DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Data Engineering and Analytics

Abnormal Events Causal Analysis from Spatio-temporal Human Mobility Network

POOREUMOE KIM

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Data Engineering and Analytics

Abnormal Events Causal Analysis from Spatio-temporal Human Mobility Network

Kausalanalysen abnormaler Ereignisse eines räumlich-zeitlichen Mobilitätsnetzwerks von Menschen

Author:	POOREUMOE KIM
Supervisor:	Prof. Dr.-Ing. Jörg Ott
Advisor:	Leonardo Tonetto
Submission Date:	15.10.2020

I confirm that this master's thesis in informatics: data engineering and analytics is my own work and I have documented all sources and material used.

Munich, 15.10.2020 POOREUMOE KIM

Acknowledgments

I wish first to thank my advisor Leonardo Tonetto: he guided and encouraged me to be professional from start to finish this thesis. With his research advice and feedback, I could consistently steer myself in the right direction.

I also would like to acknowledge the Technical University of Munich. In my first admission letter, they promised to do everything they can to develop my talents and kept the words. They recognized international students' enthusiasm, diligence, and determination and allowed them to research. Without their genuine support, I would not have finished this thesis.

Finally, I must state my sincere heartfelt appreciation to my parents and friends for presenting me with continuous support and continuous encouragement throughout my years of study and research. This achievement would not have been possible without them. Thank you.

POOREUMOE KIM

Abstract

How do social events such as the COVID-19 outbreak and Hong Kong demonstration affect and are affected by human mobility? Some researchers established the impacts of specific social events on human mobility. Others developed algorithms to detect anomalies in a dynamic network. On top of that, we suggest general methodologies that characterize the causality between events and anomalies by a dynamic mobility network analysis. We define three types of anomalies and propose five methods to characterize the three types based on machine learning, statistics, graph theory. Plus, we suggest methods to infer causality based on statistics.

Then, we apply the methods, delivering social interpretation by case studies. We choose two mobility systems: (1) Twitter users' international mobility and (2) COVID-19 virus carriers' mobility in South Korea. They are different systems, having different network properties. Through the case study, we found that

1. Our methods operate given heterogeneous mobility systems
2. The changes of the global mobility network caused by the Hong Kong protest, Russia World Cup, Indonesia earthquake
3. Quantified causality between floating population, virus carriers' traffic, and the number of confirmed cases
4. Pre- and post-signals when the mobility of the COVID-19 virus carriers surges

The result shows that our methods can analyze an abnormal event's cause and effect from a spatio-temporal mobility network and deliver meaningful interpretations for real cases, as mentioned above.

Contents

1. Introduction

When protests in Hong Kong occur, how will the global travel pattern change? Which mobility pattern of the urban population does explain a city's COVID-19 virus transmission? We want to research the causality between unusual social events and human mobility. Thus, our main research question is, **'How do social events affect and are affected by anomalies in a spatio-temporal human mobility network?'**. To answer the main question, we set two research sub-goals: (1) to suggest methodologies that characterize how an event affects and is affected by an anomaly of human mobility, (2) to apply the methods delivering social interpretation by case studies. The case studies analyze Twitter users' global travels and COVID-19 virus carriers' urban mobility in South Korea.

1.1. Terms

Before introducing the contents of our study, we should clarify the following terms:

1. Graph: a graph is an ordered pair $G = (V, E)$ where V is a set of nodes and E is a set of edges. Network is a synonym to graph.

2. Dynamic graph: Harary and Gupta proposed and classified the concept of dynamic graph model[1]. They first set the four entities of a static graph: (1) V, (2) E, (3) f (map vertices to numbers), and (4) g (map edges to numbers). If at least one of them varies by time, they define it as a dynamic graph.

3. Spatio-temporal graph: it is a type of dynamic graph. Each node stands for one spatial location, such as a country or a town. Because the mobility between the nodes varies by time, the set of time series spatial graphs becomes the spatio-temporal graph.

