향
speed	double(3,2)	속도
status	int(2)	상태
link	int(6)	맵매칭 링크
pos	double(3,2)	링크상의 위치 비율

이러한 테이블을 바탕으로 택시들의 위치를 플로팅하는 인터페이스가 구현되었으며 히스토리 데이터 파일의 가시화(Visualization)인 경우에는 각 레코드들이 시간 순서대로 각 택시의 위치를 플로팅하며 온라인 플로팅인 경우에는 레코드 보고를 받아 실시간으로 화면에 표시할 수 있다. 이 위치의 플로터 화면은 <그림 2>와 같다. 이 그림은 제주시의 시가지 지역과 제주국제공항 부분을 확대한 것으로서 각 도로는 그래프 형태로 양 끝점만을 표시하고 있으며 지도 정보에 관련된 웨이프 파일이나 데이터베이스 연산을 피하도록 한다.

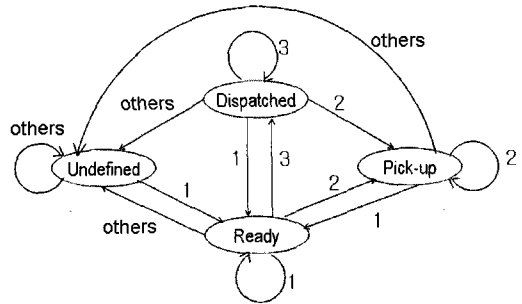


<그림 2> 택시들의 위치 히스토리 데이터에 대한 가시화 화면

#### 4. 자료의 분석

##### 4.1 레코드의 분석 구조

각 레코드는 택시의 순간적인 위치들을 나타내며 특히 택시들의 상태는 status 필드에 의해 얻어질 수 있다. 그러나 이 status 필드는 택시 운행 중에 승객의 탑승과 하차 등 자동적으로 설정되는 경우가 있고 운전자가 수동으로 조작을 해주어야 하는 경우가 있다. 수동 조작을 요하는 경우 운전자가 실수로 상태 변경을 안 한 경우는 전혀 의미없는 분석이 실행될 수도 있다. 본 논문에서 처리하고자 하는 레코드들의 시간적 흐름은 <그림 3>의 상태로 표현될 수 있다.



1: Empty 2: Passenger on D: Dispatched 4: Rest

<그림 3> 레코드들의 시간 흐름 상태도

각 레코드들은 택시의 상태를 나타내는 status 필드에 1, 2, 3, 4의 값을 포함하고 있다. 여기서 1은 택시가 승객을 태우지 않은 상태, 2는 승객을 태운 상태, 3은 배차를 받고 승객에게 가고 있는 상태, 4는 택시가 쉬고 있는 상태를 나타낸다. 이를 기반으로 <그림 3>에서 보는 바와 같이 각 택시들에 대해 Undefined, Ready, Dispatched, Pick-up 등 네 개의 상태를 정의하며 해당 택시 레코드의 status 필드에 의해 천이하도록 한다.

장치의 리셋이나 영업의 시작, 특정 상태에서 과다하게 머무는 경우 타임아웃이 되어 처음의 Undefined 상태에서 시작하게 된다. 이후 가장 기본적으로 택시는 승객을 태우기 전의 Ready 상태에 있어야 하며 status 필드에 1을 포함하는 레코드가 보고되는 즉시 Ready로 천이한다. Ready 상태에서 status가 2인 레코드가 보고되면 승객을 태운 것을 알 수 있으며 Pick-up 상태로 가게 되고 승객이 하차시 다시 Ready 상태로 오게 된다. Ready 상태에서 배차를 받으면 Dispatched 상태로 천이한 후 승객을 태우면 Pick-up 상태로 천이하는데 여기서 주의할 것은 특히 현재 택시 텔레매틱스 시스템에서는 배차를 받았다가 승객을 태우지 못한 경우도 발생한다. 이 경우 운전자가 다시 수동으로 상태를 바꾸어야 status 필드가 1로 바뀌어 Dispatched에서 Ready 상태로 하는데 운전자가 이를 잇는 경우 레코드가 계속해서 수집된다. 이에 대한 오차를 줄이기 위하여 타임아웃을 두어 일정 기간 Dispatched 상태에 머물면 Undefined 상태에서 시작하도록 한다. 또 택시가 2시간 동안 레코드를 보고하지 않으면 영업을 종료된 것으로 간주하며 추후 다시 레코드가 보고되면 Undefined 상태에서 출발한다.

