한국ITS학회논문지 J. Korea Inst. Intell. Transp. Syst.

기종점통행량(O/D) 기반의 고속도로 통행실적 산정 방법론 연구

Methodology for Estimating Highway Traffic Performance Based on Origin/Destination Traffic Volume

이 호 원* · 홍 정 열** · 최 윤 혁***

* 주저자 : 계명대학교 도시계획 및 교통공학과 석사과정
** 교신저자 : 계명대학교 교통공학전공 조교수
*** 공저자 : 한국도로공사 도로교통연구원 책임연구원

Howon Lee* · Jungyeol Hong** · Yoonhyuk Choi***

* Dept. of Transportation Eng., Kiemyung Univ.
** Dept. of Transportation Eng., Kiemyung Univ.
*** Dept. of Transportation Research, Korea Expressway Corporation Research Institute

† Corresponding author : Jungyeol Hong, jyhong9868@kmu.ac.kr

Vol. 23 No.2(2024)
April, 2024
pp.119~131

pISSN 1738-0774
eISSN 2384-1729
https://doi.org/10.12815/kits.2024.23.2.119

Received 10 January 2024
Revised 24 January 2024
Accepted 26 February 2024

ⓒ 2024. The Korea Institute of Intelligent Transport Systems. All rights reserved.

요 약

고속도로의 효율적 운영과 지속가능한 이동환경 제공을 위해 정확한 통행실적의 파악은 필수적이다. 그러나 인프라 및 기술적 제약, 추정을 통한 기존 연구방법의 한계, 통합 빅데이터 활용의 제약 등의 이유로 즉각적이고 정확한 고속도로 통행실적 산정에 어려움이 있다. 이에 본 연구는 자동요금징수시스템, 단거리전용통신 등으로부터 수집된 실시간 빅데이터를 활용하여 개별 차량의 고속도로 기종점통행량(Origin-Destination: OD) 및 주행거리를 분석하고 이를 기반으로 고속도로 통행실적을 산정하는 방법론의 틀을 제시하였다. 특히 데이터상 하나의 통행임에도 불구하고 분할된 통행으로 나타나는 데이터상 오류를 통행궤적 및 주행특성 진단을 통해 올바르게 연결함으로써 보다 신뢰성 있는 고속도로 통행 OD를 구축하고자 하였다. 연구 결과 개별 차량의 분할된 OD 통행이 20분 내 연속되는 경우 통행연결이 필요하며 통행연결, 주행거리 산정, 누락교통량 비율 보정, 비연계 구간 연계 과정을 거쳐 전국 고속도로의 일평균 통행실적은 248,624천대·km/일로 도출되었다. 이는 도로업무편람에서 제시하는 248,166천대·km/일과 비교 시 약 458천대·km/일이 높았다. 본 연구의 결과는 기존의 조사과정에서 누락된 통행실적의 보완가능성을 보여준다.

핵심어 : 기종점통행량, 주행거리, 통행실적, 빅데이터, 통행연결

ABSTRACT

Understanding accurate traffic performance is crucial for ensuring efficient highway operation and providing a sustainable mobility environment. On the other hand, an immediate and precise estimation of highway traffic performance faces challenges because of infrastructure and technological constraints, data processing complexities, and limitations in using integrated big data. This paper introduces a framework for estimating traffic performance by analyzing real-time data sourced from toll collection systems and dedicated short-range communications used on highways. In particular, this study addresses the data errors arising from segmented information in data, influencing the

individual travel trajectories of vehicles and establishing a more reliable Origin–Destination (OD) framework. The study revealed the necessity of trip linkage for accurate estimations when consecutive segments of individual vehicle travel within the OD occur within a 20-minute window. By linking these trip ODs, the daily average highway traffic performance for South Korea was estimated to be248,624 thousand vehicle kilometers per day. This value shows an increase of approximately 458 thousand vehicle kilometers per day compared to the 248,166 thousand vehicle kilometers per day reported in the highway operations manual. This outcome highlights the potential for supplementing previously omitted traffic performance data through the methodology proposed in this study.

Key words : Origin-destination trip, Travel distance, Traffic performance, Bigdata, Traffic linkage

Ⅰ. 서 론

1. 연구 배경 및 목적

고속도로의 출·도착지별(Origin-Destination: OD) 통행실적(대·km/일)은 고속도로를 이용하는 모든 차량들의 일별, OD별 통행량과 통행거리로 도출할 수 있으며, 고속도로 및 교통 시스템의 계획과 개발, 교통안전, 교통 혼잡관리, 운영 및 정책 평가 등 다양한 분야의 의사결정에 지표로써 활용할 수 있다.

그러나 고속도로를 주행하는 개별 차량들은 원래 계획한 출발지(O)와 도착지(D) 사이를 고속도로 이외에 국도, 지방도 등 등급이 다른 여러 종류의 도로를 혼용하여 통행할 수 있어 개별 차량은 하나의 통행이 아닌 다수의 통행으로 OD가 분할되어 실제 OD와 다른 결과를 나타내는 문제가 발생한다. 예를 들어 인천 남동IC에서 강릉IC를 최초 출발지(O)와 최종 도착지(D)로 계획한 차량이 영동고속도로를 통행하다가 여주IC에서 진출하여 국도 42번으로 통행하고, 다시 영동고속도로 문막IC를 진입하여 최종적으로 강릉IC에서 진출하는 경우 이 차량의 OD는 인천 남동IC에서 여주IC 간, 문막IC에서 강릉IC 간으로 분할되어 두 개의 개별 OD로 집계된다. 이러한 OD 집계방식은 고속도로 통행 OD 구축 결과의 신뢰성을 저하시킬 뿐만 아니라 고속도로 영업소 OD기반의 여객 및 화물자동차 수요추정에 정확성을 저하시키는 요인이 되며, 중간에 누락된 통행거리로 인해 고속도로의 통행실적이 과소 산정되는 문제가 발생한다.

따라서 본 연구에서는 고속도로와 타 등급의 도로를 병행하여 주행 시 집계되는 통행 OD가 분할되는 한계를 극복하기 위하여 TCS(Toll Collecting System)와 HI-PASS(Highway Payment and Administration System) 데이터, DSRC(Dedicated short-range communications) 차량 통행궤적 데이터 기반의 고도화된 통행실적 산정 알고리즘 구축을 주요 목적으로 하였다. TCS와 HI-PASS 데이터는 고속도로 요금징수 시스템으로부터 수집되며 개별 차량별, 시간대별, 일별, 월별 등 다양한 시공간 프레임으로 분석 가능하다는 장점이 있다. 또한 DSRC 데이터는 근거리 전용 통신으로 노변기지국과 차량에 탑재된 OBU(ON Board Unit) 간 통신하여 차량의 궤적을 추적하므로 OD 간의 통행거리 추출이 가능하다. 따라서 TCS 및 HI-PASS 데이터로부터 고속도로 OD 구축 및 통행량을 산정하고, DSRC를 통한 OD 간 통행거리를 도출함으로써 고속도로 차종별 통행실적(대·km/일) 산정의 자동화를 도모하였으며, 이러한 시스템을 고도화하기 위하여 분할된 개별 차량의 OD별 통행을 연결하는 방법론을 제시하고자 하였다. 또한 TCS, HI-PASS 데이터의 한계성을 보완하기 위해 TMS(Traffic Monitoring System) 통계연보 데이터를 활용하여 연계되지 않는 고속도로 구간[1]의 통행실적을 집계하여 전국 고속도로의 통행실적을 산정하였다.

