

공간 기반의 개별 차량 대용량 정보 맵핑에 관한 연구

A Study on Map Mapping of Individual Vehicle Big Data Based on Space

정 규 수*

* 주저자 및 교신저자 : 한국건설기술연구원 미래스마트건설연구본부 연구위원

Kyusoo Chong*

* Dept. of Future & Smart Construction Research, Korea Institute of Civil Engineering and Building Technology

† Corresponding author : Kyusoo Chong, ksc@kict.re.kr

Vol.20 No.5(2021)

October, 2021
pp.75~82

pISSN 1738-0774
eISSN 2384-1729
<https://doi.org/10.12815/kits.2021.20.5.75>

Received 26 August 2021
Revised 7 September 2021
Accepted 7 September 2021

© 2021. The Korea Institute of Intelligent Transport Systems. All rights reserved.

요 약

국내 2020년 기준 교통사고 건수는 약 23만 건으로, 고속도로는 비반복적 정체와 높은 주행 속도로 인해 다른 도로 대비 교통사고 발생 건수당 사망자수가 2배 이상으로 나타나고 있다. 고속도로의 교통정보는 도로의 중심선을 이용한 노드, 링크를 기준으로 제공하고 있으나 세분화된 속도정보를 주지 못하고 있다. 최근 자율주행차 뿐만 아니라 일반 차량에서도 장애물 모니터링, 위치 측정을 위한 차량용 센서 장착이 일반화되고 있어, 대용량 위치 기반 데이터를 이용한 분석은 처리속도에 따라 실시간 서비스가 가능하다. 본 연구는 대용량 위치기반 개별 차량 정보의 분석을 위한 공간 기반의 맵핑 방법을 제시하였다. 경위도 각각 2개로 분할하는 4진법 기준의 분할 방법을 적용하여 개별 차량의 공간 코드를 생성하여 지오코딩 하는 방법으로 처리 속도를 대폭 증가 시켰다. 공간이 세분됨에 따라 평균속도는 유사하였으나 속도의 표준편차는 점차 감소하였으며 9회 분할 이후는 그 감소 폭이 작아 졌다.

핵심어 : 차량 빅데이터, 링크, 지오코딩, 공간 분할, 속도 표준편차

ABSTRACT

The number of traffic accidents is about 230,000, and due to non-recurring congestion and high driving speed, the number of deaths per traffic accident on freeways is more than twice compared to other roads. Currently, traffic information is provided based on nodes and links using the centerline of the road, but it does not provide detailed speed information. Recently, installing sensors for vehicles to monitor obstacles and measure location is becoming common not only for autonomous vehicles but also for ordinary vehicles as well. The analysis using large-capacity location-based data from such sensors enables real time service according to processing speed. This study presents an mapping method for individual vehicle data analysis based on space. The processing speed of large-capacity data was increased by using method which applied a quaternary notation basis partition method that splits into two directions of longitude and latitude respectively. As the space partition was processed, the average speed was similar, but the speed standard deviation gradually decreased, and decrease range became smaller after 9th partition.

Key words : Vehicle Big Data, Link, Mapping, Space Partition, Speed Standard Deviation

I. 서론

최근 자율주행차량에는 차량용 LiDAR(Light Detection and Ranging), Radar(Radio Detection and Ranging), GNSS(Global Navigation Satellite System), 카메라 등의 센서를 이용하여 주변의 상황을 분석하여 자율주행에 활용하고 있다. 여기서 발생하는 데이터는 도로 위의 형상이나 현상을 실시간으로 구현할 수 있는 것으로써, 교통정보의 분석에 중요한 정보로 활용될 수 있다. 또한, 최근 MaaS¹⁾ 개념이 도입되면서 자동차 등 이동체와 통신 및 서비스가 융합되어 도로의 모든 이동체는 센서, 검지기의 역할을 하고 빅데이터를 생성하게 된다. 현재 ITS 등의 운영을 위한 각종 센서들은 설치 및 유지보수가 어렵고 비용이 높아, 모빌리티 개념의 움직이는 센서를 이용한다면 훨씬 정확하고, 신속하며 부가 생산 데이터의 활용에 따른 비용이 낮거나 없는 교통정보의 생성이 가능할 것이다. 자율주행차뿐만 아니라 일반 차량에서도 장애물 모니터링, 위치의 측정을 위한 차량용 Radar, GNSS, 카메라 등의 센서 장착이 일반화되고 있어, 여기서 발생하는 빅데이터도 앞서 언급한 분야에 활용될 수 있다.