4. Complex system: Rind defined the term as follows: "A complex system is literally one in which there are multiple interactions between many different components"[2]. Because of its straightforwardness, this thesis adopts the definition. Besides, Ladyman et al. researched how scientists use the term 'complex system' and suggested its features[3]. One core feature is the problematic causal relationship, which he quoted as "Complexity starts when causality breaks down"[4]. We also consider the complex system internally involves the complicated causality in this thesis.

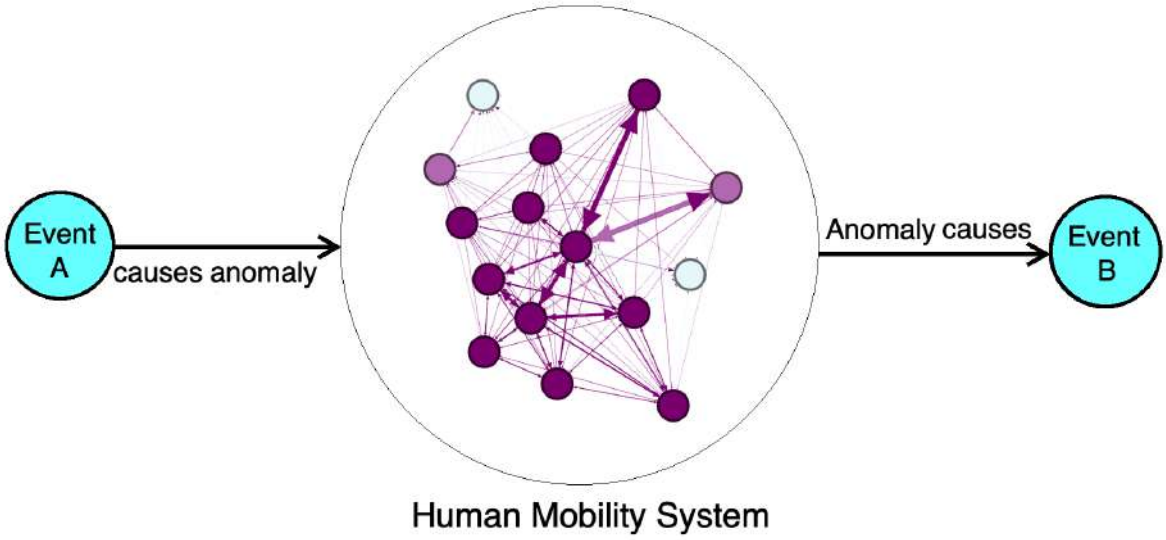

Figure 1.1.: Human mobility system and outside events

5. Exogenous event: an event out of a system that is not caused by the factor inside of a system.

6. Abnormal event: an event occurring irregularly.

According to the definitions, human mobility is also a complex system. Besides, human movement from one place to another place always involves time change. Therefore, a spatio-temporal graph is a structure that can explain such mobility, and that is why we use a dynamic graph. Plus, we use a weighted-directed graph where an edge has both a direction and traffic volume(weight).

1.2. Human mobility system

The relation between an event outside of the system and the system's internal anomaly is directional. The bi-direction means: if an abnormal event occur, it causes abnormal mobility pattern. On the other hand, there exists an opposite case too. The figure 1.1 illustrates the relation.

1.3. Types of anomaly

To understand the system's anomaly, we need to define anomaly in the spatio-temporal mobility network as follows:

1. An outlier regarding a graph metric (type A)
2. An observation different from the past mobility pattern (type B)
3. An observation different from the standard mobility pattern (type C)

One can detect the type A anomaly with data exploration. Besides the naive way, we suggest five methods that can further detect and describe each anomaly type in chapter 3 with statistical techniques to analyze causality between the event and anomaly.

1.4. Heterogeneous mobility cases

After the methods are suggested, we apply them to actual mobility data. The first case study is about Twitter users' inter-country travels. The second case is about COVID-19 patients' urban mobility in South Korea. The two mobility systems are heterogeneous as explained in the table 1.1.