1) 개방식 나들목형 영업소, 비연계 민자고속도로 및 무료구간 등

2. 기존 문헌 고찰

요금징수시스템으로부터 수집한 데이터를 활용하여 고속도로 통행량 및 OD를 산정하는 다양한 연구들이 기존에 수행되어왔다. Kim(2001)의 연구에서는 교통량 조사의 인력 및 비용 부담을 줄이기 위하여 TCS 데이터를 활용한 구간별 통행량 추정 방법을 제시하였다. 차종을 기존 5종에서 11종으로 재분류하여 차종별·구간별 통행량을 도출하였으며, 결과를 통하여 기존 고속도로 정기교통량 조사에 대한 검증이 가능함을 제시하였다. Kim(2007)은 여객통행수요 분석에서 수집된 표본조사의 한계점을 극복하기 위하여 직접 영향권에 해당하는 구간의 사회경제적 통계자료와 관광지에 인접한 고속도로 TCS 데이터를 수집·분석하였으며, 성장한계 모형을 이용한 여객통행수요 추정 모형을 제시하였다. Jung(2012)은 수납시스템(TCS) 데이터와 차량검지기(VDS) 데이터 처리 과정을 분석하였으며 TCS와 HI-PASS 데이터 통합 및 자료처리 절차와 알고리즘 개선을 통하여 수납데이터의 오류와 이상치를 제거하는 등 집계된 교통량의 정확성을 확보할 수 있도록 TCS와 HI-PASS 데이터 기반의 개선된 전처리 알고리즘을 설계하였다. Shin et al.(2013)의 연구에서는 차종구분 영상조사 자료를 활용하여 도시유형 및 권역별 차종비율을 도출하였다. Jun(2013)은 TCS 데이터를 활용한 기존의 연구들은 주로 시계열 분석을 통하여 장래 고속도로 교통량을 예측하였다면 국가교통DB센터(KTDB : Korea Transport DataBase)에서 제공하는 OD 통행량, 네트워크와 함께 TCS 데이터를 활용하여 고속도로의 장래 교통수요를 예측하였으며, TCS를 활용한 통행 배정 결과가 더 정확하다는 결론을 주장하였다. Han et al.(2017)은 DSRC 및 TCS 자료 분석을 통해 특정 구간에 대한 경로형 교통자료를 산출하였으며 상습 정체구간에 대한 분석 및 수요집중기간의 우회패턴에 대한 분석을 통하여 정체를 완화하고 도로네트워크를 효율적으로 활용할 수 있는 활용방안을 제안하였다. Kang et al.(2023)의 연구에서는 TCS 데이터로부터 고속도로 구간별 통행량을 집계하고, 저속 및 고속축중기(Low/High Weigh-In-Motion)로부터 화물적재원단위를 추정하여 고속도로 화물수송실적(tom·km/year)을 도출하는 프레임워크를 설계하였으며, Shan and LI(2008)은 통행 OD 데이터, 화물 OD 데이터 등 다양한 데이터를 기반으로 복합적인 OD 이슈를 분석하고 고속도로 화물차에 대한 OD 매트릭스를 구축하였다. Shi et al.(2016)은 기존의 교통 모델에 의존하는 연구와 다르게 베이징, 산시성, 안후이성의 실제 고속도로 톨게이트 데이터를 기반으로 고속도로망 취약성 분석을 수행하였으며 Fu et al.(2019)는 메모리 모듈, 예측 모듈, 기능 확장 모듈을 이용한 고속도로 OD 예측 신경망을 개발하여 새로운 OD 데이터 구축 및 관리 방안을 제시하였다.

이러한 기존 연구를 바탕으로 TCS 데이터를 활용하여 고속도로 구간별·차종별 통행실적 산정 및 결과의 신뢰성을 향상시키고자 노력해 왔다는 사실을 확인할 수 있다. 본 연구에서는 고속도로를 이용한 화물차들의 차종별 OD를 분석하여 영업소별·권역별·차종별 통행실적으로 도출하고, 교통 빅데이터들의 미시적 시공간 범위의 정보를 분석함으로써 통행이 분할될 수 있는 문제를 해결하여 통행실적 산정을 보다 고도화시키는데 기존 연구와 차별성이 있다.

Ⅱ. 데이터 수집

1. 데이터 수집

통행실적 산정 시 필요한 통행량을 집계하기 위하여 TCS 데이터와 HI-PASS 데이터의 고속도로 요금징수 데이터를 활용하였다. TCS는 톨게이트 영업소에 위치하여 도로 이용자로부터 통행료를 징수하는데 이용되

며 통행료, 차량정보, 요금소 정보 수집이 가능하다. HI-PASS는 차량이 고속도로를 통과할 때 통행료를 자동으로 수집하는 전자 통행료 수납시스템이다. TCS와 HI-PASS 데이터가 구분되어 집계되는 이유는 요금을 지불하는 방식의 차이로 인해 영업소의 요금수납기기 및 차로가 분류되어 있기 때문이다. 따라서 차량이 어떠한 요금수납방식의 차로를 통과하여 데이터가 수집되었는지에 대한 차이가 있으며, 이 두 데이터를 통합하여 영업소를 통과한 차량들의 통행량을 알 수 있다.

분석에 활용된 TCS와 HI-PASS 데이터의 시간적 범위는 2021년 10월 18일에서 2021년 10월 24일까지 도로법에 따른 정기조사 기간과 동일한 10월 셋째 주, 7일간의 데이터이다. TCS 및 HI-PASS 데이터는 개별 차량의 데이터가 모두 축적되는 형태로 데이터의 크기가 방대하여 가공하는데 어려움이 존재한다. 따라서 본 연구에서는 1주일 데이터를 수집하여 통행실적 구축 방법론을 제시하였으며, 향후 본 방법론에 기반한 자동화프로그램이 개발되면 1년간의 데이터를 이용한 통행실적 구축이 가능하다. 공간적 범위는 데이터 수집 범위가 한정적인 TCS, HI-PASS 뿐만 아닌 TMS 통계연보 등의 데이터를 연계하여 한국도로공사가 운영하는 356개의 영업소와 비 수탁운영 민자노선의 44개 영업소를 포함한 전국 고속도로이다. TCS 및 HI-PASS 데이터는 <Fig. 1>과 <Fig. 2>와 같이 차량 일련번호(암호화), 진입일시, 진입영업소, 진출일시, 진출영업소, 차종(5종, 12종, 16종)으로 구분된다.

