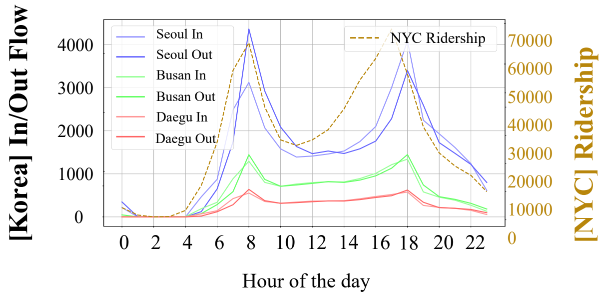
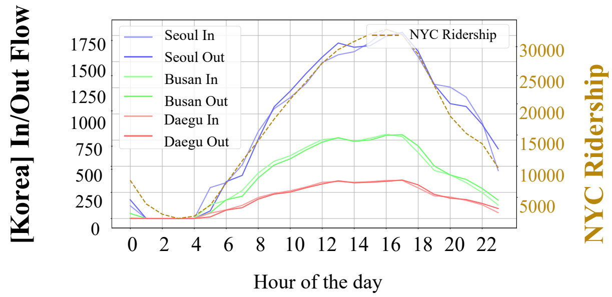
Additional Dataset Strawman Analysis

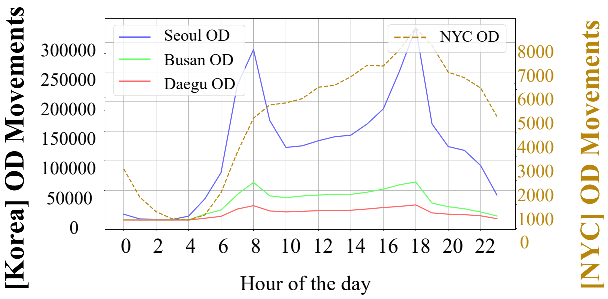
Transportation-[Seoul, Busan, Daegu, NYC] Analysis



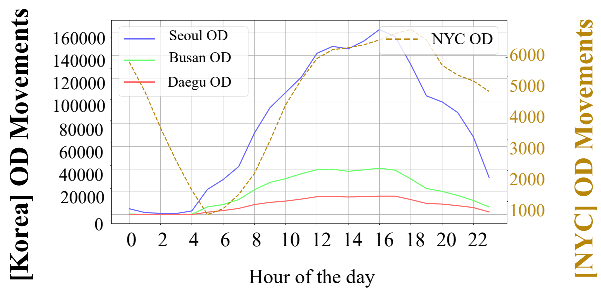
(a) [**Weekday**] Time-Series Patterns.



(b) [**Weekend**] Time-Series Patterns.



(c) [**Weekday**] OD Patterns.



(d) [**Weekend**] OD Patterns.

Figure 8: Temporal transportation patterns to weekdays and weekends. Inflow, outflow, ridership, and OD movements on all nodes are aggregated hourly during the weekdays or weekends.

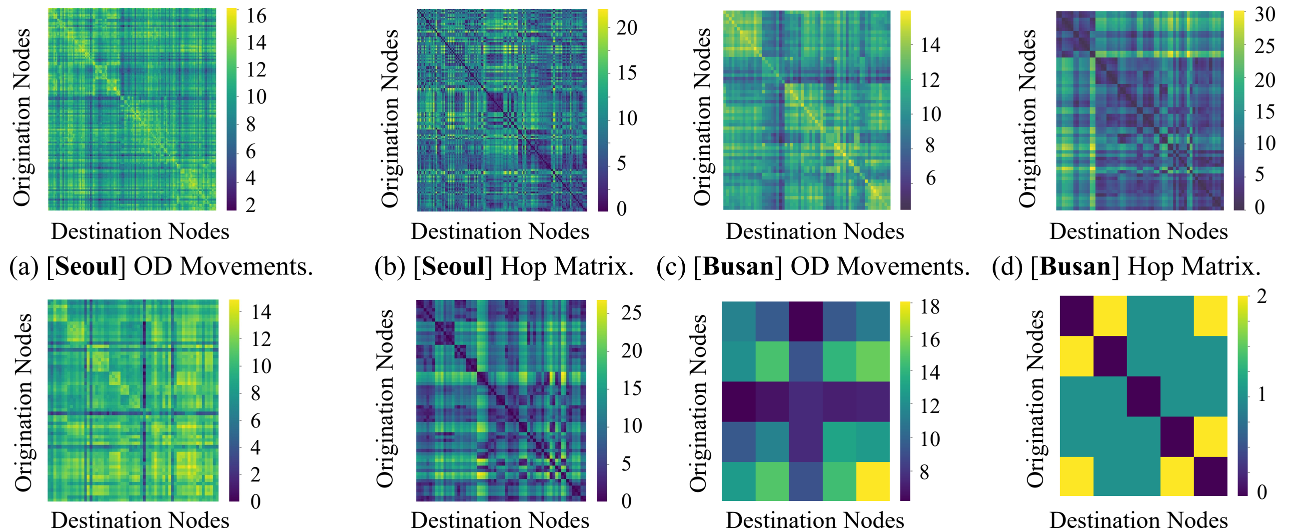
Temporal Aspects: A comprehensive analysis of the transportation datasets was conducted to investigate the temporal patterns during the weekdays and weekends in three major South Korea cities (Seoul, Busan, and Daegu) and New York city, as depicted in Figure 8. Figures 8a and 8b provide insights into the hourly inflow, outflow, and ridership of node time-series features throughout the day. The three Korea cities consistently exhibit temporal patterns of inflow and outflow, with a pronounced bi-modal distribution on weekdays, characterized by peak inflows and outflows around typical commuting hours at 8 a.m. and 6 p.m. This pattern is most notably evident in Seoul. In contrast, the weekend patterns display a smoother uni-modal distribution without sharp peaks, and the weekend demand in these cities decreases by approximately 30 % compared to weekdays. Similarly, New York demonstrates significant peaks in ridership at 8 a.m. and 5 p.m. during the weekdays, with comparable weekend patterns to the Korea cities.

