교통카드자료를 이용한 환승정류장의 개발 방향에 관한 연구

- 대구시를 중심으로 -

A Study on the Developmental Directions of Transfer Stations with Traffic Cards Data
- Focused on Daegu City -

김기혁* · 이승철**

Kim, Ki-Hyuk · Lee, Seung-Cheol

Abstract

Increasing the mode transfer volume between public transportation modes has known to be necessary for efficiency improvement of public transportation system operation and it is also found to be important to have relevant transfer point selection with reflection of current travel pattern. This study is in regards to providing a selection guideline for the location of transfer point between public transport modes. This case study has been carried out for Daegu Metropolitan City especially for public transportation users behaviour by analysis of daily usage of transportation card to identify the transfer travel pattern. A cluster analysis was applied to categorize the pattern of transfer stop which induces many users and a discriminant analysis also utilized for grouping the stops by number of transfer trip. This research produces the estimation result of transfer volume for urban railway system no.3 in Daegu City which is currently under construction. In addition, the locations of transfer center has also been proposed.

Keywords : *traffic cards, transfer stations, cluster analysis, discriminant analysis*

요 지

교통수단간 환승은 대중교통체계운영의 효율성 증대를 위해 반드시 필요한 것으로 무엇보다 중요한 것은 현실에 기초한 환승거점의 선정이라고 할 수 있다. 본 연구는 대중교통 이용자의 통행실태를 반영한 환승거점 입지선정 방향에 관한 연구로서 대구광역시를 대상으로 대중교통 이용자 1일 교통카드 사용내역 자료를 활용하여 환승통행패턴을 분석하고자 한다. 군집분석을 통해 이용객이 많은 정류소를 분류하고 환승통행량의 규모에 따라 그룹핑하여 판별분석을 실시하였다. 도출된 판별함수를 이용하여 현재 시공 중인 대구시 도시철도 3호선의 환승통행량을 추정하고 환승센터 설치 가능지점을 제시하였다.

핵심용어 : 교통카드, 환승정류장, 환승센터, 군집분석, 판별분석

1. 서 론

1.1 연구의 배경 및 목적

대도시에서 대중교통이용의 불편과 개인교통수단 선호의식으로 인한 자가용 증대는 교통 혼잡을 심화시키는 주요한 요인이 되고 있다. 교통 혼잡으로 인하여 대기오염, 열섬현상, 교통사고 증가, 녹지 감소 등 많은 도시교통문제와 환경문제가 심각해짐으로써 지속가능한 도시개발에 대한 관심이 높아지고 있다.

1990년대 후반 철도역과 주요 버스정류장들을 포함하는 대중교통결절점을 중심으로 하는 도시개발의 개념이 국내에 도입되었다. Feter Calthope에 의해 처음 정립된 대중교통지향형도시개발(Transit-Oriented Development: TOD)은 토지이용과 교통의 연관성을 강조하고 대중교통중심의 복합적 토

지이용과 보행친화적인 교통 환경으로 유도하고자 하는 방식이다. 도시계획적인 측면에서 TOD는 무분별한 도시의 외연적 확산을 억제하고 승용차 중심의 통행패턴을 대중교통 및 녹색교통 위주의 통행패턴으로 변화시키는 기법으로 인식된다. 미국이나 유럽에서 TOD개념을 적용하여 개발된 사례가 많이 있다. 국내에서 TOD개념이 도입되기 전 역세권 개발을 통해서 고밀도개발을 해왔다. 역을 중심으로 다양한 상업 및 업무활동이 이루어지도록 혼합지역으로 개발한다는 점에서 TOD개념의 하나로 볼 수 있다. 역세권 개발의 의미로 최근 환승센터 건설이 대두되고 있다. 환승센터는 대중교통수단 간의 연계 및 환승편의 제공, 승용차 등의 개인교통수단과의 연계 및 환승편의 제공 등 대중교통 이용활성화로 도시교통문제 해소에 기여하고 있으나, 이 기능을 다하기 위해서는 환승센터에 대한 체계적인 연구와 입지 기준의 정립

*정회원 · 교신저자 · 계명대학교 교통공학과 교수 (E-mail : ghkim@kmu.ac.kr)

**계명대학교 교통공학과 박사 (E-mail : nemophila@nate.com)

이 필요하다.

대구시는 교통카드 사용률이 91%에 달하고 있으나, 교통카드자료를 활용하는데 있어 많은 제약이 따르고 있다. 버스의 운행 스케줄의 업데이트 문제로 인하여, 승객의 O/D를 파악할 수 없어 교통카드를 통해 수집할 수 있는 자료가 매우 부정확한 실정이다. 따라서 본 연구에서는 교통카드자료에서 수집할 수 없는 위치 정보를 대구시 BMS자료를 이용하여 환승자료를 구축하고, 이를 토대로 환승정류장의 특성을 파악하였다. 현재 공사 중인 도시철도 3호선의 환승량을 추정하기 위하여 환승자료를 이용하여 판별분석을 실시하였으며, 대구시 실정에 맞는 환승정류장의 입지를 찾고자 하였다.

1.2 연구의 범위 및 방법

본 연구의 범위는 대구광역시내의 시내버스와 도시철도를 이용하는 승객들의 교통카드를 대상으로 이루어졌다. 대구시 교통카드의 경우 서울의 스마트카드와는 달리 통행시간, 통행거리, 기종점 통행량 등의 자료를 내포하지 않기 때문에 환승정류장을 파악하기 어렵다. 따라서 본 연구에서는 대구시 교통카드자료를 환승유형별로 구분한 뒤, 추출된 자료의 카드결제시간을 가공하여, 차량별 상세 통신 정보의 버스정류장 도착시간과 Matching할 수 있도록 하였다.