The first difference is the mobility scale. The first case describes the international trips while the second one describes urban mobility. The global travel involves more noises in that even minor local accident affects the mobility system. Therefore, it is more complex and harder to understand causality for an abnormal event. For example, the California wildfire in 2018 must have affected global travel pattern, but we could not see clear causal result from the Twitter geotagging. On the other hand, the patients urban mobility is less complex. We could successfully recognize the casual relationships between floating population, patients mobility and massive epidemic spread.

Next, the Twitter dataset contains trend due to the data collection volume, and seasonality. Some of our analytic methods underperform due to the two time-series property. The confirmed cases of COVID-19 includes only seasonality, which is not the yearly season, but the two waves of massive outbreak.

Also, the Twitter dataset is centered in the US. Most travel takes place in or around the United States. Also, the spatio-temporal graph from the Twitter geotags is more dense. This is because the dynamic graph from COVID-19 data only involves patients, not the normal population.

Plus, there are not many confirmed COVID-19 cases in South Korea, so the second graph is more sparse. We suggest the specific density values in the section 4.2 and 5.2.

Finally, the Twitter geotags were collected from 2016 to 2019. On the other hand, the second case is from Jan.2020 to June.2020.

Mobility system	**Twitter users' travel**	**COVID-19 patients movement**
Scale	International travels	Urban mobility in Seoul
Complexity	High	Low
Trend	Yes	No
Seasonality	Yes	Yes
Hub node	Yes(USA)	No
Graph density	Dense	Sparse
Time interval	2016−2019	Jan−June.2020

Table 1.1.: Comparison of two human mobility systems

2. Related Work

Regarding our research topic, we introduce recent research trends and present what makes our research distinguished from them. There are four relevant fields: (1) anomaly detection from a dynamic graph, (2) causal inference from time series data, (3) mobility analysis through Social Network, and (4) human mobility research considering COVID-19. Theoretical details are in chapter 3, and the applications are in chapters 4 and 5.

2.1. Anomaly analysis from a dynamic graph

According to Ranshous et al.'s article, they summarized the conventional approaches for anomaly detection in a dynamic network[5]. There are commonly two stages for the analysis. The first is mapping a network to real numbers. In the second stage, an anomaly detector is applied, such as Support Vector Machine(SVM) or Local Outlier Factor(LOF). Mathematical expressions are as follows:

$$\Omega \colon D \to \mathbb{R}^d \tag{2.1}$$

$$f \colon \{[\mathbb{R}^d]^*, [\mathbb{R}^d]^t\} \to [0, 1] \tag{2.2}$$

where D is the domain-specific representation, d is the number of feature dimensions, $[\mathbb{R}^d]^*$ denotes the history of the data, and $[\mathbb{R}^d]^t$ does the current time point. The equation 2.2 computes anomaly score. Researchers often use traditional machine learning(ML) ways to implement the two steps.

Our study also apply the following methods: (1) SVM-based anomaly detection by Akoglu and Faloutsosto[6], and (2) tensor rank decomposition with LOF by Lin et al.[7] We use the first to characterize one node's dynamic neighborhood and find an anomaly's representative features. Considering the second, we use it to find a normal mobility state and draw additional insights for abnormality.

Besides the traditional ML way, we characterize type C anomaly by graph prediction using neural network(NN). The approaches have become a notable stream of modern graph prediction to capture the relationships between elements in a complex system. The most prominent algorithm is Graph Convolution Network(GCN) by Kipf and Welling[8]. Briefly, the GCN updates each nodes' features by considering the local neighborhood to model the

relational complexity within a graph. We describe the core operation with the mathematical formula in chapter 3.