통행실적 산정을 위하여 필요한 고속도로 OD 간 통행거리는 한국도로공사에서 제공한 OD 간 최단경로 데이터, DSRC 데이터, DsrcS 프로그램을 이용하였다. 이 중 OD 간 최단경로 데이터는 운전자가 일반적으로 최단경로를 선호한다는 가정하에서 적용되었으며 최근 운전자들이 고속도로 운행 시 티맵, 카카오맵, 내비게이션 등 최적 및 최단거리를 제공해주는 기능들을 이용한다는 것을 고려할 때 최단거리 데이터로 분석가능하다는 것에 타당성을 더한다. DSRC 데이터는 전용 단거리 통신 기술로, 차량 대 차량 간 통신을 통하여 차량의 개별 통행궤적을 수집한 데이터이다. DSRC 데이터를 OD화하여 OD별 실제 주행거리를 추출할 수 있는 DsrcS 프로그램(The Korea Expressway Corporation, 2022)을 이용하여 실제 주행거리를 추출하였으며 연계되지 않아 누락된 통행 데이터는 TMS 통계연보 데이터를 활용하여 집계하였다.

car_num	in_date/time	in_office	out_date/time	out_office	car_type		
					5	12	16
3031EAB1B031313631	20211018114900	soul	20211018115330	busan	1	1	1
3031EAB3A036363133	20211018204000	soul	20211018205039	busan	1	1	1
3031EAB5AC33313030	20211018110000	soul	20211018144841	busan	1	1	4
3031EB848831373537	20211018085800	soul	20211018091520	busan	1	1	1
3031EB848831373537	20211018121300	soul	20211018123059	busan	1	1	3
3031EB848831373537	20211018091600	soul	20211018114904	busan	1	1	3
3031EB848831373537	20211018123700	soul	20211018175902	busan	1	1	3
3031EB848832373736	20211018181600	soul	20211018185253	seoul	1	1	3
3031EB848832373736	20211018080900	soul	20211018084151	seoul	1	1	3
3031EB848837383231	20211018092900	soul	20211018094033	seoul	1	1	3
3031EB88432333933	20211018061300	soul	20211018070749	seoul	1	1	3
3031EB8BA431363530	20211018172500	soul	20211018181506	seoul	1	1	4
3031EB8D9433373334	20210000000000	soul	20211018083559	seoul	1	1	2
3031EB8D9433373334	20210000000000	soul	20211018000352	seoul	1	1	3
3031EB8D9433373334	20210000000000	soul	20211018165246	seoul	1	1	3
3031EB919038323234	20211018062400	soul	20211018085841	seoul	1	1	3
3031EB919038323234	20211018120500	soul	20211018151308	seoul	1	1	3
3031EB9DBC35343034	20211018091600	soul	20211018092456	seoul	1	1	3
3031EB9DBC35343034	20211018211100	soul	20211018212106	seoul	1	1	3

<Fig. 1> Data structure of TCS

car_num	in_date/time	in_office	out_date/time	out_office	car_type		
					5	12	16
3031EAB08033333435	20211018143026	seoul	20211018150218	sungname	1	NH	4
3031EAB08033333435	20211018073159	seoul	20211018074825	sungname	1	NH	4
3031EAB08037343532	20211018144319	seoul	20211018164023	sungname	1	1	4
3031EAB1B031373931	00000000000000	seoul	20211018155724	sungname	1	1	3
3031EAB1B031373931	20211018174457	seoul	20211018175526	sungname	1	1	3
3031EAB1B031373931	20211018160640	seoul	20211018161816	daegu	1	1	3
3031EAB1B031373931	00000000000000	seoul	20211018181218	daegu	1	1	3
3031EAB1B032343635	00000000000000	seoul	20211018074504	daegu	1	1	4
3031EAB1B035393031	20211018161303	seoul	20211018162152	daegu	1	1	3
3031EB848832343136	00000000000000	seoul	20211018073527	daegu	1	NH	3
3031EB848832343136	00000000000000	seoul	20211018074413	daegu	1	NH	3
3031EB848835333436	20211018111545	seoul	20211018113126	gwangju	1	1	3
3031EB88437393631	20211018190906	seoul	20211018191753	gwangju	1	NH	3
3031EB888438303436	00000000000000	seoul	20211018192057	gwangju	1	NH	4
3031EB888438303436	00000000000000	seoul	20211018091348	gwangju	1	1	4
3031EB888438303436	00000000000000	seoul	20211018090144	gwangju	1	1	4
3031EB88843830343 6	00000000000000	seoul	20211018193002	gwangju	1	1	4
3031EB8D94313637 38	00000000000000	seoul	20211018092728	gwangju	1	1	4
3031EB8D94313637 38	20211018094234	seoul	20211018095240	gwangju	1	1	4
3031EB8D94313637 38	00000000000000	seoul	20211018071004	gwangju	1	1	4

<Fig. 2> Data structure of HI-PASS

2. 기존 통행실적 검토

통행실적은 몇 대의 차량이 얼마만큼의 거리를 통행하였는가에 대한 지표이며, 여객 및 화물 수송실적을 산정하는데 필수적이다. 현재 고속도로를 포함하여 일반국도, 지방도에 대한 통행실적을 도로업무편람에서 공표하고 있다. 도로업무편람의 고속도로 통행실적은 매년 10월 셋째 주 목요일의 구간교통량(VDS), TCS 데이터를 대푯값으로 이용하여 고속도로 구간별 교통량 데이터를 구축하고, 고속도로 상시조사를 통해 월별,

요일별 보정계수를 산정 및 적용하여 교통량을 추정하는 방식을 지니고 있다. 본 연구에서 제시하는 통행실적 산정방법은 추정이 아니라 실시간 수집되는 실증 데이터를 기반으로 교통량 데이터를 집계함으로써 보다 현실적인 통행실적 산정이 가능하다. 실증 데이터를 기반으로 한 방법론으로 산정한 고속도로 통행실적과 도로업무편람에서 제시한 고속도로의 통행실적을 비교하였으며, <Table 1>과 같이 도로업무편람에서 제시한 2021년 고속도로 통행실적은 248,166(thousand·veh·km)이다.