수집된 대구시 교통카드자료는 2011년 9월 7일 평일 자료로서 버스와 도시철도 이용을 포함한 1,746,718건의 자료이며, 또한 2011년 9월 7일 기준으로 운행된 버스차량대수는 1,533대이며, 운행된 1,533대에 대한 통신정보자료를 수집하였다. 환승자료구축을 위해서 차량별 상세 통신정보의 GPS 시간과 교통카드 결제시간을 Matching하여 위치정보를 수집하였다.

환승정류장에 대한 변수로서 노선수, 평균대기시간, 일승차 인원, 도심으로부터의 거리, 평균접근거리, 환승시간, 환승량 대하여 특성 분석을 실시하였다.

환승센터의 입지는 권역별 통행특성, 환승패턴, 개발여건, 도시공간구조상의 위계, 토지규모 및 형태 등을 감안하여 환승다발지점들을 대상으로 하여 선정하게 된다. 따라서 본 연구에서는 일승차인원을 기준으로 이용객이 많은 정류장을 군집분석을 실시하여 분류한 뒤, 환승량별 그룹을 나누어 판별분석을 실시하였다. 분석된 판별함수를 이용하여 현재 공사 중인 대구시의 도시철도 3호선에 대한 환승량을 추정하여, 환승센터 설치 가능지점을 제시하였다.

2. 이론적 고찰

2.1 선행연구고찰

권영중(2004)은 우리나라에도 TOD의 도입이 시급한 실정이며, 성공적인 도입을 위해서 관련제도의 개선과 설계지침, 추진체계 정립 등이 따라야 한다고 보고 있다. 외국의 사례 검토와 도시개발 문제점 진단 및 원익 분석 등을 주요 내용으로 다루었다.

오영택(2008)은 대중교통의 공급특성과 이용지수와의 연관성을 규명하였다. 서울시 249개의 도시철도역이 역세권별로 대중교통공급수준이 다르며, 또한 토지이용특성이나 지형적 특성도 다르게 나타나고 있기 때문에 역세권별로 보완해야

할 요인이 다를 것이라 판단하여, 역세권의 공급현황에 따른 차별된 개선방안을 제시하였다.

문영일(2011)은 역세권 TOD의 물리적 계획요소 특성을 요인분석을 통하여 파악하고, 그 특성이 역세권의 승차인원과 연관성을 지니고 있는지를 요인-회귀분석을 통해 도출하였다. 요인분석결과 토지이용복합도, 개발밀도수준, 대중교통시설수준, 가로설계수준, 녹색교통연계시설수준, 지하철연계시설수준, 대중교통운행수준으로 7개 요인으로 분석되었다. 요인-회귀분석 결과에서는 대중교통운영수준만 부정적(-)영향으로 나타났으며, 이는 지하철과 버스의 배차간격이 증가할 경우 이용수요가 감소하기 때문인 것으로 판단하였다.

성현곤(2008)은 TOD의 중요 구성요소인 밀도(Density) 및 다양성(Diversity) 등의 토지이용특성과 도시설계특성이 대중교통이용증대에 미치는 영향을 파악하기 위하여 요인분석과 다중회귀분석을 실시하였다. 업무 및 상업 중심의 고밀도 특성과 토지이용의 복합화, 그리고 협소한 가로망이 밀집되어 있는 특성은 대중교통이용증대에 매우 긍정적인 영향이 있는 것으로 나타났으며, 접근성이 좋지 않거나 상업시설 등이 없이 주거 중심의 특성, 간선도로 위주의 큰 규모의 건물군이 형성되어 있는 경우 대중교통이용증대에 부정적인 영향이 있는 것으로 나타났다.

2.2 군집분석과 실루엣 크기

군집화 분석(Cluster Analysis)은 자료 뒤에 내재되어 있는 내재적 그룹들을 드러내는 데 쓰이는 탐색적 자료 분석 기법이다. 개체(Units, Records)간 거리에 근거하여, 서로 가까운 개체들은 같은 군집으로 묶이고 서로 떨어진 개체들은 다른 군집에 속하도록 한다. 군집화에는 여러 방법이 있지만, 계층적 군집화(Hierarchical Clustering)방법들은 모든 쌍의 개체간 거리를 메모리에 저장해 놓고 업데이트하기 때문에 규모가 큰 데이터 마이닝에서는 사용되지 않는다. 대신 데이터 마이닝에서는 비계층적 군집화(Nonhierarchical Clustering)방법인 K-means, 신경망 기법의 하나인 Kohonen, 그리고 메모리 사용이 효율적인 Two step을 활용한다. 비계층적 군집방법은 분할방법(Partitioning Method)이라고도 하며, 군집의 수를 사전에 지정하고 대상 객체들을 적절한 군집에 배정하는 방법이다. 군집분석에서 군집의 수를 선택하는 것은 가장 기본적인, 매우 어려운 문제이다. 실제 군집의 수는 자료의 분포와 크기, 형태에 따라 의존하거나 연구자가 원하는 군집의 결과로 인하여 정확한 선택이 어렵다. 이러한 문제로 본 연구에서는 실루엣의 너비를 측정하여 군집의 수를 정하였다.

실루엣(Silhouette)은 군집 내 데이터가 어느 정도 가깝게 대응하는 것과 이웃 군집 데이터가 어느 정도 대응되는지에 대한 측도이다(Rousseeuw, 1987). 실루엣 너비는 -1부터 1까지의 값을 가진다. 1에 가까우면 관찰값은 주변 군집보다 지금 속한 군집에 더 잘 속하며, -1에 가까우면 군집이 잘 되지 않았다고 판단할 수 있다. 1번째 관찰값의 실루엣의 너비는 식과 같다.