On top of the original GCN, researchers developed various applications. They also designed models to forecast the traffic state. But many suggestions are for a static graph such as one adjacency matrix describing non-dynamic roads, which violates our mobility system state. For example, we applied T-GCN suggested by Zhao et al.[9] for traffic prediction, but it only shows at most 0.3090 of accuracy. In this test, we assumed the static adjacency matrix whose element is zero if there is no traffic in 2018 2019 from Twitter geotags database. The result indicated that the models including the static assumptions are not feasible.

On the other hand, there are models for dynamic graph prediction. We mainly use models of Pareja et al.[10] Although their model successfully predicts link existence in the future, it does not work for a sparse graph.

Other studies do not provide models for a weighted, directed, dynamic graph. For instance, Lei et al. suggested the GCN-GAN model for a weighted dynamic graph[11], which did not operate in our case.

In addition to ML and NN methods, our distinctive study applies non-conventional methods to not only detect but further describe anomaly in a spatio-temporal mobility graph. For example, we measure social-geological diversity in a region to characterize abnormalities. Besides, we generalize the gravity model of international trade in international economics to analyze bilateral traffic between nodes.

2.2. Causal inference from time series data

Causal inference is to figure out a causal relationship between cause and effect. When analyzing social data, researchers often cannot experiment given a controlled environment, but they observe what happened. They need to quantify to clarify the causal relations through statistical methods.

One common approach is to infer causality from time-series observations. Initially, Granger suggested a classical and intuitive causality concept for the pair of time series variable[12]. Other modern researchers still practice it thanks to its straightforward definition and description. For example, Jakob introduced this method to analyze a causality of time series data from complex systems[13]. Furthermore, we need to analyze observations over a time interval on many cross-regional units like countries. Therefore we adopt the Hurlin and Dumitrescu's concept[14]. They extended classical Granger causality for the panel data. Finally, Eichler provided a comprehensive description of the causal inference in time series analysis[15], and Pearl wrote a thorough overview of the topic of causal inference[16].

Our analysis uses the typical statistical approach because (1) our observations have ideal

shape to measure Granger causality, and (2) the approaches are intuitive and interpretable. We want to quantify plausible causal relationships between floating populations, the virus carriers' traffic, and the number of new confirmation from COVID-19 dataset.

2.3. Mobility analysis through Social Network

Scientists have researched crowd mobility through Social Network Service(SNS). Primarily, Twitter is an excellent proxy for human mobility because it relies on publicly available data and provides high resolution positioning when users opt to geotags their tweets with their current location[17]. Through Twitter, some researchers studied how people respond to a social event like a natural disaster. For example, Wang and Taylor uncovered spatio-temporal human mobility patterns in different disasters by leveraging geotagged tweets. Martín et al. also collected posts on Twitter and estimated how people evacuated from Hurricane Matthew[18].

On the other hand, our studies suggest and applied anomaly analytics to a dynamic mobility network, describing how abnormal events affect the network structure. To do this, we use the database created by Sen and Dietz[19]. They collected Twitter users' geotags and check-in data on new locations and built the trip database. We transform the data as a dynamic graph, which we described with details in chapter 4.

2.4. Human mobility research considering COVID-19

Since the COVID-19 pandemic hit humanity, researchers have studied mobility regarding the epidemic. We first introduce papers to measure how human mobility changed by the epidemic. Schlosser et al. measured how the mobility network structure varies according to a lockdown policy[20]. Similarly, Galeazzi et al. studied the effect of mobility lockdown in France, Italy, and the United Kindom. They also analyzed the human mobility network over time[21].

In contrast, some articles show that human mobility affects the virus spread. Iacus et al. quantified how the mobility impacts on the virus spread[22]. They focus on building regression models over time and observe how much the mobility explains death cases. Kraemer et al. maintain the control measures implemented in China mitigated the spread of COVID-19 [23].

The above articles indicate the relation between human mobility and the pandemic outbreak. Some researchers further analyzed the causality and found that we may see an early signal before the virus spread or predict it. According to Chu et al. suggestion, network density may be a metric to predict pandemic. They simply measured the network density to assess

connectivity of different countries' confirmed cases' changes[24]. On the other hand, Kapoor et al. used the spatio-temporal mobility network to forecast the outbreak[25].