<Table 1> Traffic performance in 2021 road work manual

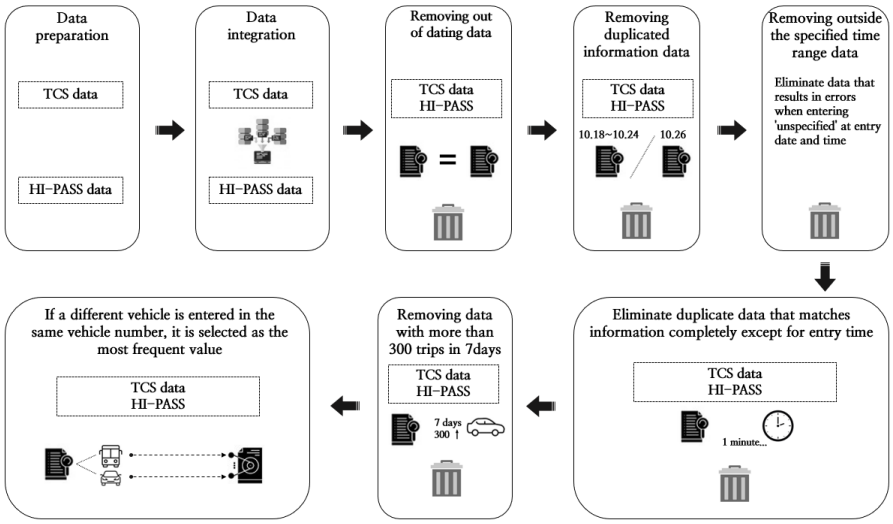
Year	Road type	Road length (km)	Average daily traffic volume (veh/day)	VKT (thousand veh·km)	VKT rate (%)	VKT/km of road (thousand·veh)
2021	Highway	4,866	51,004	248,166	48.9	51.0

Ⅲ. 방법론

1. 데이터 전처리

TCS 데이터와 HI-PASS 데이터를 전처리 과정에서 통합하였으며, 이것을 '수납데이터'라 정의하였다. 수납데이터는 국내 고속도로의 일주일간 통행 데이터(2021. 10. 18 – 2021. 10. 24)이며, 총 31,596,788(통행/7일)로 일일 통행량으로 환산 시, 약 4,513,827(통행/일)로 나타났다.

데이터 전처리를 위한 첫 번째 과정으로 차량별 진출입 통행 정보가 완전히 중복되는 데이터를 제거하는 과정을 진행하였다. 동일 차량이 동일 OD 및 시간에 통행한 데이터는 기기 오류로 인한 문제 데이터로 간주하여 제거하였으며, 중복제거를 통해 제거된 데이터는 총 24,464개의 통행이다. 두 번째로 진입 및 진출일시에 이상치 데이터를 제거하였다. 여기서 이상치 데이터란 지정된 시간적 범위를 벗어난 데이터를 의미한다. 단, 선정된 시간 범위 이전 시간에서 시작되어 선정 범위 내에 끝나는 경우, 선정된 범위 내에서 통행을 시작하여 선정된 시간 범위 이후에 통행이 끝나는 경우가 존재할 수 있으므로 데이터 가공 시 2021.10.17.에서 2021.10.25.를 시간적 범위로 확대 지정하였으며, 위 과정에서 총 11,392개의 데이터를 제거하였다. 세 번째, 진입일시와 진입영업소를 제외한 다른 모든 데이터가 동일한 경우에 대한 전처리를 진행하였다. 이는 한 차량이 출발 정보를 제외한 도착 일시, 도착 시간, 도착 장소 등이 모두 동일한 경우를 말한다. 이 경우는 진입영업소와 진입일시가 데이터 수집의 오류로 인해 누락되어 다른 통행 데이터로 분류된 것으로 의심할 수 있다. 네 번째로 차량 일련번호, 진출영업소, 진출일자가 동일하지만 진출시간이 상이한 경우, 진출시간을 기준으로 차이가 60초 미만인 통행은 제거하였다. 본 데이터의 경우 또한 한 차량이 동일 지점을 동일 시간에 여러 번 통행하였다는 중복데이터로 식별하였다. 다섯 번째로 고속도로를 통행한 개별 차량의 통행 수를 분석한 결과, 7일간 비정상적으로 많은 통행이 검지된 차량을 전처리하였다. 확인 결과 2대의 차량이 식별되었으며 7일간의 통행이 380번과 5,266번의 통행을 한 것으로 분석되었다. 이 경우의 데이터는 차량 일련번호를 암호화하는 과정에서 문제가 발생하여 생긴 오류 데이터로 간주하였다. 마지막으로 차량 일련번호 기준으로 구분된 차종에 문제가 있는지 확인하였으며 한 차량이 2개의 차종으로 분류되는 데이터가 식별되었다. 이는 데이터 기초 가공 및 암호화 과정에서 문제가 발생한 것으로 간주하고, 일련번호 기준으로 동일 차종 라벨링이 많은 데이터를 채택하였다. <Fig. 3>의 과정을 거쳐 데이터 전처리를 수행하였으며, 원시데이터 31,596,788개의 데이터 중 약 99%에 해당하는 총 31,273,647개의 통행 데이터가 구축되었으며 <Table 2>에서 확인가능하다.



<Fig. 3> Procedure of data preprocessing

<Table 2> Summary of the data preprocessing

Process	Number of observations
Raw data collection	31,596,788
Removal of duplicated observations	31,290,685
Removal of outlier	31,273,647
Final dataset	31,273,647

TCS 최단거리 데이터는 OD 간으로 제공되었으며 매칭되지 않은 일부 OD 간의 거리가 존재한다. 이러한 이유는 최단거리 데이터가 약 160,000개 가량의 모든 OD 조합에 대한 정보를 포함한 것이 아니므로 일부 OD에 대한 최단거리 정보가 누락되었기 때문이다. 또한 차량이 개방식 영업소나 무료구간 이용 등의 이유로 진입영업소를 알 수 없는 경우에도 OD 간 거리 데이터가 매칭될 수 없다. 이렇게 누락된 OD의 최단거리를 재산정하기 위하여 가중평균을 이용하였다. 가중평균 거리는 진출영업소 기준으로 구축하였으며, 산정 식은 Eq. (1)로 정의한다.

$$\frac{\sum_{OD=1}^{N} (\text{진입영업소}_{OD} \sim \text{진출영업소}_{OD} \text{ 간 최단거리}) \times \text{통행량}_{OD}}{\sum \text{진출영업소통행량}_{OD}} \quad \cdots\cdots (1)$$

통행실적 산정의 정확성을 높이기 위하여 DSRC 실제 주행거리 데이터를 이용하여 보정하였다. DSRC 데이터는 개별 차량 통행 경로에 대한 궤적정보가 축적된 형태이기 때문에 차량 궤적을 추적하여 OD의 형태로 만들어주는 DsrcS 프로그램을 활용하여 OD 간의 주행거리 데이터를 추출하였다. 구축된 OD 간 주행거리 데이터는 다양한 차량의 경로가 포함되므로 한 OD에 여러 개의 주행거리가 산정된다. 따라서 통행량을 이용한 가중치를 적용하여 OD별 주행거리의 대푯값을 선정하였다. DSRC 데이터를 이용한 OD별 실제 주행거리 산정 식은 Eq. (2)며 OD별 실제 주행거리를 산정하였다.