실루엣 너비를 통하여 평균 실루엣 너비를 구할 수 있다. Kaufman과 Rousseeuw는 평균 실루엣 너비를 가장 크게 하는 군집의 개수가 가장 최적이라고 판단하였다. 평균 실루엣

너비가 0.5 이상이면 군집을 구분하는데 유의하며, 0.2보다 작은 경우는 군집을 구분하는데 합리적이지 않다고 보았다.

2.3 판별분석

판별분석은 사전에 정의된 하나의 집단변수(group variable) 개개의 변수값을 판별하는데 중요한 역할을 하는 판별 변수(discriminant variable)의 선형결합(linear combination)인 판별함수(discriminant function)를 만든 후, 이 판별함수에 새로운 개체의 특성을 대입하여 어떤 집단에 속할지를 판별하는 것이다. 판별분석은 기준변수나 종속변수의 범주를 보다 명확하게 구분하기 위한 판별함수를 개발하거나 예측 변수나 독립변수들의 선형결합관계를 찾아내기 위하여 실시한다. 판별분석은 다음과 같은 과정으로 분석된다.

- ① 어떤 집단에 속하는 구별하는데 중요한 역할을 하는 판별 변수들을 독립변수들 중에서 택한다.
- ② 선택된 판별변수들의 선형결합(판별함수)을 만들어 개체들을 집단으로 분류하는 기준을 마련한다.
- ③ 판별함수의 의하여 개체들이 얼마나 정확히 분류되는지 그 판별정도를 파악한다.
- ④ 새로운 개체가 유입되었을 때 판별변수와 판별함수에 적용하여 어떤 집단에 속할지를 판단한다.

3. 환승 DATA 구축

대구시 교통카드자료를 원시데이터로 하여, 환승데이터를 추출하였다. 추출된 자료의 카드결제시간을 가공하여, 차량별 상세 통신 정보의 버스정류장 도착시간과 Matching할 수 있도록 하였다. BMS자료의 하나인 정류장별 위치정보는 Arc-GIS에서 정류장들의 위치좌표와 속성자료를 입력하여, 분석이 용이하도록 하였다.

3.1 교통카드자료

대구시 교통카드자료는 2011년 9월 7일 평일 데이터로서 전 시간에 대한 자료로서 버스와 도시철도 이용을 포함한 1,746,718건의 자료를 수집하였다.

수집된 자료를 환승유형별로 구분하였다. 환승은 버스에서

```
ID,TR_DATE,TR_TIME,STATE_MARK,LINE_ID,CAR_ID,CARD_CMP_CODE
1,20110907,064442,04,004010,701228,00
2,20110907,164156,04,008360,701645,00
2,20110907,211957,04,007190,705314,00
4,20110907,064952,04,004270,702605,00
4,20110907,071008,01,004270,702605,00
4,20110907,173943,14,004270,702697,00
4,20110907,172702,04,002730,702645,00
4,20110907,173644,01,002730,702645,00
4,20110907,071644,14,002730,702612,00
10,20110907,180744,01,003030,701554,00
10,20110907,175417,04,003030,701554,00
10,20110907,183250,01,006090,701935,00
10,20110907,182336,14,006090,701935,00
14,20110907,170433,20,1,-,00
14,20110907,174531,30,1,-,00
14,20110907,122029,06,001160,707311,00
17,20110907,075823,06,009390,707394,00
18,20110907,133443,04,007300,702699,00
18,20110907,123611,01,007300,702644,00
18,20110907,133950,01,007300,702699,00
18,20110907,123325,14,007300,702644,00
18,20110907,122435,01,004020,701848,00
18,20110907,122101,04,004020,701848,00
18,20110907,135407,01,003040,702412,00
18,20110907,134406,14,003040,702412,00
26,20110907,131202,04,001020,701063,00
26,20110907,220612,04,006230,705204,00
28,20110907,080542,08,007080,707425,00
28,20110907,085836,01,002320,701081,00
28,20110907,085435,04,002320,701081,00
28,20110907,090134,14,005030,701134,00
28,20110907,113436,34,009370,705037,00
28,20110907,110540,20,2,-,00
28,20110907,112325,30,1,-,00
35,20110907,062945,20,1,-,00
35,20110907,065200,30,1,-,00
35,20110907,163114,20,1,-,00
35,20110907,164832,30,1,-,00
39,20110907,101742,20,2,-,00
39,20110907,113400,30,2,-,00
39,20110907,112416,20,2,-,00
39,20110907,103621,30,2,-,00
39,20110907,135353,05,001050,702049,00
44,20110907,194539,08,001120,707412,00
44,20110907,182449,08,001180,707432,00
```

그림 1. 교통카드자료

버스, 버스에서 도시철도, 그리고 도시철도에서 버스 이상 세 가지 유형으로 구분하여 분석을 실시하였다. 현재 대구시는 도시철도노선이 2개만 운영되고 있어, 환승이 대구 도심인 반월당역 한 곳에서만 이루어지고 있다. 따라서 본 분석에서는 도시철도에서 도시철도의 경우는 제외하고 분석을 실시하도록 하였다.

3.2 대구시 BMS자료

BMS 자료 중의 하나인 차량별 상세 통신정보는 운행한 노선의 이력이 담겨있는 자료이다. 버스의 위치정보는 주요 교차로 통과시나, 정류장 출발과 도착, 그리고 돌발 상황 시 이벤트 통신으로 위치 정보가 수집된다.