Figures 8c and 8d illustrate the total flows for OD movements throughout the day, revealing that the three Korea cities have more daily OD movements than the node feature time series. The OD movements also exhibit similar bi-modal distribution patterns during weekdays and uni-modal distribution patterns on weekends, consistent with the node time-series features. However, in New York, although the weekend patterns are similar between node time-series features and OD movements, the weekday patterns differ. These differences can be attributed to the source time-series dataset derived from public transportation, such as subway, tram, and railway, which are primarily used for commuting, and the source OD movements dataset from taxis in New York.

Spatial Aspects: Figure 9 represents the relationships between OD movements and the hop matrix. In Busan and Daegu, areas with higher (brighter) OD movements are associated with lower (darker) hops in the hop matrix, indicating a negative correlation in the *Transporation-[Busan, Daegu]* datasets. However, the *Transporation-[Seoul, NYC]* datasets do not exhibit a clear correlation, likely due to the complexity of Seoul's spatial network and the simplicity of NYC's spatial network, making it challenging to discern the relationships between OD movements and the spatial network.

Epidemic-[Korea, NYC] Analysis

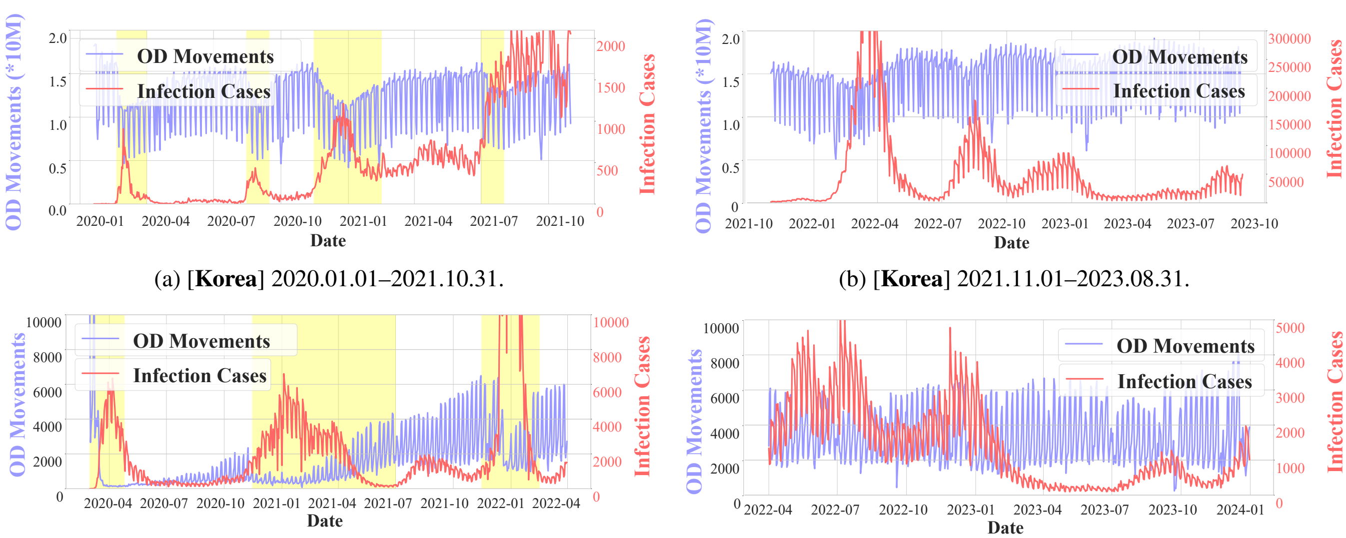
Temporal Aspects: The epidemic datasets, depicted in Figure 10, reveal distinct patterns during the early and later stages of the COVID-19 pandemic. Figures 10a and 10c demonstrate that during the early stages, OD movements and infection cases were negatively correlated, with infection cases increasing as OD movements decreased. This trend is particularly evident during the peaks of infection cases, which correspond to sharp declines in OD movements. However, Figures 10b and 10d illustrate that the



(a) [**Seoul**] OD Movements. (b) [**Seoul**] Hop Matrix. (c) [**Busan**] OD Movements. (d) [**Busan**] Hop Matrix.