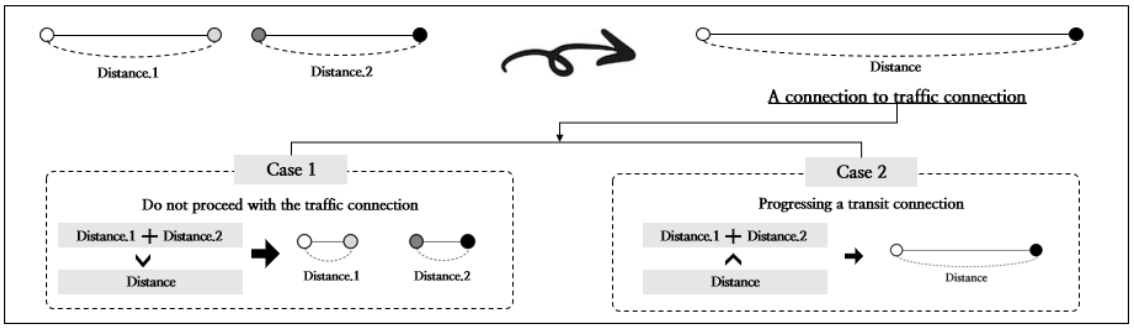
$$\frac{\text{주행거리}(total\ distance)_N \times \text{통행량}(count)_N}{\sum_{OD=1}^{N} \text{통행량}(count)_{OD}} \quad \cdots\cdots (2)$$

2. 통행연결 방법론 구축

데이터 가공을 통해 전처리가 완료된 수납데이터를 이용하여 통행을 연결하는 방법론을 구축하였다. 통행연결의 목적은 개별 차량이 통행함에 있어서 다양한 경로를 이용할 수 있는데, 이때 국도와 고속도로를 병행하는 경우, 지정 경로 이탈로 인한 방향 전환[2] 등을 목표로 영업소를 이용하는 이유로 하나의 분할된 통행을 실제 하나의 연결된 통행으로 재구축하기 위함이다. 통행연결 기준은 동일한 일련번호를 가진 차량을 기준으로 데이터를 분류하여 각 차량 통행별 진출시간과 진입시간의 차이 및 최단거리를 이용하였고 다음 두 가지 조건에 맞게 지정하여 통행을 연결하였다.

첫 번째 조건은 차량 일련번호별 진출시간과 진입시간을 기준으로 시간 차이를 비교하여 10분 간격으로 분석을 진행하도록 한다. 차량의 일련번호를 이용하였기 때문에 개별 차량별로 데이터를 구분할 수 있고 개별 차량의 진출시간에서 바로 이어지는 다음 통행의 진입시간을 빼면 몇 분 뒤에 다음 통행이 시작되었는지 확인할 수 있다.

두 번째 조건은 <Fig. 4>와 같이 통행연결된 OD의 최단거리와 연결되기 전 OD 간 거리의 합을 비교하여 통행연결된 OD의 거리가 긴 경우에 한하여 통행연결을 진행하는 것이다. 이 조건은 통행연결을 했을 시, 거리 데이터가 유실되는 경우를 보완하기 위한 과정이다. 예를 들어, 출발지와 목적지가 같은 경우(짧은 시간 내에 목적을 달성하고 귀가한 경우)나 동승자를 수송하는 경우(학원, 출장, 마중 등) 주행거리가 '0'이 되는 문제가 발생하게 된다. 위 조건을 통하여 짧은 시간 안에 목적을 달성하고 다시 본 위치로 돌아오는 등의 경우에 발생하는 거리 데이터 유실 문제점 보완이 완료되었다.



<Fig. 4> Criteria for measuring travel distances

3. 주행거리 매칭 및 통행실적 산정 방법론

일반적으로 운전자는 최단경로를 선호한다는 가정하에 OD 간 최단거리 데이터를 OD별 교통량 데이터와 매칭하여 통행실적 구축을 위한 기반 데이터를 구축하였다. 구축된 최단거리에 현실성을 반영하기 위해 DSRC 데이터를 이용하였으며 앞서 구축한 OD 간 실제 주행거리 데이터를 OD별 교통량 데이터에 추가로

2) 운전자 실수 또는 우회를 목적으로 한 경로 이탈을 의미함

매칭하고 최단거리 값과 실제 주행거리 값을 비교하였다.

최단거리 데이터는 영업소 간 이동하는 좌표를 이용하여 산정한 최소한의 통행거리이므로 DSRC를 이용한 실제 주행거리가 최단거리보다 짧다면 오류 데이터로 인식하고 최단거리를 이용하였으며 최단거리보다 실제 주행거리가 길 경우, 실제 주행거리로 값을 대체하여 보정을 진행하였다.

매칭이 완료된 OD 간 주행거리와 OD별 통행량을 이용하여 통행실적을 산정하였으며 산정하는 식은 Eq. (3)으로 정의한다.

$$\sum_{OD=1}^{N} (\text{주행거리}_{OD} \times \text{통행량}_{OD}) = \text{총 통행실적} \quad \cdots\cdots (3)$$

Ⅳ. 분석 결과

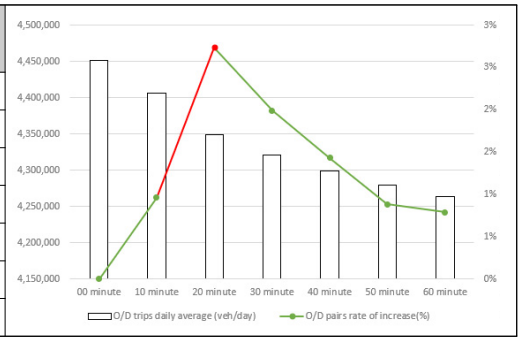
1. 통행연결 결과 분석

본 연구에서 제시한 기준을 바탕으로 통행연결 전과 후의 통행실적 결과를 각각 비교하였다. <Table 3>에서 제시되는 바와 같이 통행연결은 10분 간격으로 증가하여 최대 60분 이내에 동일 차량의 통행이 분할되어 있을 경우 하나의 통행으로 연결하였다.