2011년 9월 7일 기준으로 운행된 버스차량대수는 1,533대이며, 운행된 1,533대에 대한 통신정보자료를 수집하였다. 환승자료구축을 위해서 차량별 상세 통신정보의 GPS시간과 교통카드 결제시간을 Matching하여 위치정보를 수집하였다.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	차량별 상세 통신정보													
2	운행일자 : 2011년 09월 07일													
3	버스회사 : 한일운수(주)													
4	차량번호 : 1969													
5	CRF_ID : 1644*****													
6	CDMA번호 : 010-*****													
7														
8	노선번호	순위	방향	방면	노드	GPS시간	속도(km/h)	이벤트	운행정보	돌발상황	GPS좌표(X)	GPS좌표(Y)	GPS상태	수집시간
9	909	13	역	영남대 (7121005700)	조물공사	23:19:51	68	버스정류소도착	공차		128,458862	35,49667	정상 (01,00)	23:19:56
10	909	13	역	영남대 (7121005700)	강제돌회관	23:19:37	53	버스정류소도착	공차		128,458405	35,497932	정상 (01,00)	23:19:42
11	909	13	역	영남대 (7121005700)	영남대돌문	23:18:55	70	버스정류소도착	공차		128,458817	35,502216	정상 (01,00)	23:19:02
12	909	13	역	영남대 (7121005700)	암랑한리빌라2	23:18:21	42	버스정류소도착	공차		128,457596	35,504257	정상 (01,00)	23:18:26
13	909	13	역	영남대 (7121005700)	암랑한리빌라트	23:18:10	33	버스정류소도착	공차		128,456985	35,504753	정상 (01,00)	23:18:12
14	909				75번돌점	23:17:46	51	버스정류소출발			-	-	정상	23:17:52
15	909	13	역	영남대 (7121005700)	75번돌점	23:17:41	49	버스정류소도착	공차		128,454666	35,5037	정상 (01,10)	23:17:47
16	909				영남대	23:17:13	0	버스정류소출발			-	-	정상	23:17:47
17	909	13	역	영남대 (7121005700)	영남대	23:16:36	22	버스정류소도착	막차		128,452316	35,502003	정상 (01,00)	23:16:41
18	909				영대오거리(영대방향)	23:15:25	38	버스정류소출발			-	-	정상	23:15:39
19	909	13	역	영남대 (7121005700)	영대오거리(영대방향)	23:15:19	42	버스정류소도착	막차		128,447617	35,500416	정상 (01,00)	23:15:28

그림 2. 차량별 상세 통신정보

그림 3. 환승유형별 환승통행량

표 1. 자료수집

환승유형	통행	정류장(개)
버스-버스	118,768	1,933
버스-도시철도	36,366	755
도시철도-버스	38,801	638

3.3 환승 DATA 구축

교통카드자료와 BMS자료를 이용하여 환승유형별로 자료를 구축하였다. 유형별로 첫 번째 수단에서 승차와 하차, 그리고 두 번째 수단에서 환승의 순서가 이루어지도록 자료를 구축하였다.

구축된 환승자료는 모두 버스정류장을 기준으로 정리하였다. 예를 들어 버스에서 도시철도로 환승하는 경우 환승정류장이 도시철도역이 되지만, 장래 도시철도 계획 시 환승수요를 추정하기 위하여 버스정류장 기반으로 정리를 하였다. 구축된 자료는 그림 3과 같으며, 원의 크기는 환승량의 크기를 나타낸다.

버스에서 버스로 환승하는 경우는 도심 가까이에서 많이 이루어지고 있으며, 버스에서 도시철도, 도시철도에서 버스로 환승하는 경우 도심부와 종점부에서 많이 이루어지는 것으로 나타나고 있다.

4. 환승통행의 특성 분석

4.1 환승통행 변수 분석

환승통행의 행태를 분석하기 위하여 표 2와 같은 환승통행 변수를 파악하였다.

노선수는 각 버스정류장에서 서비스되고 있는 총 노선의 수로서 대구광역시 버스노선안내시스템을 이용하여 정류장의 노선수를 파악하였다. 평균배차시간은 해당 정류장 노선들의

평균배차시간이며, 일승차인원은 해당 정류장의 일일 승차인원에 대한 정보로서 교통카드자료를 가공하여 수집된 자료이다. 도심으로부터의 거리는 대구시에서 지리적인 중심점인 반월당역을 기준으로 한 떨어진 거리이며, Arc-GIS를 이용하여 두 지점간의 직선거리를 측정하여 적용하였다. 환승량은 환승 수단에서 환승을 하는 승객수이며, 접근거리는 첫 번째 수단을 이용한 정류장에서 환승정류장까지의 직선거리이다. 환승시간은 첫 번째 수단에서의 결제시간부터 환승수단에서의 결제시간까지로 이동 및 대기시간이 포함되어 있다.

환승량과 일승차인원과 관계에서 버스에서 버스로 환승하는 경우의 정류장들과 도시철도에서 버스로 환승하는 경우의 정류장들은 선형관계가 있는 것으로 나타났다. 그리고 버스에서 도시철도로 환승하는 경우는 일승차인원과 노선수에 대해서 선형관계가 있는 것으로 나타났다.

평균접근거리는 버스에서 도시철도로 환승하는 경우의 정류장이 3.71km로 가장 짧았으며, 버스에서 버스로 환승하는 경우 4.62km, 도시철도에서 버스로 환승하는 경우 8.45km 순으로 나타났다.