## 4.2 배차 처리 정보의 조회

택시가 관제 센터로부터 배차를 받은 후 승객을 탑승할 때까지의 시간은 승객에게는 콜 요청 후 택시를 기다리는 시간이 되어 이 시간이 짧을수록 승객의 만족도는 올라간다고 할 수 있다[7]. 배차 시

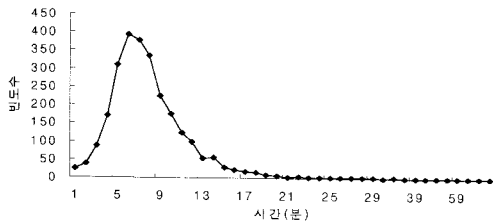
나리오에서 택시의 상태는 Ready→Dispatched→Pick-up의 천이를 갖는다[6]. 이때 택시의 상태가 Dispatched 상태에서 머무는 시간을 측정하면 배차 후 승객 탑승까지의 시간을 추산할 수 있다. 그러나 실제 데이터에서 실험한 바에 의하면 배차 취소나 승객의 미탑승 등으로 승객 탑승에 실패한 경우 택시 기사가 택시의 상태를 수동으로 Dispatched인 상태에서 Ready로 바꾸지 않으며 Dispatched인 상태에서 비정상적으로 오래 머무르게 된다. 이때 택시의 상태는 계속해서 배차 상태로 남아있게 된다. 이후 추가적인 배차를 받든 그냥 받은 승객을 태우면 그때서야 Pick-up로 바뀌게 된다. 따라서 상태가 Dispatched인 상태를 측정하면 위와 같은 문제점을 내포하고 있다. 극단적으로 실제 데이터를 분석해보면 배차상태에서 해당일의 영업을 종료하고 다음날 첫 승객을 태울 때까지 Dispatched인 상태로 보고되는 택시도 있었다.

<그림 3>의 상태도에서 Dispatched 상태에 차량의 상태가 있는 시간을 측정하여 dispatch라는 테이블을 <표 2>와 같이 생성하고 관련 레코드들을 저장하였다. 물론 택시들의 보고 간격이 1분이므로 분 단위의 오차를 포함할 수 있다. 이 테이블의 데이터를 기반으로 하여 빈도수를 플롯하면 <그림 4>와 같다. 그래프에서 보는 바와 같이 대부분의 배차 후 탑승까지의 시간은 13분 이내에 분포하고 있으며 이 시간들의 평균은 8.055분이다. 단, 20분 이상의 시간이 소요된 비정상적인 상태를 제외하면 7.670분으로 감소한다. 비정상적인 경우 93분 만에 승객을 태운 경우가 발생하기도 하였다. 이는 배차받은 후 승객을 태우지 못하고 운전자가 상태를 바꾸지 않은 상태에서 운행하다가 우연히 승객을 태운 경우로 판단된다.

배차 후 승객을 태울 때까지의 시간과 더불어 택시가 배차 후에 이동한 거리를 추정할 수가 있다. 이는 배차 상태에서 각 레코드들에 포함된 경도와 위도 좌표를 기반으로 계산할 수 있는데 배차 알고리즘의 효율성을 측정하는데 유용한 자료가 된다. 그러나 택시로부터의 좌표들이 1분마다 보고가 되기 때문에 실제로 이동한 경로를 분석하기는 불가

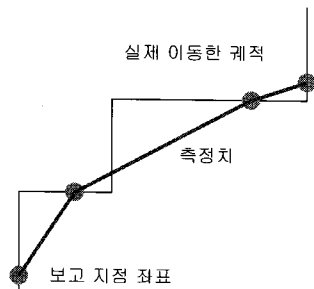
<표 2> Dispatch 테이블의 구조

필드	타입	설명
tid	char(6)	택시 Id
accepted	datetime	배차받은 시간
dine	datetime	배차후 승객을 태운 시간
distance	double(3,2)	배차후 이동 거리
accx	double(11,10)	배차받은 위치 경도
accy	double(11,10)	배차받은 위치 위도
x	double(11,10)	승객탑승 위치 경도
y	double(11,10)	승객탑승 위치 위도
completed	int (8)	분으로 환산된 Dispatch 시간



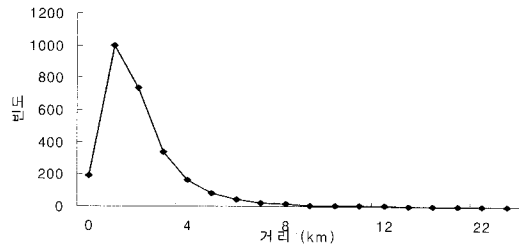
<그림 4> 승객 대기시간의 분포

능하며 따라서 거리에 대한 정확한 계산도 불가능하다. 단 <그림 5>에서 보는 바와 같이 거리를 추정할 수 있다. 여기서 경도는 1초당 31 m, 위도는 1초당 25 m의 보정을 사용하였다. 물론 두 좌표간에 최단 경로를 사용한다면 좀더 정확해질 수는 있겠지만 이 경우도 정확성을 완벽히 보장하지는 못한다.