대용량 데이터의 분석을 위해서는 좌표 기반의 맵 매칭을 위한 다양한 방법이 사용되고 있다. 맵 매칭은 지점과 지점을 매칭하는 Point-to-point 알고리즘(Bernstein and Kornhauser, 1998), 지점과 선을 연결하는 Point-to-curve 알고리즘(Bernstein and Kornhauser, 1998, White et al., 2000), 선과 선을 매칭하는 Curve-to-curve 알고리즘(Bernstein and Kornhauser, 1998, White et al., 2000)이 있다. 그 외 토폴로지 분석을 기반으로 한 맵 매칭 알고리즘(Quddus et al., 2007)과 확률적 맵 일치 방법인 Probabilistic 알고리즘(Quddus et al., 2007)이 있다. 이러한 맵 매칭 방법은 각각의 알고리즘 별 오차 발생 등의 단점이 있고 개별 차량의 데이터와 도로 선형 간의 공간분석이 필수적이라는 점에서 대용량 데이터의 위치 기반 처리에는 적합하지 않다.

개별 위치 좌표를 이용한 지도 내 자신의 위치를 빠르게 색인하기 위한 방법으로 공간 인덱스 기법이 있다. 전통적인 공간 인덱스 접근 방법은 크게 균등 분할, 비중첩 분할, 중첩 분할로 분류되며 각각 그리드 기반, R-Tree 기반, QuadTree 기반 인덱스에 해당된다(Singh and Bawa, 2017). 최근, 위치 기반 색인을 위한 지오해시(geohash)는 공간 빅데이터에서 공간정보 검색 속도를 높이기 위해 전 세계 경위도 두 값의 좌표를 한 개의 숫자로 변환하는 기법으로, 32진법으로 표현되는 지오 코딩방법의 제시(Niemeyer, 2000), 64진법, Hilbert 등 다양한 변형을 사용(Vukovic, 2016)하고 있다. 하지만, 이러한 색인 기법은 개별 위치를 그리드 등 일정 범위로 나타내 주는 것으로 그동안 지역 내 교통 패턴 분석 외에 선형 교통 분석에는 사용하지 않았다.

본 연구에서는 개별 차량에서 발생하는 위치, 속도, 방위각, 가감속 등의 빅데이터를 선형 교통 분석에 활용하기 위해 위치 좌표 기반 색인 방법 중 하나인 지오해시(geohash)를 기반으로 공간 기반의 맵핑 방법을 제시하고자 한다. 개별 차량의 속성정보를 기반으로 하여 위치 좌표의 공간 맵핑을 이용한 교통 분석은 기존 연구와 큰 차별성이 있으며, 이동객체의 공간 매칭 시 색인 기법은 기존 맵 매칭 기법 대비 월등한 속도를 나타내, 대용량 데이터의 실시간 맵핑이 가능할 것으로 판단된다.