버스에서 버스로 환승하는 경우의 정류장들의 도심거리가 5km 이하인 정류장들이 환승량이 많은 것으로 나타났으며, 버스에서 도시철도, 도시철도에서 버스로 환승하는 경우의 정류장들은 5km 이상에서 환승량이 많은 것으로 나타났다.

버스에서 도시철도로 환승하는 경우의 정류장들은 환승시간이 5분 이하일 때 환승량이 많은 것으로 나타났으며, 나머지 유형에서는 환승시간이 10분 이하일 때 환승량이 많은 것으로 나타났다.

4.2 환승유형별 판별 분석

환승센터의 입지는 권역별 통행특성, 환승패턴, 개발여건, 도시공간구조상의 위계, 토지규모 및 형태 등을 감안하여 환

표 2. 환승통행의 변수

구 분	내 용	비 고
노선수	해당 정류장의 노선수(개)	대구버스노선안내시스템
평균배차시간	해당 정류장 노선들의 평균배차시간(분)	대구버스노선안내시스템
일승차인원	해당 정류장의 일승차인원(명)	교통카드자료(가공)
도심거리	대구 중심으로부터의 직선거리(km)	GIS자료
환승량	환승유형별 환승량(명)	환승자료
접근거리	해당 정류장의 평균접근거리	환승자료
환승시간	하차정류장에서 환승정류장까지의 이동시간 및 대기시간(분)	환승자료

표 3. 상관분석결과

버스-버스 환승정류장 상관분석							
변수	노선 수	평균 배차시간	일 승차 인원	환승량	도심 거리	환승 시간	접근 거리
노선수	1	-.040	.588	.577	-.052	-.274	-.125
평균 배차시간		1	-.010	-.009	-.179	.044	-.062
일. 승차인원			1	.791	-.172	-.179	-.145
환승량				1	-.206	-.209	-.130
도심거리					1	.082	.434
환승시간						1	.166
접근거리							1

버스-도시철도 환승정류장 상관분석							
변수	노선 수	평균 배차시간	일 승차 인원	환승량	도심 거리	환승 시간	접근 거리
노선수	1	-.095	.598	.169	-.053	.018	.090
평균 배차시간		1	-.101	-.032	.078	-.071	.012
일. 승차인원			1	.181	-.186	.019	.043
환승량				1	.025	-.187	-.035
도심거리					1	.001	.191
환승시간						1	.254
접근거리							1

도시철도-버스 환승정류장 상관분석							
변수	노선 수	평균 배차시간	일 승차 인원	환승량	도심 거리	환승 시간	접근 거리
노선수	1	-.019	.596	.212	-.082	-.104	-.037
평균 배차시간		1	-.047	-.009	.067	-.079	.052
일. 승차인원			1	.492	-.177	-.065	-.029
환승량				1	.000	-.224	.015
도심거리					1	-.056	.421
환승시간						1	-.013
접근거리							1

승다발지점들을 대상으로 하여 선정하게 된다. 따라서 본 연구에서는 일승차인원을 기준으로 이용객이 많은 정류장을 군집분석을 실시하여 분류한 뒤, 환승량별 그룹을 나누어 판별 분석을 실시하였다.

적정 군집수를 찾기 위하여 실루엣 통계량(silhouette statistic) 기준을 적용하였다. 실루엣은 군집 내 데이터가 어

느 정도 가깝게 대응하는 것과 이웃 군집 데이터가 어느 정도 대응되는지에 대한 척도이다. 평균 실루엣 너비가 0.5 이상이면 군집을 구분하는데 유의하며, 0.2보다 작은 경우는 군집을 구분하는데 합리적이지 못하다고 할 수 있다.

4.2.1 버스에서 도시철도로 환승하는 경우 판별 분석

4.2.1.1 군집분석

버스에서 도시철도로 환승하는 경우의 정류장들에 대하여 일승차인원을 기준으로 군집분석을 실시하였다. 각 분석기법들 중 실루엣 크기가 가장 큰 군집분석방법을 선택하였다.

3가지 군집분석 방법 중 실루엣 크기가 가장 크게 나타난 k-means 방법을 채택하였다. k-means 방법은 5군집으로 군집을 나누었으며, 군집중심은 각각 216.18, 7354.5, 2466.6, 1020.9, 4603.8로 나타났으며, 샘플수는 각각 497, 2, 41, 209, 6개로 분류되었다.

5군집 중 일승차인원이 가장 적은 군집-1을 제외하고 군집중심이 1,000 이상인 군집-2, 3, 4, 5에 대하여 환승량을 그룹화하였다.

k-means 방법이 0.867로 실루엣 크기가 가장 높게 나타났으며, 그에 따른 적정 군집 수는 5개로 나타났다.

군집1은 군집중심이 1464로 환승량이 가장 많은 군집이며, 도시철도2호선 종점역이다. 군집2는 군집중심이 23.07로 환승량이 가장 작은 군집이며, 가장 많은 정류장이 분류되었다. 군집3은 군집중심이 874.67로 환승량이 세 번째로 많은 군집이며, 군집4는 환승량이 가장 적은 군집이다. 군집5는 군집중심이 988로 환승량이 두 번째로 많은 군집이며, 도시에 위치해있다.

