

빅데이터 분석을 통한 서울시 지하철 지연 분석

박수민⁰, 백재순^{*}, 김성진(교신저자)^{*}

⁰명지전문대학 ICT융합공학과,

^{*}명지전문대학 ICT융합공학과

e-mail: parksoomin1201@mjc.ac.kr⁰, {hisoon99, ict214548}@mjc.ac.kr^{*}

Seoul Subway Delay Analysis through Big Data Analysis

Soo-Min Park⁰, Jai-Soon Baek^{*}, Sung-Jin Kim(Corresponding Author)^{*}

⁰Dept. of ICT Convergence Engineering, Myongji College,

^{*}Dept. of ICT Convergence Engineering, Myongji College

● 요약 ●

본 논문은 진접선 개통 이후 급증하는 서울 지하철 4호선의 혼잡 문제와 현재 진행 중인 장애인 차별 반대 시위를 다룬다. 네이버의 지도 API를 활용해 위도와 경도 데이터를 추출하고 지하철 노선별 장애인 승객 수와 최대 지연시간을 시각화한다. 2호선과 4호선의 혼잡도가 표시되어 문제의 심각성을 알 수 있다. 평균 출퇴근 시간 탑승 및 하차 수치를 분석하여 4호선 편의시설 개선, 2·4호선 열차 운행 횟수 늘리기, 환승역 운영 최적화 등 전략적 권장 사항을 제시한다. 제안된 대책은 서울시 지하철 시스템의 접근성 향상, 혼잡완화, 전반적인 효율성 제고를 통해 보다 폭넓은 교통시설 개선과 승객 편의 증진에 기여하는 것을 목표로 하고 있다.

키워드: 지하철 지연(Subway Delays), 혼잡(Congestion), 데이터 분석(Data Analysis)

I. Introduction

서울 지하철 4호선은 진접선 개통 이후 혼잡도가 162%에서 186%로 증가했는데, 이는 현재 1~8호선 중 가장 혼잡한 상황이다. 공사는 이번 조치로 혼잡도가 167%로 감소할 것으로 예상하고 있다[1]. 전국장애인인권연대는 2021년 12월부터 삼각지역과 4·5호선 왕십리역에서 장애인 차별에 항의하는 시위를 이어오고 있다[2]. 이태원 사태 이후 안전 우려에 대응하기 위해 서울시는 서울교통공사와 협력해 혼잡한 지하철역을 점검해 혼잡도가 높은 곳을 드러낸다[3]. 매년 증가하고 있는 서울시 교통 문제에 대한 서울 시민들의 불만은 계속 늘어나고 있다. 본 논문에서 교통 지연 문제를 개선하기 위해 특정 호선 및 역 방면으로 시각화한다. 시각화를 통해 공통점을 찾아낸 후 결과를 바탕으로 지연을 감소시킬 수 있는 개선방안을 제시한다.

II. Preliminaries

조성근(2015)의 연구에서는 서울지하철 2호선 열차의 지연 및 혼잡 해소방안을 제시하고 있다. 2호선의 만성적 지연을 해소하기 위하여 2호선 혼잡구간 및 시간대에 성능이 우수한 ATO 열차를 연속배치하거나 승하차를 돕는 질서도우미를 배치하는 것의 효과성을 살펴보고 있다[4]. 이성희(2022)는 지하철의 혼잡 관리 모델을 통해 정차시간 절감 효과 분석에 대해서 살펴보고 있다[5].

III. The Proposed Scheme

1. 위도 및 경도

네이버에서 제공하는 지도 API를 사용하여 위도, 경도 데이터를 추출한다. 아래의 Fig. 1처럼 Python 기반의 Corab에서 네이버 지도 API URL을 불러온 뒤 API 요청 과정을 시작한다. 파라미터 설정과 JSON 데이터 파싱을 해준 뒤 나의 API 인증정보인 ID와Secret을 입력한다. 주소 입력창을 Input 설정을 하고 지오코딩 함수를 호출한 뒤 위도 및 경도 데이터를 추출한다. 위도 및 경도 값을 알고 싶은 장소의 도로명 주소를 입력하면 입력한 해당 도로명 주소와 위도 및 경도 값의 정보를 순서대로 나타내준다.