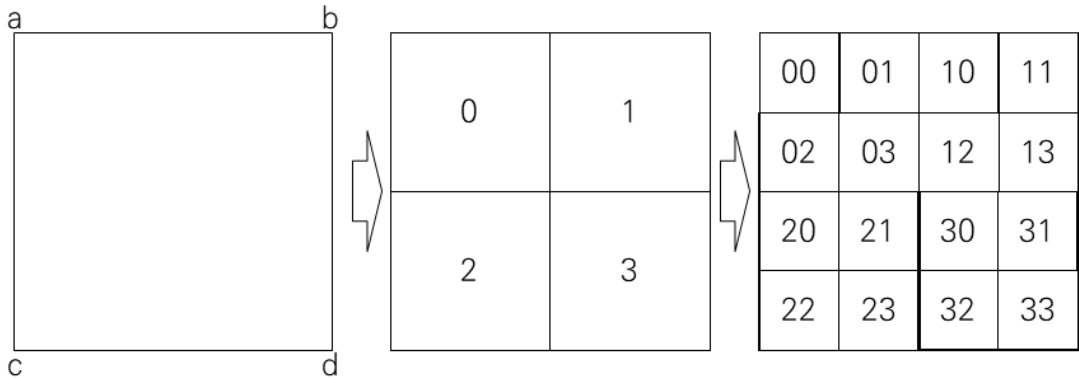
II. 공간 기반 맵핑 방법

지오해시는 전 세계를 기준으로 경도·위도를 분할하여 하나의 코드화 하는 방법으로, 우리나라의 경도·

1) Mobility as a Service, 서비스로서의 모빌리티, 즉 스마트폰 하나로 카셰어링뿐만 아니라 철도, 택시, 자전거, 전동스쿠터, 주차장 등에 이르기까지 포괄적인 이동 서비스를 제공하는 개념을 말한다.

위도 범위 내에서 공간 분할 및 코드 부여, 맵핑 방법을 제시하고자 한다. 우리나라의 본토 범위는 경도 126°와 130°사이, 위도 34°와 38°사이로 하고, 공간 분할은 개별 차량의 GPS 좌표 정보를 TM(Transverse meractor) 좌표로 변환 후 공간에 지오코딩 하기 위해 하나의 코드로 생성하였다. 우리나라 영역에 맞는 경위도 범위에서 소규모 지역의 대용량 데이터 맵핑에는 과도한 계산을 수행할 수 있으므로, 지엽적 좌표 적용 및 공간 세분화를 위한 분할 규모의 축소가 필요하다. 따라서 위도, 경도 각각에 대해 포함되는 공간을 0 또는 1의 이진수로 부여하여 나열하고 다시 4진수로 변경하여 1개의 코드를 생성하였다. 분할 단계의 차이를 최소화하기 위해 4진법 사용하여 우리나라 형태에 맞는 정사각형에 가까운 공간 분할을 적용하고, 위경도 방향으로 동일하게 2개의 공간으로 분할하였다. <Fig. 1>은 4진법 기반 공간분할 및 코드부여 과정을 보여주며 단계별 방법은 다음과 같다.

[1단계] 공간을 1회에 4개로 분할한다고 할 때, <Fig. 1>에서 코드 2 공간의 경우 위도는 아래쪽인 '1', 경도는 왼쪽인 '0'이며, 이진수 '10'은 4진수 '2'가 된다.



<Fig. 1> The method of space partition based on quaternary notation

[2단계] 주어진 GNSS 좌표를 분할하여 코드를 생성한 후 자릿수당 코드를 비교하여 사용한다.

[3단계] 좌표의 코드가 123123123123 일 경우 초기 공간은 1이며, 5회 분할 공간의 코드는 12312로, 동일 코드를 가진 좌표는 동일 공간으로 계산한다.

이와 같이 개별 차량의 위치정보를 이용한 공간의 분할은 분할된 공간에서 다시 4개의 공간으로 분할하는 방법으로 점진적으로 진행되며 각각의 공간의 4개 모서리 좌표를 기준으로 다음 식과 같이 계산할 수 있다.

- $(a+c)/2 > Y; (a+b)/2 > X \rightarrow 0$ (1)
- $(a+c)/2 > Y; (a+b)/2 < X \rightarrow 1$ (2)
- $(a+c)/2 < Y; (a+b)/2 > X \rightarrow 2$ (3)
- $(a+c)/2 < Y; (a+b)/2 < X \rightarrow 3$ (4)

여기서 a, b, c, d = 각각 공간의 좌측상단, 우측상단, 좌측하단, 우측하단 모서리의 좌표
 Y, X = 위도좌표, 경도좌표

이상의 분할 방법에 따라 공간분할을 하면 분할 횟수와 동일한 자리수의 공간코드가 생성되는데, 1회 분할할 경우 CH0, CH1, CH2, CH3의 4개 공간이 생성되고 2회 분할하면 CH00, CH01, CH02, CH03 등 16개의 공간이 생성된다. 12회 분할하면 CH000000000000, CH00000000000 1, CH000000000002, CH000000000003 등 16,777,216개의 공간이 생성된다.