(e) [**Daegu**] OD Movements. (f) [**Daegu**] Hop Matrix. (g) [**NYC**] OD Movements. (h) [**NYC**] Hop Matrix.

Figure 9: Spatial patterns of the OD movements and hop matrix in the *Transportation-[Seoul, Busan, Daegu, NYC]* datasets. (a), (c), (e), and (g) display the log-scaled sum of OD movements between nodes. (b), (d), (f), and (h) are matrixes based on hops, indicating the number of nodes to be traversed from one node to another on the spatial network.



(a) [**Korea**] 2020.01.01–2021.10.31. (b) [**Korea**] 2021.11.01–2023.08.31.

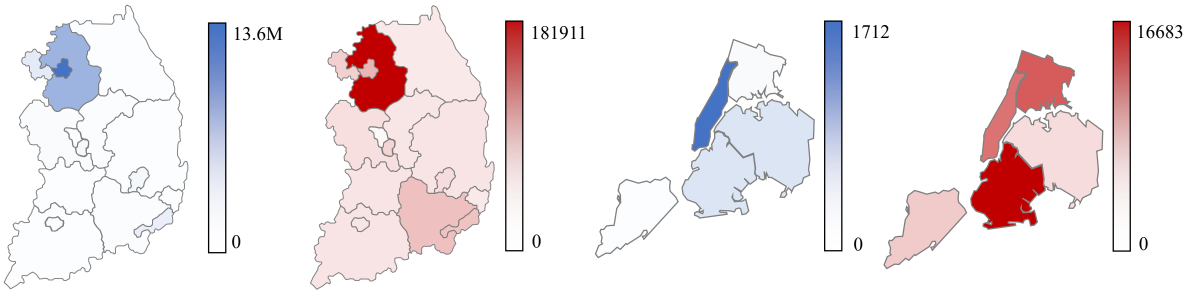
(c) [**NYC**] 2020.03.01–2022.03.31. (d) [**NYC**] 2022.04.01–2023.12.31.

Figure 10: Temporal epidemic and movement patterns in South Korea and NYC. Infection cases and OD movements about all nodes are summed daily (M: million).

later stages of the COVID-19 pandemic exhibit a different pattern compared to the early stages. Despite severe peaks in infection cases, OD movements do not display as strong a correlation as observed in the early COVID-19 period.

Spatial Aspects: Figure 11 presents a comprehensive visualization of all three components of the mobility networked time series on the day when the infection cases peaked for each Epidemic dataset (Korea: 03/17/2022 and NYC: 01/03/2022). The analysis reveals that nodes in close spatial proximity do not necessarily guarantee similar values for OD movements and infection cases. Furthermore, the areas with the highest OD movements do not always correspond to those with the highest infection cases, as observed in both Korea and NYC. However, in the *Epidemic-Korea* dataset, high OD movements typically indicate significant population exchanges in specific regions, which tend to correlate with areas having a high number of infection cases. Conversely, the *Epidemic-NYC* dataset shows a different pattern, where Brooklyn has the highest number of infection cases despite Manhattan having high OD movements.

In summary, the analysis and interpretation of the temporal and spatial aspects of the transportation and epidemic datasets



(a) [**Korea**] OD Movements. (b) [**Korea**] Infection Cases. (c) [**NYC**] OD Movements. (d) [**NYC**] Infection Cases.

Figure 11: Spatial visualizations of the OD movements and infections in the *Epidemic-[Korea, NYC]* datasets. In terms of each node, (a) and (c) display the sum of both total origin movements that belong to the node as the destination and total destination movements that belong to the node as the origination. (b) and (d) are the sum of the infection cases in a maximum infection day (M: million).

highlight the importance of mobility networked time-series forecasting and the need for fusion methodologies. Mobility networked time-series forecasting allows for the development of models that can effectively capture the intricate temporal and spatial dependencies in human mobility data, adapt to evolving patterns and relationships, and provide accurate predictions to support data-driven decision-making in various domains. Moreover, fusion methodologies enable the integration of multiple data sources, providing a more comprehensive and holistic understanding of human mobility patterns and their complex relationships with various factors, such as spatial proximity, connectivity, and external events.