주소를 입력해주세요: 경기 남양주시 별내3로 345
주소: 경기도 남양주시 별내3로 345
위도: 37.6678115
경도: 127.1160450

Fig. 1. Extract latitude and longitude

2. 시각화

1.1 지하철 호선별 장애인 탑승객 수

Fig. 2처럼 지하철 1~9호선 별로 2021년 12월 ~ 2023년 9월까지 데이터를 그룹화 한 뒤 장애인 탑승 승객수의 총합을 계산하고 막대그래프로 시각화한다.

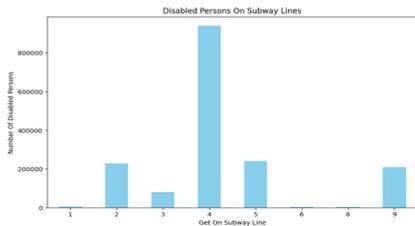


Fig. 2. Number of disabled passengers by subway line

1.2 지하철 호선별 최대 지연 시간

위 Fig. 2를 바탕으로 아래의 Fig. 3처럼 지하철 2, 4, 5호선 별로 2023년 1월 ~ 9월까지 데이터를 각 호선 별로 최대 지연 시간 평균을 계산하고 막대그래프로 시각화한다.

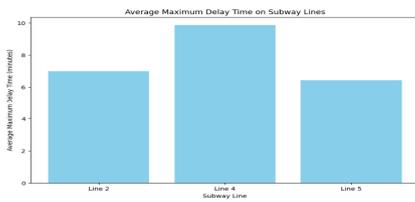


Fig. 3. Average maximum delay time by subway line (minutes)

2.1 2호선 혼잡도 지도

아래의 Fig. 4처럼 혼잡도 데이터에서 2호선에 해당하는 데이터를 필터링한다. 색상을 지정하고 지도의 중심 좌표를 설정한다. Folium 맵과 MarkerCluster 레이어를 생성한다. 혼잡도에 따라서 100단위로 마커가 추가되도록 반복문을 설정한다. 혼잡도 값에 따라서 마커 색상을 지정해준 뒤 마지막으로 맵을 저장하고 Html로 나타낸다.



Fig. 4. Line 2 congestion map

2.2 4호선 혼잡도 지도

Fig. 4와 같이 앞서 설명해드린 진행방식과 같기 때문에 똑같이 진행하되 4호선에 해당하는 데이터를 필터링하여 Fig. 5처럼 나타낸다.

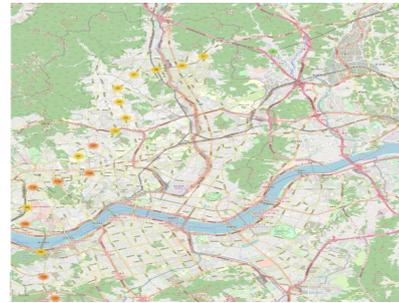


Fig. 5. Line 4 congestion map

3.1 출퇴근시간 승하차 평균 승객 수

아래의 Fig. 6처럼 지하철 1~9호선 별로 2023년 1월 ~ 9월까지 출퇴근시간 승하차 승객 수 데이터를 승차 승객과 하차 승객 값을 더한 뒤 2로 나눠서 평균값을 계산한다. 출근시간 승하차 평균 승객 수와 퇴근시간 승하차 평균 승객 수를 색깔로 구분하여 이중 막대그래프로 시각화한다.

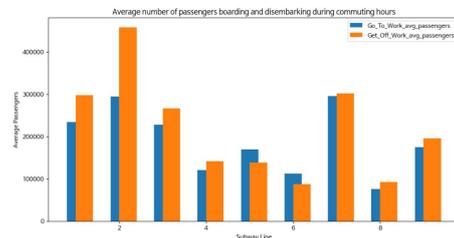


Fig. 6. Average number of passengers boarding and disembarking during rush hour

3.2 2호선 출퇴근시간 승하차 평균 승객 수

Fig. 6을 바탕으로 아래의 Fig. 7처럼 2호선 기준으로 세분화하여 2호선의 환승역들의 출퇴근시간 승하차 평균 승객 수를 나타낸다. 다른 점은 추가로 2호선만 데이터 필터링 과정을 거치고 데이터프레임으로 저장한다는 점이다.