4.2.1.2 판별분석

환승정류장 특성 변수를 사용하여 판별분석을 실시한 결과

표 4. 일승차인원 군집분석

분석기법	실루엣 크기	군집 수
k-means	0.731	5
TwoStep	0.693	5
Kohonen	0.546	12

표 5. 환승량 군집분석

분석기법	실루엣 크기	군집 수
k-means	0.867	5
TwoStep	0.864	3
Kohonen	0.649	12

그림 4. 군집분석과정

그림 5. 버스-도시철도 군집별 환승량도

표 6. Wilks의 람다

함수의 검정	Wilks의 람다	카이제곱	자유도	유의확률
1	.913	23.173	4	.000

유의확률이 0.000으로써 집단 간 평균의 차이가 있으며, 통계적으로 유의하다고 볼 수 있다. 버스에서 도시철도로 환승하는 경우의 정류장들에 대하여 판별분석을 실시한 결과 모형은 다음과 같은 식으로 표현된다.

$$z=0.001*d_{per}-1.485 \quad (1)$$

식 (1)을 이용하여 나온 값이 함수의 집단 중심점에 가까운 값을 선택하여 집단을 분류할 수 있게 된다. 판별함수모형을 이용하여 집단을 판별하는 방법 외에 R. A. Fisher가 제시한 분류함수(Classification function)방법이 판별함수를 이용하여 집단을 판별하는 것보다 쉽기 때문에 많이 사용하고 있다. Fisher의 선형 판별함수에 변수값을 대입하여 분류함수 중 가장 큰 값을 갖는 분류집단에 해당하는 집단으로 판별하는 방법이다. 본 분석에서는 Fisher의 분류함수를 이용하였으며, 분류함수는 표 8과 같다.

$$z1(군집-1)=0.001*d_{per}-6.162 \quad (2)$$

표 7. 정준판별함수 비표준화계수

변수	정준판별함수계수
d_per	0.001
(상수)	-1.485

주) d_per : 일승차인원

표 8. Fisher의 분류함수계수

구분	군집-1	군집-2	군집-3	군집-4	군집-5
d_per	.001	.002	.002	.002	.005
(상수)	-6.162	-1.157	-6.164	-3.698	-16.617

주) d_per : 일승차인원

$$z2(군집-2)=0.002*d_{per}-1.157 \quad (3)$$

$$z3(군집-3)=0.002*d_{per}-6.164 \quad (4)$$

$$z4(군집-4)=0.002*d_{per}-3.698 \quad (5)$$

$$z5(군집-5)=0.005*d_{per}-16.617 \quad (6)$$

5개의 군집 중 환승량이 가장 적은 집단 군집-2는 228개 중 223개가 분류되어 97.8% 정확도를 보이고 있으며, 모형 전체 판별력은 86.8%로 나타났다.

4.2.2 도시철도에서 버스로 환승하는 경우 판별 분석

4.2.2.1 군집분석

버스에서 도시철도로 환승하는 경우와 동일한 방법으로 분석을 실시하였다.

표 9. 분류결과

구분		군집	예측 소속집단					전체
			1	2	3	4	5	
원래값	빈도	1	0	1	0	0	0	1
		2	0	223	0	0	5	228
		3	0	3	0	0	0	3
		4	0	24	0	0	0	24
		5	0	1	0	0	1	2
	%	1	.0	100.0	.0	.0	.0	100.0
		2	.0	97.8	.0	.0	2.2	100.0
		3	.0	100.0	.0	.0	.0	100.0
		4	.0	100.0	.0	.0	.0	100.0
		5	.0	50.0	.0	.0	50.0	100.0

표 10. 일승차인원 군집분석

분석기법	실루엣 크기	군집 수
TwoStep	0.767	2
k-means	0.702	5
Kohonen	0.542	12

3가지 군집분석 방법 중 실루엣 크기가 가장 크게 나타난 TwoStep 방법을 채택하였다. TwoStep 방법은 2군집으로 군집을 나누었으며, 군집중심은 각각 1998.2, 388.48로 나타났으며, 샘플수는 각각 105, 533개로 분류되었다. 2군집 중 일승차인원이 많은 군집-1에 대하여 환승량을 그룹화하였다.

TwoStep 방법이 0.816으로 실루엣 크기가 가장 높게 나타났으나, k-means 방법과 큰 차이도 없고, 실루엣 크기도 상당히 설명력이 있기 때문에 버스에서 도시철도로 환승하는 경우와 동일한 군집수를 유지해주기 위하여 k-means 방법을 채택하였다. 그에 따른 적정 군집 수는 5개로 나타났다.

군집1은 군집중심이 1220.5로 환승량이 세 번째로 적은 군집이며, 군집2는 군집중심이 62.41로서 가장 많은 정류장이 분류되었다. 군집3은 군집중심이 2115로 환승량이 가장 많은 군집이며, 도시철도 2호선의 종점역이다. 군집4는 군집중심이 578.17로 환승량이 두 번째로 적은 군집이며, 군집5는 환승량이 두 번째로 많은 군집이다. 도시철도 1호선 동쪽 종점과 남서측 종점 부근에 위치하고 있다.

4.2.2.2 판별분석

표 11. 환승량 군집분석

분석기법	실루엣 크기	군집 수
TwoStep	0.816	2
k-means	0.814	5
Kohonen	0.66	11

표 12. Wilks의 람다

함수의 검정	Wilks의 람다	카이제곱	자유도	유의확률
1	0.761	28.212	8	.000
2	.912	9.540	3	.023

일승차인원과 도심거리 2개의 변수를 사용하여 판별분석을 실시한 결과 2개의 함수 모두 유의확률이 0.05 이하로써 집단 간 평균의 차이가 있으며, 통계적으로 유의하다고 볼 수 있다. 고유값의 비율(분산비율)이 첫 번째 함수는 설명력이 67.2%이고 두 번째 함수의 설명력은 32.8%이다. 그리고 정준상관계수는 고유값과 비슷한 의미를 가지는데 1에 가까울수록 설명력이 크다는 것을 의미한다. 따라서 첫 번째 함수의 설명력이 더 크다.