Ⅲ. 개별 차량 정보 맵핑 결과 및 분석

1. 데이터 전처리

본 연구는 도시 지역 등 구간의 길이가 짧고 제한속도가 상대적으로 낮으며 단속류에 따른 교통 흐름의 변화가 있는 지역보다 구간의 길이가 상대적으로 길고 제한속도가 높아 연속류의 분·합류부, 사고, 비반복적 정체가 발생하는 고속도로를 대상으로 한다. 또한, 차량 각각의 위치정보와 속도정보 등 데이터의 수집이 가능하고 고속도로 상의 데이터 수집이 충분히 가능하며, 지정차로 등 주행 차로의 영향이 적은 데이터가 필요하다. 택시 DTG(Digital Tacho Graph) 데이터의 경우 도시지역에 주로 분포하며, 승객의 탑승, 하차, 대기의 영향에 의해 임의의 추정차가 발생할 수 있고, 버스 DTG의 경우도 도시지역에 주로 분포하며 도시 간 버스의 경우 버스전용차로 등에 따른 주행 차로의 영향에 따라 주변 교통류와 다른 특성으로 나타날 수 있어, 화물차 DTG 자료를 활용하였다. 화물차의 경우도 주행속도가 상대적으로 낮으며, 주행 차로도 우측인 경우가 많으나 고속도로를 포함한 전국 도로에 고르게 분포하여 비교적 정확한 데이터를 수집할 수 있다. 또한, 화물차는 DTG를 의무적으로 설치고 정보를 공개하게 되어 있어 데이터의 수집이 가능하나, 실시간 수집과 주행 속도, 차량 특성 등이 일반 차량을 대표할 수 없는 한계점이 있다.

DTG 데이터에는 주행 노선정보가 없기 때문에 고속도로 노선별 데이터의 추출이 필요하였다. 이를 위해 노드링크는 지능형교통체계관리시스템의 표준노드링크를 사용하였다. 추출한 데이터는 고속도로 차로 당 3.5m의 버퍼를 두었으며 졸음쉘터, 휴게소 등 주행차를 고려하여 속도 데이터가 0kph인 경우와 제한속도를 20% 초과하는 경우는 제외하였다. DTG 데이터 중 위치정보는 GPS 경위도 좌표로, 평면 직각 좌표로 변환하여 사용할 필요가 있다. 경위도 좌표계는 WGS84이며 이를 중부원점(N38, E127), TM 좌표계로 변환하여 사용하였다. 실제 GPS의 오차는 차로 기준을 넘을 수 있어 데이터 집계에 한계가 있지만, 대용량 데이터의 평균값을 사용하므로 발생 가능한 GPS 오차는 고려하지 않았다. 또한, 차량의 상하행 방향성을 분리하기 위해 방위각이 정북을 0도로 했을 때 90도에서 270도 사이는 하행, 반대는 상행으로 분리하였고 본 연구에서는 하행 데이터만 적용하였다.