<Table 3>의 00 minute이란 수납데이터의 차량별 ID를 기준으로 특정 차량이 두 개 이상의 통행으로 분할된 경우 첫 번째 통행의 진출시간과 두 번째 통행의 진입시간과의 차이를 의미한다. 이러한 경우에 통행을 연결하여 구축한 통행량 데이터는 대/7일 단위이며 7일 치 데이터를 일평균 통행량 데이터로 구축하기 위해 7로 나누었다. 이것을 'Average daily O/D trips'로 정의하였다. 분석 결과 통행을 연결하는 시간 범위가 증가할수록 OD의 수가 많아지는 반면, OD 통행량은 다소 감소하는 추세로 분석되었다. 해당 결과를 <Table 3>의 그래프로 시각화한 결과 분할된 통행이 20분 내에 있을 경우, 통행연결 시 기울기의 변곡점이 발생하는 것을 볼 수 있다. 이는 통행연결 후, 급격한 변화율이 발생하는 구간을 의미하며 이후에는 변화율이 점차 안정된다는 것을 확인할 수 있다. 이를 통하여 통행연결 데이터는 20분 이후로 통행량 및 OD 수의 변화가 안정화된다.

<Table 3> Results of linking segmented O/D trips by 10-minute units

Time	O/D pairs (number)	Rate of increase(%)	Average daily O/D trips (veh/day)
00 minute	104,164	-	4,450,851
10 minute	105,170	0.96%	4,405,791
20 minute	108,128	2.74%	4,349,051
30 minute	110,329	1.99%	4,320,781
40 minute	111,927	1.43%	4,298,395
50 minute	112,921	0.88%	4,279,277
60 minute	113,815	0.79%	4,263,019

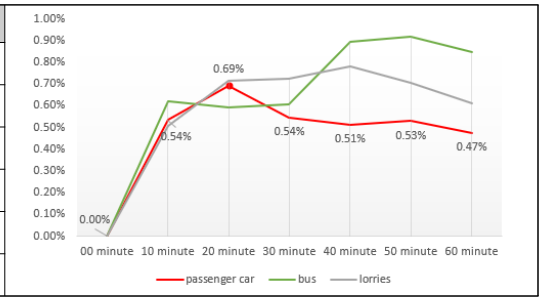


또한 <Table 4>의 그래프 확인 결과, 통행이 분할된 차량들 중 영업소 진출 후 20-30분 사이에 다시 고속도로 영업소를 진입하는 경우가 가장 많았다. '화물차량 상·하역 시간 단축을 위한 일괄 하역장비 기술개발

최종보고서'에서 제시한 평균 하역시간과 도시 내부 진출입 시간을 고려해 보았을 때, 화물차가 경유지에서 정차하는 시간은 약 26분이다. 경유지를 통행하는 화물차는 각각의 OD로 통행을 분할하는 것이 타당하므로 통행연결의 기준점은 20분으로 선정하였다. 버스의 경우에는 승용차 및 화물차에 비해 근거로 제시할 수 있는 경유시간 정보 구득에 한계가 있다. 그러나 시외버스 및 고속버스는 화물차와 동일하게 경유지가 발생하고 여객을 상하차 한다는 특성을 가지고 있으며, 통행연결 시 통행량이 적어 다른 차종과 다르게 통행실적의 변화가 크게 발생하지 않은 이유로 동일하게 20분을 기준으로 통행을 연결하였다.

<Table 4> Analysis of traffic performance changes by Vehicle Type

Time	Passenger car(%)	Bus(%)	Trucks(%)
10 minute	0.54%	0.62%	0.51%
20 minute	0.69%	0.59%	0.71%
30 minute	0.54%	0.61%	0.73%
40 minute	0.51%	0.90%	0.78%
50 minute	0.53%	0.92%	0.71%
60 minute	0.47%	0.85%	0.61%



2. 통행실적 결과 분석

통행연결에 대한 OD와 통행량의 분석을 마친 후, 고속도로 수송실적에 기반이 되는 통행실적을 구축하였다. 시행과정은 앞서 말했듯 구축된 OD별 통행량과 한국도로공사에서 제공받은 TCS 최단경로를 거리로 적용하였으며 DSRC 데이터를 이용하여 구축한 실제 주행거리 데이터로 최단경로의 거리 데이터를 보완한 후, 완성된 OD별 주행거리 데이터와 OD별 통행량의 곱에 누락된 교통량 비율 10% 보정 후, 산출하여 대·km/일 단위의 고속도로 통행실적을 구축하였다. 여기서 누락된 교통량 비율 10%란 한국도로공사에서 제시한 일평균 고속도로 통행량 4,964,421(veh/day)과 본 연구에서 집계한 수납데이터의 일평균 고속도로 통행량 4,450,851 (veh/day)을 비교하였을 때 나타난 차이이다. 이러한 수치적 차이는 고속도로 통행료 수납기기 오류 등의 이유로 인하여 발생한 것으로 전국 고속도로 통행실적 산정을 위해서는 보정이 필요하다.

<Table 5>에서 보여주듯이 통행을 연결하였을 때, 통행량은 감소하는 반면 통행실적이 점차 증가하는 것을 확인할 수 있다. 이는 통행연결을 통하여 등급이 다른 도로를 주행한 경우 등의 누락된 구간의 통행실적

<Table 5> Traffic performance of linking segmented O/D trips by 10-minute units

Time	Number of ODs (number)	Rate of change in the number of ODs(%)	Traffic volume (veh/day)	traffic performance (veh·km/day)
00 minute	104,164	-	4,895,936	178,656,786
10 minute	105,170	0.96%	4,846,370	179,602,485
20 minute	108,128	2.74%	4,783,956	180,852,959
30 minute	110,329	1.99%	4,752,859	181,935,383
40 minute	111,927	1.43%	4,728,235	183,026,916
50 minute	112,921	0.88%	4,707,205	184,106,177
60 minute	113,815	0.79%	4,689,321	185,066,679

이 보완되었다는 것을 시사한다. 분석 결과, 20분을 기준으로 통행연결을 진행한 경우의 OD 개수가 총 108,128개이며 이때의 OD 개수 변화 비율이 가장 높은 수치인 2.74%로 도출되었다. 이는 목적통행 중 발생한 수단통행이나 추적을 할 수 없는 누락된 구간을 통행하는 경우가 20분 이내로 가장 많이 발생한다고 해석이 가능하며 이를 20분 기준 통행연결의 근거로 제시하는 바이다.

3. 차종별 통행실적 비교

<Table 6>의 차종별 통행실적 결과와 같이 20분 기준으로 통행을 연결한 후 산정된 통행실적은 승용차가 124,028,590(veh·km/day)로 차종별 통행실적 중 가장 큰 값을 가지며 버스가 3,799,099(veh·km/day)로 가장 작은 통행실적을 가진다. 차종별 통행실적 변화를 분석한 결과 통행연결이 진행될수록 통행연결 시간이 증가할수록 통행실적 또한 증가하므로 차종별로 누락되었던 OD간 주행거리 데이터가 보완됨을 확인할 수 있다.