<그림 5> 거리 추정의 오차 요인

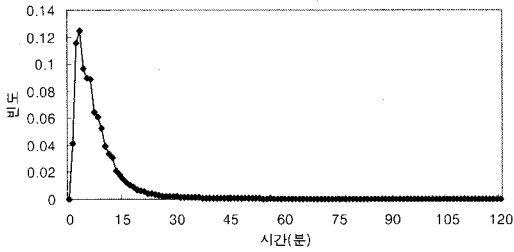
이를 바탕으로 하여 배차 후 이동한 거리의 추정치를 빈도수로 나타내면 <그림 6>과 같다. 배차 후 이동거리는 대부분 10 km 내에 분포하고 있으며 그림에서 보는 바와 같이 배차 후 택시의 평균 이동 거리는 2.055 km 이며 비정상적인 경우를 10km 이상 이동한 경우로 추정한다면 평균 이동 거리는 1.989 km 정도로 감소한다. 이동거리 역시 추후 텔레매틱스 시스템의 업그레이드 시 성능의 판단 기준으로 사용될 수 있다.



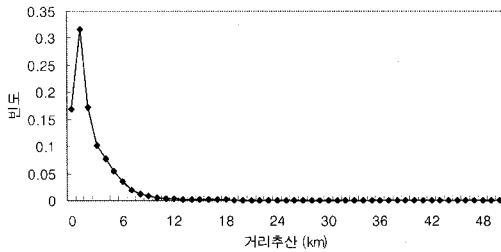
<그림 6> 배차 후 승객 탑승까지 차량이 이동한 거리 분포

### 4.3 택시 운행기록의 검색

택시들의 운행 기록은 승객 탑승에 대한 자료를 제시해주며 탑승 시간과 거리에 대한 시간적 지리적 분포를 제시할 수 있어서 택시 운전사들의 경험에 의지하는 승객 대기 시간과 위치를 좀더 체계적으로 계획할 수 있게 한다. 택시 운행 기록의 검색을 위해 <그림 3>의 상태도에서 택시가 Pick-up 상태에 머무르는 시간과 거리를 4.2의 방식과 같이 측정하였으며 그 결과는 <그림 7>과 <그림 8>에 각각 도시되어 있다. 승객의 탑승시간은 한나절 혹은 하루 대차의 경우 2시간 이상으로 계산되는 경우가 있었으며 2 시간 이내의 확률 분포를 보면 15분 이내의 탑승이 대부분을 차지한다. 거리도 10 km 이내에서 대부분 위치하고 있다. 더욱이 오차 포함 2km 이내를 이동하는 단거리 이동이 40 % 가까이 분포한다. 물론 여기서 2km란 실제 이동 거리가 아니고 <그림 5>에서 보는 추정거리이므로 실제 거리는 이보다 는 길 것으로 예상된다.



<그림 7> 승객 탑승 시간의 확률 분포



<그림 8> 승객 탑승 거리추산의 확률 분포

## 5. 결론

본 논문은 제주 택시 텔레매틱스 사업에서 수집된 차량의 히스토리 데이터를 온라인 혹은 오프라인으로 저장 및 처리하기 위한 프레임워크를 제시하였으며 저장된 데이터를 가시화하고 효율적으로 분석하기 위한 스테이트 다이어그램을 정의하였다. 실제 데이터에 있어서 처리가 어려운 부분을 파악한 후 이에 기반하여 택시의 운행과 배차에 대한 분석을 수행하였으며 상태도 상에서 배차 후 승객 탑승까지의 시나리오 분석에 의해 이 상태에 머무르는 시간과 이동거리를 계산하였다. 분석 결과는 배차 후 약 7분, 약 1.9km 이내에 승객에게 도달할 수 있는 것을 보이고 있다. 또 상태도 분석에 의해 각 택시들이 승객을 태우고 이동한 시간과 거리 등에 대한 분석도 수행하였다.