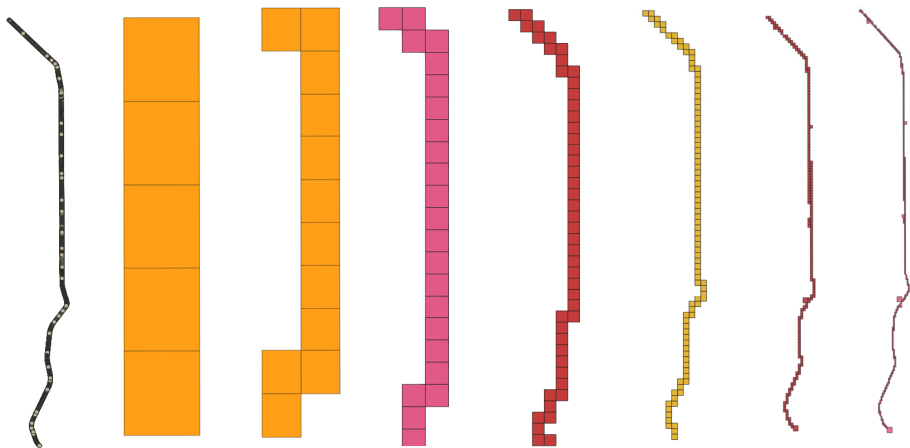
2. 적용 및 분석

본 연구는 개별 차량 대용량 정보를 빠르게 맵핑하여 속도의 변화구간, 차량의 운행 형태, 개별 운전 특성 등 교통정보 분석에 활용하기 위한 것으로, 다양한 목적으로 활용하기 위해 공간 분리의 정밀도를 다르게 할 수 있다. 따라서 개별 차량의 속도 정보를 대상으로 차량간 속도 편차가 적고 지오코딩의 효율성을 높이기 위한 공간 분할을 적용하였다. 경부고속도로의 경기도 지역의 DTG 데이터 중 GPS 좌표를 이용하여 지오코딩을 하였다. 개별 차량의 DTG 데이터 속성을 모두 포함하여 각 차량의 1초 간격으로 9시30분부터 5분간 누

적 데이터를 적용하였다. <Fig. 2>의 좌측부터 순서대로 DTG 데이터, 6회, 7회, 8회, 9회, 10회, 11회, 12회 분할한 공간의 결과이며, <Table 1>은 공간 분할 단계에 따라 공간의 개수, 동일 공간으로 지오크딩 된 개별 차량의 대수, 각 공간 내의 평균속도 합을 전체 공간 개수로 나눈 평균속도 및 각 공간 내 표준편차 합을 전체 공간 개수로 나눈 평균표준편차를 보여준다.

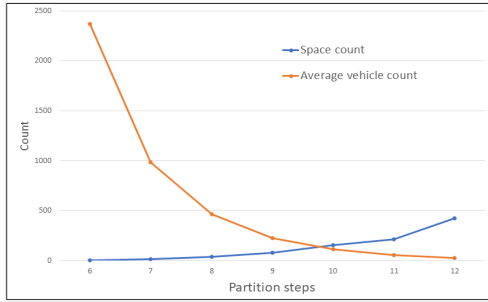
<Table 1> The attribution of geocoding spaces

Partition steps	The number of space	The number of average vehicle	Average speed (km/h)	Average standard deviation (km/h)
6	5	2368	72.79	24.52
7	12	987	73.60	20.61
8	38	465	78.44	14.04
9	79	222	79.16	12.14
10	156	113	79.56	10.86
11	212	56	75.98	9.81
12	423	28	76.33	9.15