<Table 6> Traffic performance by vehicle type on expressway

Time	Passenger car(veh·km/day)	Bus(veh·km/day)	Truck(veh·km/day)	All vehicles(veh·km/day)
00 minute	122,520,635	3,753,574	52,382,576	178,656,786
10 minute	123,176,557	3,776,805	52,649,122	179,602,485
20 minute	**124,028,590**	**3,799,099**	**53,025,270**	**180,852,959**
30 minute	124,703,381	3,822,217	53,409,786	181,935,383
40 minute	125,343,605	3,856,501	53,826,809	183,026,916
50 minute	126,007,475	3,892,001	54,206,702	184,106,177
60 minute	126,602,801	3,924,954	54,538,924	185,066,679

<Table 7>는 승용차, 버스, 화물차 차종별로 통행실적이 가장 많은 구간을 추출한 데이터이다. 분석 결과 통행실적이 가장 높은 차종은 승용차로 나타났으며 승용차의 통행실적은 약 444,187(veh·km/day)로 김해부산-대동 영업소의 구간이 가장 높게 산출되었다. 버스의 통행실적은 풍세상2-서울 영업소 구간이 가장 높았으

<Table 7> The top 10 expressway segments with highest traffic performance by vehicle type

Passenger car		Bus		Truck	
Top section (origin-destination)	(veh·km/day)	Top section (origin-destination)	(veh·km/day)	Top section (origin-destination)	(veh·km/day)
GimhaeBusan-Daedong	444,187	Pungsesang 2 - Seoul	41,954	GimhaeBusan-Daedong	166,865
Pungsesang 2-Seoul	341,266	GiheungDongtan-Seoul	22,532	Pungsesang 2-Seoul	106,279
Mokpo-DongGwangsan	323,249	SuwonSingal-Seoul	22,292	Mokpo-DongGwangsan	100,745
DongGwangsan-Mokpo	293,591	Cheongju-Seoul	21,109	DongGwangsan-Mokpo	92,967
Seoul-Cheonan	264,640	Seoul-Cheongju	20,306	NamNonsanha-Gwangju	88,679
Cheonan-Seoul	250,278	Seoul-GiheungDongtan	19,687	NamNonsanha-DongSuncheon	87,534
DongGwangju-Suncheon	229,017	Cheonan-Seoul	19,158	NamSuncheon-SeoYoungam	70,303
SuwonSingal-Seoul	228,353	NakdongJC2-Seoul	19,100	SeoYoungam-NamSuncheon	65,980
Suncheon-DongGwangju	222,062	Seoul-Cheonan	18,844	NakdongJC2-DongSeoul	65,676
Dongsan-Sokcho	214,463	Chilwon-Seoul	18,817	Suncheon-DongGwangju	64,335

며, 화물차는 승용차와 마찬가지로 김해부산-대동 영업소 간의 통행실적이 가장 높은 것으로 도출되었다. 화물차의 통행실적은 주로 항만이나 산업단지로 연결되는 김해부산, 목포, 남순천 등을 기종점으로 하는 구간이 높게 나타났으며, 버스는 서울을 기종점으로 천안, 기흥동탄, 수원신갈 간의 출퇴근목적의 여객수송량이 많은 구간의 통행실적인 높게 나타나는 것을 알 수 있다.

4. 고속도로 통행실적 결과 도출

수납데이터는 데이터 수집에 어려움이 있는 비연계 구간에 대해서는 교통량 데이터가 집계되지 않는다는 한계성이 존재한다. 이 한계성을 보완하기 위해 TMS 통계연보에서 비연계 구간의 통행실적을 산정하여 합산한 값을 전국 고속도로의 통행실적으로 제시하였다. 비연계 구간의 통행실적은 67,771,555(veh·km/day)로 집계되었으며, TCS 기반 고속도로 통행실적 180,852,959(veh·km/day)와 비연계 구간의 통행실적 67,771,555(veh·km/day)를 합산하여 전국 고속도로 통행실적 248,624,514(veh·km/day)를 산정하였다. 이 방법론을 통해 기존에 데이터가 누락된 구간에 대한 통행실적 추출이 가능해져 신뢰성있는 전국 고속도로의 통행실적이 산정되었다. 결과적으로 본 연구에서는 2021년 TCS, HI-PASS, DSRC 등의 데이터를 이용한 전국 고속도로 통행실적을 약 248,624(thousand·veh·km/day)로 제시하였다. 도로업무편람에서 제시하는 2021년 고속도로 통행실적 248,166(thousand·veh·km/day)과 비교하였을 때, 본 연구에서 산정한 전국 고속도로의 통행실적이 약 458(thousand·veh·km/day)가 크다는 것을 확인할 수 있다.

이렇듯 여러 통행으로 분할된 개별 차량의 통행궤적을 연결하여 통행실적이 약 1% 미만으로 미미하게 증가하였으나 이는 본 연구에서 7일치의 수납데이터가 활용되었으며, 수납데이터와 DSRC 주행거리의 매칭 비율이 약 40% 정도로, 나머지 60%의 OD는 DSRC로부터의 실제 주행거리가 아닌 OD간 최단거리를 적용하였다는 점에서 일부 한계성이 있다. 향후 표본이 늘어 100% 매칭이 가능해진다면 통행실적 수치는 현재보다 훨씬 증가할 것으로 기대된다. 이는 기존의 통행실적 산정을 위한 조사방법의 한계로 과소 추정되었던 통행실적 값을 실증빅데이터를 통하여 OD를 구축하고 실주행거리 산정을 통해 보다 정확한 값으로 보정 가능하다는 것을 보여준다.

Ⅴ. 결 론

국내 고속도로는 여객 및 화물 수송의 중추적인 역할을 담당하고 있음에도 불구하고 타 수단과 수송실적을 정확히 비교할 수 있는 적정한 지표가 미흡한 실정이다. 이에 본 연구는 TCS, DSRC 등 데이터를 활용하여 고속도로 통행 OD 구축과정을 포함한 통행실적 산정 방법론의 틀을 정립하는데 목적이 있으며, 궁극적으로는 향후 철도, 항공 등 타 수단과 비교 가능한 수송실적 지표 값을 제시하고자 하였다.