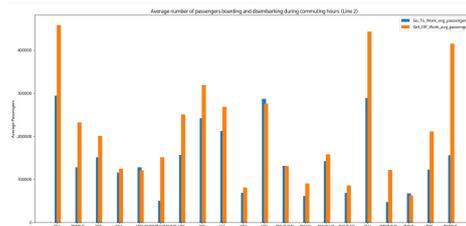


Fig. 7. Average number of passengers getting on and off Line 2 during rush hour

3.3 2호선 환승역 출퇴근시간 지하철 운행 횟수

2호선 바탕으로 지하철 운행 횟수를 알아보기 위해서서울교통공사 사이버 스테이션에서 제공하는 정보를 이용하였다[6]. 서울시 지하철 노선도에서 해당역을 클릭했을 때 역정보, 열차시간표, 내부시설, 주변지도 등 자세한 정보를 제공해주는데 본 논문에서는 열차시간표를 활용하여 2호선의 내선 및 외선의 운행횟수를 직접 Excel을 활용해서 간단하게 표로 시각화 해보았다. 앞서 Fig. 7에서 그래프를 다시 봤을 때 강남역, 잠실역, 홍대입구역, 선릉역, 신림역 총 5개의 역 기준으로 지연을 최소화 하기 위하여 지하철 운행 수를 증가하여야 한다는 이유를 명확히 제시한다. 아래의 Fig. 8은 서울교통공사 사이버 스테이션 정보를 나타내고 Fig. 9는 2호선 출퇴근시간 운행 횟수를 나타낸다.



Fig. 8. Seoulmetro cyberStation Information

지하철역	출근시간(출근시간(퇴근시간(퇴근시간(의선
강남	51 57 50 46
건대입구	51 54 52 46
교대	51 57 50 46
당산	51 53 52 47
대림	49 59 48 45
동대문역사	48 52 50 47
사당	51 57 53 46
선릉	51 57 53 46
영수	51 72 52 67
신당	49 53 49 48
신림	50 59 49 45
명동포구	51 52 52 47
왕신리	49 52 49 47
을지로3가	49 53 50 47
을지로4가	49 53 50 47
잠실	51 54 51 46
중합운동장	51 55 51 47
충정로	50 52 51 47
합정	51 52 51 47
홍대입구	51 53 52 47

Fig. 9. Line 2 Transfer Station Commute Hours Subway Operation Frequency

1.2. 2호선 및 4호선에서의 열차 빈도 증가

2호선 및 4호선에서의 열차 운영을 증가시켜 혼잡을 완화하고 전반적인 지하철 효율성을 향상시킨다.

1.3. 환승역에서의 운영 최적화

러시 아워 동안 환승역에서의 운영을 최적화 하기 위해 승하차를 위한 안내를 제공하고 열차 운영을 신속하게 진행하기 위해 추가적인 직원을 배치한다.

2. 최종 결론

결론적으로 서울시의 교통 시설을 개선하고 최소화하며 승객의 편의성을 향상시키는 큰 목표에 기여할 것으로 예상된다. 접근성, 혼잡도, 러시 아워 운영에 대한 대응으로 이러한 관장 사항을 통합적으로 실현함으로써 효과적으로 서울시 지하철 시스템을 개선에 도움을 기여할 수 있을 것으로 기대된다.

REFERENCES

- [1] Seoul official website. (From : <https://news.seoul.go.kr/traffic/archives/510099>).
- [2] Institute for the Rights of the Disabled. (From :https://www.cowalknews.co.kr/bbs/board.php?bo_table=HB41&wr_id=226).
- [3] World Newspaper. (From : <https://m.segye.com/view/20221102513787>).
- [4] Cho Seong-geun, & Jeong Il-bong. (2015). Study on measures to relieve train delays and congestion on Seoul Subway Line 2. Seoul Urban Research, 16(1), 123-135.
- [5] Seonghee Lee. (2022). [Ph.D. thesis introduction] Analysis of the effect of reducing stopping time using a subway congestion management model. Transportation Technology and Policy, 19(4), 64-64.
- [6] Seoulmetro cyberStation (From : <http://www.seoulmetro.co.kr/kr/cyberStation.do#stationInfo>).

IV. Conclusions

1. 권고 사항

1.1. 4호선의 편의 시설 증가

서울시 내에서 엘리베이터, 에스컬레이터, 휠체어리프트의 전략적인 배치를 촉진하여 접근성을 향상시키고 지연을 최소화한다.