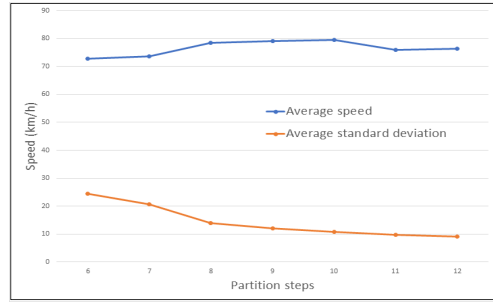


<Fig. 2> The result of geocoding spaces

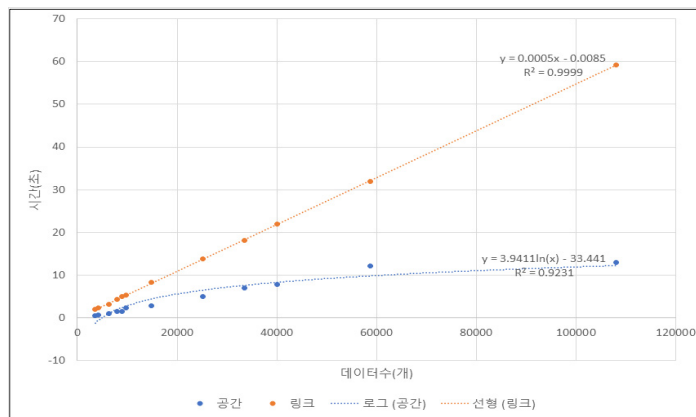
<Fig. 2>에서와 같이, 공간 분할의 횟수에 따라 공간 개수는 증가하였으며 동일 공간에 지오크딩 된 차량의 대수는 감소하게 되는데, 효율적인 지오크딩을 위해서는 적정 차량 대수와 공간의 개수를 고려하여 결정하여야 한다. 교통량이 적은 경우 즉 원활한 교통 흐름일 경우 공간의 개수를 적게 할 수 있으며, 교통량이 많은 경우는 불안정 교통류인지 안정 교통류인지에 따라 공간의 개수를 결정할 수 있을 것이다. <Fig. 3>은 공간 분할 단계에 따른 공간당 차량의 평균 대수 및 총 공간 개수를 나타내며, <Fig. 4>는 공간 분할 단계에 따른 공간당 평균속도와 표준편차의 평균값을 나타낸다. 평균속도는 공간 분할 횟수와 관계없이 일정하게 나타났으며 동일 공간의 속도 표준편차는 공간 분할 횟수에 따라 감소하였는데, 9회 분할 후에는 표준편차의 차이가 크지 않았다.



<Fig. 3> The comparison space with vehicle



<Fig. 4> The comparison speed with deviation



<Fig. 5> The comparison of processing speed between space and link

<Table 2> The comparison of processing speed between space and link

The number of data	6,281	3,565	4,217	7,995	8,981	9,721	14,838	25,163	33,523	108,003	58,686
Space	1.01	0.51	0.68	1.50	1.44	2.29	2.85	5.03	7.06	12.90	12.10
Link	3.24	1.99	2.37	4.33	5.06	5.28	8.29	13.76	18.17	59.25	31.93

본 연구는 도로 즉 선형 기반으로 하는 차량 정보를 공간 기반으로 맵핑하는 것으로서, 대용량 데이터의 신속한 처리가 가능하다. 데이터 처리 신속도를 검증하기 위해 point-to-curve 방식인 개별 차량의 위치 데이터를 이용해 가장 근접한 링크에 맵핑하는 방법과 비교하였다. DTG 데이터의 개수와 맵핑 노선을 달리하여 적용하였으며, 본 연구 방법은 10만개 데이터일 때 약 4.7배 빠른 처리가 가능하였다. <Fig. 5> 및 <Table 2>는 처리 데이터의 개수에 따른 본 연구 방법과 링크 매칭 방법에 의한 처리 시간을 비교한 것이다. 링크 기준의 공간처리 방법은 선형 증가하는 반면, 공간 기반 처리 방법은 데이터량 증가 대비 처리속도의 증가폭은 작아져, 대용량 개별 차량 정보의 효율적 맵핑이 가능함을 검증하였다. 링크 기반 처리 방법은 개별 차량의 위치 각각에 대한 링크와의 공간분석이 필요한 것이 가장 큰 원인이라 할 수 있다. 이에 반하여 본 연구 방

법은 개별 차량의 위치를 공간 코드로 계산하여 공간분석 없이 맵핑하는 것으로, 차량의 속도 등 속성 기준의 분석을 위한 최적의 공간 분할이라는 것과 함께 본 연구의 특징 및 장점이라 할 수 있다. 하지만, 공간 분리의 특성 상 분리되는 경계와 가까이 있는 경우는 유사한 특성을 가지더라도 다른 공간으로 분리될 수 있고, 경계선 상에 있는 경우는 공간 맵핑이 어렵다는 한계점이 있다.