본 연구는 고속도로 실증 데이터를 이용함으로써 통행량, 주행거리 등을 추출하였으며, 하나의 통행이 분할되어 나타나는 오류를 해결하고자 다수의 끊어진 OD를 연결하는 기준을 제시하여 보다 신뢰성있는 고속도로 통행실적을 도출하고자 하였다. 개별 차량이 20분 내 연속되는 OD가 다수 있을 경우, 통행을 연결하는 것이 가장 타당하다고 제시하였으며, 통행량이 집계되지 않는 비연계 구간에 대한 통행실적 보완을 통하여 전국 고속도로 통행실적은 248,624,514(veh·km/day)로 도출되었다. 현재 도로업무편람에서 제시하는 고속도로 통행실적 248,166(thousand·veh·km/day)과 비교하였을 때, 본 연구에서 제시하는 고속도로의 통행실적이 약 458(thousand·veh·km/day)만큼 더 크다는 사실을 확인할 수 있다. 이는 기존에 산정되고 있는 고속도로의 통

행실적에서 누락된 거리 및 통행량이 있으며 이를 TCS 및 DSRC 데이터 분석을 통해 보완할 경우 정확한 고속도로의 통행실적을 산정가능하여 보다 정확한 결과를 도출할 수 있다는 것을 시사한다. 또한, 본 연구에서는 시스템상 연계되지 않아 통계연보를 활용하여 산출했던 비연계 구간의 통행실적을 한국도로공사와 민간기업의 데이터 공유를 통해 데이터가 연계된다면 더욱 신뢰성 있는 고속도로 통행실적이 집계될 것으로 기대된다.

TCS 데이터를 통해 도출한 승용차, 버스, 화물차의 차종별 통행실적을 확인한 결과 비연계 구간을 제외한 전체 구간의 총 값은 각각 124,028,590(veh·km/day), 3,799,099(veh·km/day), 53,025,270(veh·km/day)로 산정되었으며, 차종별 비교 결과 승용차의 통행실적이 버스와 화물차의 통행실적보다 각각 32배, 2배 이상 높음을 확인할 수 있다. 차종별 통행실적이 가장 많은 구간을 확인한 결과, 승용차와 버스의 경우 서울 및 경기 권역에 통행실적이 높게 산정되었으며, 화물차의 경우 경남 또는 전남 등 항구와 가까운 지방에 통행실적이 높게 산정되었다. 이는 여객 수송을 주로 담당하는 승용차와 버스가 서울 및 경기 권역의 고속도로에 많이 분포한다는 특성을 확인할 수 있으며, 화물차는 화물 이동이 많은 산업단지 및 항만 주변에 통행이 많이 분포되고 있다는 것을 보여준다. 이러한 결과는 향후 고속도로 차종별 사고 분석, 화물차 전용차로 개설 등 다양한 정책 방안으로 제시될 수 있을 것이다.

향후 본 연구를 통하여 고속도로 여객 및 화물에 대한 수송실적 지표를 산정할 수 있게 되었으며, 나아가 국도 및 지방도를 포함하는 공로로 범위를 확장하여 도로 부문의 여객 및 화물 수송실적을 산출할 수 있을 것으로 기대된다. 현재는 표준화된 고속도로 통행실적 구축 방법론이 정립되어 있지 않기 때문에 COVID-19 발병과 같이 고속도로 통행량에 영향을 미치는 이벤트에 대한 통행실적 변화를 파악하는데 어려움이 존재한다. 그러나 본 연구에서 정립한 방법론을 활용하여 미시적 시공간 범위의 고속도로 통행실적 산정 및 변화량에 대한 분석이 가능해졌다. 이를 통하여 고속도로 통행에 대한 즉각적인 분석 및 대안 제시가 신속히 이루어질 수 있을 것으로 기대된다.

마지막으로 본 연구에서 이용한 빅데이터는 7일(2021년 10월)간이라는 한계성이 존재한다. 향후, 장기간 누적된 데이터 분석이 수행된다면 지정 연도의 실증 통행실적을 산정할 수 있으며 정확한 고속도로 통행실적을 산정할 수 있게 될 것으로 기대한다. 또한 통행실적 구축 시 적용되었던 OD간 차량 주행거리는 DSRC 데이터 표본 수의 한계로 한국도로공사에서 제공하는 영업소간 최단경로 거리 데이터와 통합하여 분석에 활용하였으나 향후 DSRC 데이터의 표본을 확장하여 OD별 개별 차량의 실제 주행거리를 산정한다면 고속도로 통행실적의 정확도를 보다 향상시킬 수 있을 것이다.

ACKNOWLEDGEMENTS

본 연구는 한국도로공사 도로교통연구원 “고속도로 수송 경쟁력 분석체계 구축 및 사례분석 연구” 지원으로 수행하였습니다.

REFERENCES

Fu, X., Yang, H., Liu, C., Wang, J. and Wang, Y.(2019), “A hybrid neural network for large-scale expressway network OD prediction based on toll data”, *Public Library of Science*, vol. 14, no.

5, pp.1-15.

Han, D. H., Kim, S. H., Park, J. J., Lee, J. H. and Kim, J. H.(2017), *A Study on the Traffic Analysis Method Using Vehicle Trajectory Data*, p.168.

Jun, J. W.(2013), *Closed Highway Traffic Prediction Method Using TCS Traffic Using TCS*, Masters degree, Hanyang University, Korea.

Jung, H. S.(2012), *A study of data preprocessing algorithm using TCS/HI-PASS data*, Masters degree, Korea University, Korea.

Kang, U. J., Hong, J. Y. and Choi, Y. H.(2023), "Algorithm for Freight Transportation Performance Estimation on Expressway Using TCS and WIM Data", *The Journal of The Korea Institute of Intelligent Transport Systems*, vol. 22, no. 3, pp.116-130.

Kim, J. J.(2001), *Estimation of link volumes on the freeway using the TCS(toll collection system) data*, Masters degree, Hanyang University, Korea.

Kim, J. S.(2007), *A Study on Demand Estimation of Leisure Trip through Highway TCS DATA at a Tourism destination*, Mmasters degree, Kwandong University, Korea.

Korea Railroad Research Institute(2017), *Final Report on the Development of Batch Unloading Equipment Technology to Reduce Loading and Unloading Time for Cargo Vehicles*, p.391.

Ministry of Land, Infrastructure and Transport(2022), *Road Work Manual*, p.535.

Park, W. S.(2008), *Driving Characteristics Classification of TCS Data Based on Fuzzy c-means Clustering Algorithms*, Masters degree, Kyungwon University, Korea.

Shan, J. and Li, X. H.(2008), "Estimating the Highway Freight Origin-destination Matrix from Multi-source Data Based on Fuzzy Programming Theory", In *2008 International Workshop on Modelling, Simulation and Optimization*, IEEE, pp.204-209.

Shi, F., Bian, K. and Xie, K.(2016), "Vulnerability analysis of highway traffic networks using origin-destination tollgate data", In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, pp.1957-1963.

Shin, S. J., Park, D. J., Choi, Y. H., Jung, S. Y., Heo, E. J. and Ha, D. I.(2013), "Estimation of Expressway O/D Matrices from TCS data by Using Video Survey Data for Vehicle Classification: Focused on Truck", *The Journal of The Korea Institute of Intelligent Transport Systems*, vol. 12, no. 1, pp.136-146.

The Korea Expressway Corporation(2022), *A Study on Establishing a Performance Evaluation of Highway Network System based on Mobility Big Data*, p.287.