We distinguish our research from the existing one by three points. First, our study analyzes virus carriers' dynamics(not public mobility) through a spatio-temporal network. Second, we found pre- and post-signals of the outbreak by using anomaly analytics in the network. Third, we interpret/unveil the anomalies from the perspective of graph analysis.

3. Methodologies

This section introduces methods to characterize abnormal events and quantify its causality from a dynamic mobility graph. In the first section, we provide five characterizing methods with pros and cons in the application's perspective. We classified the five ways according to the anomaly types in 1.3. Later, we propose methods for causal inference. This part includes one core approach, Granger causality, and two supporting methods of hypothesis test. The figure 3.1 shows types of these methodologies in a dendrogram format

3.1. Methods for anomaly characterization

3.1.1. Naive data exploration

It is the most basic way to understand anomaly to distinguish type A anomaly. For example, one can measure traffic volume in one place and observe when the volume abnormally changes. The basic graph metric includes a node's in/out degree and weight. We also call the method as naive Exploratory Data Analysis(EDA). It has the following pros and Cons of the method:
Pros of the method:

1. Simplicity: no advanced algorithm is necessary.
2. Essential data understanding: it provides general grasp of a given mobility system.

Cons of the method:

1. Shallow analysis: the naive way does not detect all anomaly.

3.1.2. Social-geological metrics

To recognize type A anomaly, we propose four social-geological metrics quantifying social roles and diversity of areas. Hristova et al. initially suggested four metrics that need social relation network and the geological graph[26]. We adopt and adjust their idea. Instead of a social relation graph, we utilized the locations of travelers' home. We named them as follows: brokerage, popularity, entropy, and homogeneity. To calculate the entropy and homogeneity,