도시철도에서 버스로 환승하는 경우의 정류장들에 대하여 판별분석을 실시한 결과 모형은 다음과 같은 식으로 표현된다.

$$z1(\text{군집-1})=0.004*\text{일승차인원}+0.563*\text{도심거리}-11.473 \quad (7)$$

$$z2(\text{군집-2})=0.002*\text{일승차인원}+0.438*\text{도심거리}-2.780 \quad (8)$$

$$z3(\text{군집-3})=0.003*\text{일승차인원}+1.156*\text{도심거리}-14.913 \quad (9)$$

$$z4(\text{군집-4})=0.002*\text{일승차인원}+0.407*\text{도심거리}-5.090 \quad (10)$$

$$z5(\text{군집-5})=0.002*\text{일승차인원}+1.016*\text{도심거리}-11.910 \quad (11)$$

5개의 군집 중 환승량이 가장 적은 집단 군집-2는 81개

표 13. 고유값

함수	고유값	분산의 %	누적 %	정준 상관
1	.198	67.2	67.2	.406
2	.097	32.8	100.0	.297

표 14. Fisher의 분류함수계수

구분	군집-1	군집-2	군집-3	군집-4	군집-5
d_per	.004	.002	.003	.002	.002
cen_dist	.563	.438	1.156	.407	1.016
(상수)	-11.473	-2.780	-14.913	-5.090	-11.910

주) d_per : 일승차인원

cen_dist : 도심으로부터의 거리

그림 6. 도시철도-버스 군집별 환승량도

표 15. 분류결과

구분	군집	예측 소속집단					전체
		1	2	3	4	5	
원래값	빈도	1	1	3	0	0	4
		2	3	78	0	0	81
		3	0	1	0	0	1
		4	1	19	0	0	20
		5	0	2	0	0	2
	%	1	25.0	75.0	.0	.0	100.0
		2	3.7	96.3	.0	.0	100.0
		3	.0	100.0	.0	.0	100.0
		4	5.0	95.0	.0	.0	100.0
		5	.0	100.0	.0	.0	100.0

중 78개가 분류되어 96.3% 정확도를 보이고 있으며, 모형 전체 판별력은 73.1%로 나타났다.

4.3 도시철도 3호선 환승량 추정

도시철도 3호선역 주변 반경 500m 이내에 위치한 버스정류장 280개소를 수집하여 환승량을 추정하였다. 버스에서 도시철도로 환승하는 경우와 도시철도에서 버스로 환승하는 경우 모두 도출된 판별함수를 이용하여 분류한 결과 환승량이 가장 적은 군집으로 분류되었다.

따라서 환승량이 동일한 분포를 따를 것이라는 가정 하에 도시철도 3호선역 주변 버스정류장의 일승차인원을 기준으로 환승량을 분포시켰다. 일승차인원은 환승량과의 관계에 있어서 선형관계를 나타내고 있어 일승차인원을 기준

표 16. 분류된 군집의 환승량 분포

구분	최소값	최대값	평균	표준편차
버스-도시철도	1	150	23.07	31.38
도시철도-버스	1	280	66.73	78.43

표 17. 추정된 환승량(상위 10개소)

구분	일승차인원	버스-도시철도 환승량	도시철도-버스 환승량
그랜드호텔앞	1288	189	482
섬유회관앞	1030	151	387
명덕역(3번출구)	1013	149	381
팔달시장앞	996	146	375
명덕역(4번출구)	906	133	342
서문시장	897	132	338
명덕네거리1	887	130	335
큰장네거리1	852	125	322
섬유회관건너1	825	121	312

으로 환승량을 분포시키는 데 문제가 없을 것으로 판단하였다.

4.4 환승량 순위

교통카드자료를 이용하여 수집된 환승자료와 추정된 도시철도 3호선 환승량을 기준으로 환승량도를 나타내면 그림과 같다. 환승량도에서는 인접한 환승정류장들에 대하여 환승량을 합하여 도식화하였다. 대규모심의 약령시 정류장이 환승량이 가장 많은 것으로 나타났으며, 도시철도 2호선의 종점인 사월역이 두 번째로 많은 것으로 나타났다.

표 18. 환승정류장의 환승량 순위와 비율

환승량 순위	구분	환승량			환승비율(%)			합계
		버스-도시철도	도시철도-버스	버스-버스	버스-도시철도	도시철도-버스	버스-버스	
1	약령시앞	1,028	0	1,235	45.43%	0.00%	54.57%	2,263
2	사월역(2번출구)	0	2,115	0	0.00%	100.00%	0.00%	2,115
3	안심역(4번출구)	0	1,726	0	0.00%	100.00%	0.00%	1,726
4	상인현대맨션건너	0	1,702	0	0.00%	100.00%	0.00%	1,702
5	경상감영공원앞	948	0	668	58.66%	0.00%	41.34%	1,616
6	아양교역(1번출구)	900	590	0	60.40%	39.60%	0.00%	1,490
7	사월역(4번출구)	1,464	0	0	100.00%	0.00%	0.00%	1,464
8	두류역(달성고등학교앞)	282	439	737	19.34%	30.11%	50.55%	1,458
9	감삼네거리2	325	375	710	23.05%	26.60%	50.35%	1,410
10	약령시건너	0	1,368	0	0.00%	100.00%	0.00%	1,368
11	그랜드호텔앞	0	1,349	0	0.00%	100.00%	0.00%	1,349
12	대곡역(1번출구)	0	1,319	0	0.00%	100.00%	0.00%	1,319
13	만촌육교1	0	1,101	0	0.00%	100.00%	0.00%	1,101
14	아양교역(2번출구)	0	1,094	0	0.00%	100.00%	0.00%	1,094
15	대곡역(2번출구)	912	0	0	100.00%	0.00%	0.00%	912
16	명덕네거리1	0	897	0	0.00%	100.00%	0.00%	897
17	서부정류장 (LG전자서비스센터)	0	894	0	0.00%	100.00%	0.00%	894
18	영대병원역	0	861	0	0.00%	100.00%	0.00%	861