IV. 결 론

국내 교통사고는 고속도로에서 발생 건수 당 사망자수가 2배 이상으로 나타나고 있는데, 이는 비반복적 정체와 높은 주행속도로 인한 것이며, 실제 차량의 정보를 이용한 구간별 세분화 된 비반복적 정체 정보의 제공이 반드시 필요하다. 하지만, 도로의 중심선을 이용한 노드, 링크를 기준으로 제공하고 있으나 세분화된 속도정보를 주지 못하고 있다. 본 연구에서는 개별 차량에서 발생하는 위치, 속도, 방위각, 가감속 등의 빅데이터를 선형 교통 분석에 활용하기 위해 위치 좌표 기반 색인 방법 중 하나인 지오해시(geohash)를 기반으로 공간 기반의 맵핑 방법을 제시하였다. 경위도 각각 2개로 분할하는 4진법 기준의 분할 방법을 적용하고, 개별 차량의 위치 좌표를 하나의 공간 코드로 생성하여 맵핑 방법을 제시하였다. 경기도 지역의 경부고속도로 상의 화물차 DTG 자료를 이용하였으며, 개별 차량의 공간 코드를 이용한 분할 단계별 맵핑 결과를 제시하였다. 공간 분할 단계가 반복됨에 따라 공간의 개수는 증가하였으며 동일 공간에 맵핑된 차량의 대수는 감소하였다. 또한 평균속도는 공간분할 단계에 관계없이 일정하게 나타났으며 동일공간의 표준편차는 감소하였는데, 8단계까지 급격히 감소하다가 9단계부터는 감소의 차이가 크지 않아, 개별 차량의 속도 정보 분석을 위한 공간 기반 맵핑은 9단계가 적합함을 알 수 있었다. 또한, 대용량 데이터의 맵핑 시 본 연구 방법인 공간 기반 맵핑이 매우 효율적임을 검증하였다.

본 연구에서는 개별 차량의 위치좌표를 노드, 링크로 구성된 도로망에 맵핑하는 기존 방법과 달리, 개별 차량의 위치정보만을 이용한 공간 기반 맵핑을 하였으며 이는 기존 교통정보 맵핑과 큰 차이가 있다. 본 연구의 공간 기반 맵핑은 교통량이 적은 경우 즉 원활한 교통 흐름일 경우 공간의 개수를 적게 할 수 있으며, 교통량이 많은 경우는 불안정 교통류인지 안정 교통류인지에 따라 공간의 개수를 결정할 수 있을 것이다. 향후, 본 연구 결과를 활용하여 적정 차량 대수와 공간의 개수의 상관관계를 이용한 분석 대상에 따른 공간 분할단계 결정하는 효율적인 지오코딩을 위한 공간 분할 단계 연구가 필요하다.

REFERENCES

- Bernstein D. and Kornhauser A.(1998), An Introduction to Map Matching for Personal Navigation Assistants, *Technical Report*, New Jersey TIDE Center.
- Chen S., Ooi B. C., Tan K. L. and Nascimento M.(2008), “ST2B-tree: A Self-tunable Spatio-temporal b+-tree Index for Moving Objects,” *ACM SIGMOD International Conference on Management of Data*, pp.29-42.
- Niemeyer G., <https://en.wikipedia.org/wiki/Geohash>, Wikipedia, 2020.12.10.
- Quddus M. A., Ochieng W. Y. and Noland R. B.(2007), “Integrity of Map-matching Algorithms,” *Transportation Research Part C: Emerging Technologies*, vol. 14, pp.283-302.

- Singh H. and Bawa S.(2017), “A Survey of Traditional and MapReduce Based Spatial Query Processing Approaches,” *ACMSIGMOD Record*, vol. 46, no. 2, pp.18-29.
- Vukovic T.(2016), *Hilbert-Geohash-Hashing Geographical Point Data Using the Hilbert Space-Filling Curve*, Master's Thesis, NTNU, Trondheim, Norway.
- White C. E., Bernstein D. and Kornhauser A. L.(2000), “Some Map Matching Algorithms for Personal Navigation Assistants,” *Transportation Research Part C: Emerging Technologies*, vol. 8, no. 1-6, pp.91-108.