결국 본 논문에서 제시한 레코드 저장 방식과 분석 방식은 막대한 양의 데이터가 매일 쌓이고 있는 제주 택시 텔레매틱스 사업에 있어서 보다 부가가치가 높은 분석을 수행하는데 기반이 될 것으로 예

상되어 과거의 배차 기록을 토대로 새로운 배차 알고리즘의 개발 및 테스트가 가능할 것으로 예상된다. 네트워크 거리, 후보 택시들의 방향성 배차시의 공평성 등 다양한 요소들을 고려할 수 있으며 배차 시간, 배차 거리, 택시간 공평성 등 다양한 성능 인자에 대한 사전검증이 가능하다. 또 시간대별 장소별 승객의 탑승에 대한 분석은 운전자들로 하여금 승객의 탑승을 미리 예측하여 대기하거나 주행 계획을 세울 수 있도록 할 것이다. 이러한 기능에서 발전하여 향후 교통 정보 예측[15], 혼잡상황 회피, 궤적에 대한 클러스터링[16] 등의 다양한 응용을 개발하고 테스트하는데 유용하게 활용이 될 것이다.

## 참고문헌

1. <http://en.wikipedia.org/wiki/Telematics>
2. <http://www.microsoft.com/windowsautomotive/default.msp>
3. Lee, J., Park, G., Kim, H., Yang, Y., Kim, P., and Kim, S., "A telematics service system based on the Linux cluster," Lecture Notes in Computer Science, Vol. 4490, 2007, pp. 660-667.
4. Green, P., "Driver Distraction, Telematics Design, and Workload Managers-Safety Issues and Solutions," Proceedings of the 2004 International Congress on Transportation Electronics, 2004, pp. 165-180.
5. Imamura, M., Kobayashi, K., and Watanabe, K., "Real-time positioning by fusing differential-GPS and local vehicle sensors," SICE Annual Conference, 2003.
6. Liao, Z., "Real-time taxi dispatching using global positioning systems," Communication of the ACM, Vol. 46, 2003, pp. 81-83.
7. Kiruthivasan, S., Deepakumar, C., and Althaf, S., "Decision Support System For



- Call Taxi Navigation Using GIS-GPS Integration,” MAP India, 2006.
8. SK텔레콤컨소시엄, 『2차년도 제주텔레매틱스 시범도시 구축사업 완료보고서』, 2006년 8월
  9. Lee, J., Kang, E., and Park, G., “Design and implementation of a tour planning system for telematics users,” ESUC (ICCSA), Lecture Notes in Computer Science, Vol. 4707, 2007, pp. 179-189.
  10. Horowitz, E., Sanhi, S., and Rajasekaran, S., Computer Algorithms, W. H. Freeman & Co., 1997.
  11. Goldberg, A., Kaplan, H., and Werneck, R., “Reach for A\*: Efficient point-to-point shortest path algorithms,” MSR-TR-2005-132. Microsoft Research, 2005.
  12. Hariharan, R. Toyama, K, “Project Lachesis: Parsing and modeling location histories,” 3rd International Conference on GIScience, 2004
  13. Krumm, J., “The geographic context browser,” International Workshop on Exploiting Context Histories in Smart Environments, 2005.
  14. Krumm, J., “The Microsoft multiperson location survey,” MSR-TR-2005-103, Microsoft Research, August 2005,
  15. Lee, S., Lee, B., and Yang, Y., “Estimation of Link Speed Using Pattern Classification of GPS Probe Car Data,” Proc. International Conference on Computational Science and its Applications, 2006, pp. 495-504.
  16. Kim, S., Won, J., Kim, J., Shin, M., Lee, J., and Kim, H., “Path prediction of moving objects on road networks through analyzing past trajectories,” KES, Lecture Notes in Computer Science, Vol. 4693, 2007, pp. 379-389.

**이정훈**

1988년 서울대학교 컴퓨터공학과(공학사)  
 1990년 서울대학교 컴퓨터공학과(공학석사)  
 1996년 서울대학교 컴퓨터공학과(공학박사)  
 1990~1992, 1996년 대우통신 종합연구소 전송연구실(선임연구원)  
 1997~현재 제주대학교 전산통계학과 교수  
 관심분야 : 실시간 통신, 텔레매틱스, 지리정보

**권상철**

1986년 서울대학교 지리교육과(문학사)  
 1989년 Pennsylvania State University 지리학과(이학석사)  
 1995년 Ohio State University 지리학과(문학박사)  
 1995년~현재 제주대학교 사범대학 지리교육전공 교수  
 관심분야 : 지리자료분석, 환경교육, 도시지리