그림 7. 환승량도

5. 결 론

본 연구는 대구시의 교통카드자료를 기반으로 환승정류장의 특성을 파악하고, 환승정류장의 적정 분류를 통한 입지 특성을 제시하여 원활한 환승정류장 체계를 구축하고자 하는 것이다. 대구시 교통카드 자료는 서울시에서 운영하고 있는 카드자료의 형태와는 다르게 카드ID, 결제시간, 차량번호, 운송회사, 노선번호에 대해서만 정보를 가지고 있다. 따라서 승차정류장을 파악하기 위해 BMS 자료의 하나인 차량별 버스정류장 도착 시간을 수집하여, 교통카드자료와의 Matching을 시도하였다.

환승정류장의 개발을 위한 규모를 고려하기 위하여 1차적으로 일승차인원에 대한 군집분석을 실시하여 이용자들이 많은 정류장을 분류하였다. 그리고 2차적으로 이용자가 많은 정류장 중 환승량에 대한 군집분석을 실시하여 환승량 규모별 군집을 나누었다. 비계층 군집분석은 분석가가 사전에 군집수를 지정한다는 단점이 있지만 본 연구에서는 실루엣 크기를 지표로 하여 적정 군집수를 정하도록 하였다.

버스에서 도시철도로 환승하는 정류장들을 대상으로 일승차인원에 대하여 군집분석을 실시한 결과 5군집으로 나누었으며, 다시 일승차인원이 많은 군집을 대상으로 환승량에 대하여 군집분석을 실시하였다. 각 군집분석 방법 중 k-means 군집분석으로 실시하여 5군집일 때 실루엣 크기가 0.867로 가장 크게 나타나, 5군집을 적정 군집수로 지정하고 판별 분석을 실시하였다. 도시철도에서 버스로 환승하는 정류장들 역시 동일한 5군집으로 나타나, 5개의 환승량 군집을 종속변수로 하는 판별 분석을 실시하였다.

두 가지 환승 유형에 대하여 판별분석을 실시하였고, 버스에서 도시철도, 도시철도에서 버스로 환승하는 경우 정확도는 각각 86.8%, 73.1%로 나타났다. 도출된 판별함수를 기준으로 도시철도 3호선에 대하여 환승량을 추정한 결과 두

가지 유형 모두 환승량이 가장 적은 집단으로 분류되었다.

일승차인원과 환승량을 고려하여 환승정류장의 설치 지점을 제시하였으며, 설치 가능한 지점들을 대상으로 지리적인 위치와 환승형태의 비율을 고려하여 대구시 실정에 맞는 환승정류장의 입지를 구분 할 수 있다. 본 연구에서는 대구시의 교통카드자료만을 이용하여 분석하였기 개인교통수단과 대중교통수단 간의 비율을 고려하지 못하는 한계가 있다.

감사의 글

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2011-0026412)

참고문헌

- 권영중, 오재학(2004) 대중교통지향형 도시개발과 교통체계 구축방안, 교통개발연구원 2004-03.
- 문영일, 노정현(2011) 서울시 역세권의 TOD환경과 대중교통이용 수요 관계분석, 한국도로학회 논문집, 한국도로학회, 제13권 제4호, pp. 211-220.
- 박서현, 황기연(2007) 대중교통지향형 도시개발(TOD)의 성공요건, 대한국토·도시계획학회 추계정기학술대회 논문집, 대한국토·도시계획학회.
- 박지형(2008) TOD 계획요소별 대중교통 이용효과 분석, 한양대학교 대학원.
- 박진영, 김동준(2006) 대중교통정책수립에 있어서 교통카드 자료 활용방안 연구, 한국교통연구원.
- 성현곤, 김동준, 박지형(2008) 서울시 역세권에서의 토지이용 및 도시설계특성이 대중교통이용증대에 미치는 영향 분석, 대한교통학회지, 대한교통학회, 제26권 제4호, pp. 135-147.
- 신성일(2007) 교통카드 Data를 활용한 대중교통 평가체계 구축방안, 서울시정개발연구원, 시정연 2007-R-09.
- 안광호, 임병호(2006) SPSS를 활용한 사회과학조사방법론(2006 개정판).
- 오영택(2008) 대중교통지향형 도시개발을 위한 역세권 개선 방향에 관한 연구, 한양대학교 도시대학원.
- 원태연, 정성원(2010) SPSS PASW 18.0 통계조사분석, 한나래출판사.
- 이혜림(2010) K-평균법의 군집개수 선택기준에 관한 비교 연구, 동국대학교 대학원 통계학과.
- 허명희, 이용구(2008) 데이터마ining 모델링과 사례 제2판, 한나래출판사.
- Clementine Manual. SPSS Korea 컨설팅팀.
- Calthope, F. (1993) The Next American Metropolis: Ecology, Community, and the American Dream, Princeton Architectural Press.
- Kaugamanm, L. and Rousseeuw, P. J. (1990) Finding groups in data: an introduction to cluster analysis, New York, Wiley.

(접수일: 2012.5.14/심사일: 2012.5.21/심사완료일: 2012.9.19)