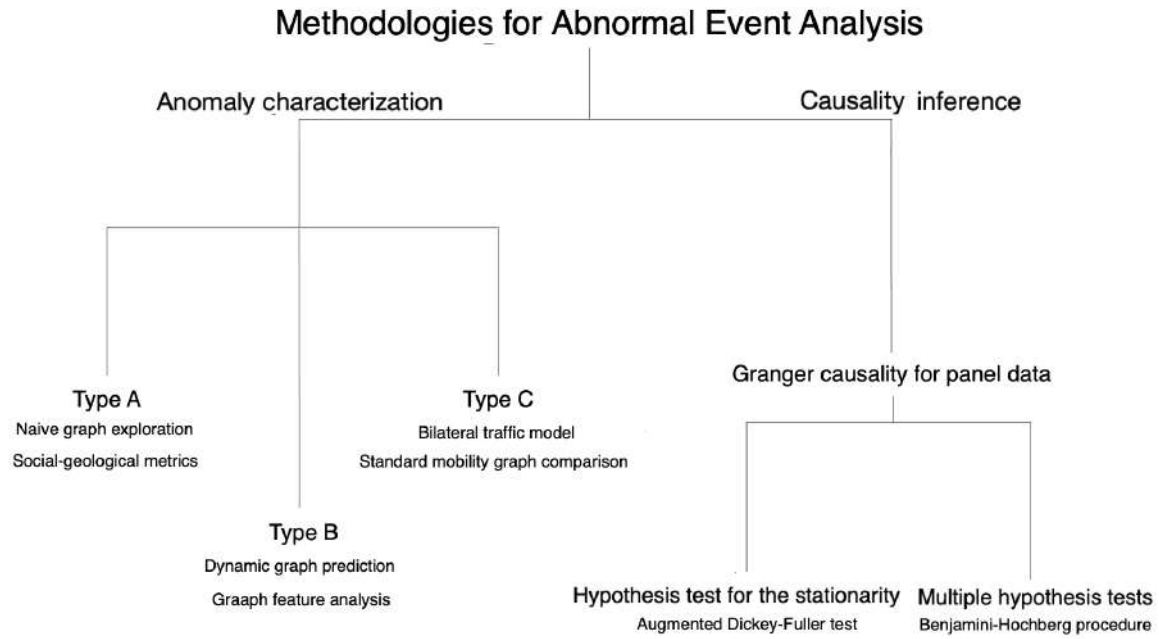

Figure 3.1.: Types of anomaly analysis methodologies

we need a dynamic home-destination graph as well as the original mobility graph like Figure 3.2. In this graph, an edge weight from node u to v means the traffic volume where travelers' home location is node u and the destination is v.

1. Brokerage: it is the between centrality of a node which represents a degree to which nodes stand between others. This thesis named it brokerage because a node with a high between centrality plays as a broker connecting two different regions. In order to design brokerage, the original paper used Burt's redundancy measure that focuses on neighborhood nodes in a social relation network. On the other hand, our brokerage

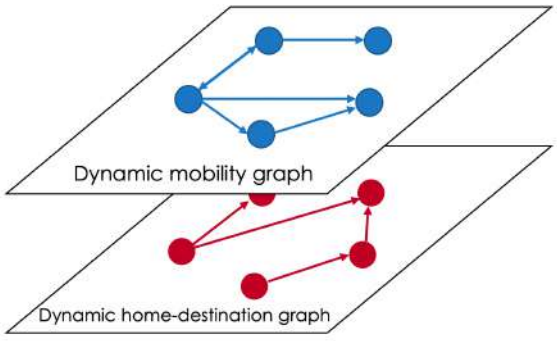

Figure 3.2.: Two dynamic graphs to measure entropy and homogeneity

counts all short paths in graphs and uses people's home information. Also, it does not require the social relation of friends. The formula is same as Brandes' between centrality[27], but we added on time notation:

$$B_t(l) := \sum_{u,v \in V_t} \frac{\sigma(u,v|l)}{\sigma(u,v)} \tag{3.1}$$

where V_t is the set of nodes at time t, $\sigma(u,v)$ is the number of shortest paths from node u to v, and $\sigma(u,v|l)$ is the number of those paths passing through some node l other than u,v. If $u = v, \sigma(u,v) = 1$, and if $l \in \{u,v\}$, then $\sigma(u,v|l) = 0$.

2. Popularity: it is an average incoming traffic of a node over its incoming edges. High popularity means it attracts more people from its neighborhoods on average. The popularity of node l at time t is

$$P_t(l) = \frac{\sum_{u,l \in V_t} W_{u,l}}{\sum_{u,l \in V_t} 1_{\{W_{u,l}>0\}}} \tag{3.2}$$

where $W_{u,l}$ is edge weight between node u and l and $1_{\{.\}}$ is an indicator.

3. Entropy: the entropy of a place describes how diverse visitors' home places are. For example, if a home country of visitors to node A is only country B, it becomes 0. Its mathematical formula is based on Shannon entropy but has time notation:

$$H_t(l) = -\sum_{u \in N} p_t(l,u) log p_t(l,u) \tag{3.3}$$

where $p_t(l,u)$ is a probability that a check-in in place l is made by visitors whose home place is u at time t. N is the set of all home places.

4. Homogeneity: it is a mean similarity between home regions' travel patterns. It is also a social diversity measure of an area. In other words, if a node has a high homogeneity, it means that the visitors' home countries(or cities) have similar travel destinations. The mathematical formula is as follows:

$$S_t(l) = \frac{1}{|V_t(l)|(|V_t(l)|-1)} \sum_{a,b \in V_t(l)}, sim(H^t_{(a,)}, H^t_{(b,)}) \tag{3.4}$$

where H^t is a home adjacency matrix such that H^t_{ij} means the number of visits to node j at time t by travelers whose home place is i, and $V_t(l) = \{i|H^t_{ij} > 0\}$. Furthermore, the similarity measure $sim(u,v) = \frac{1}{D_c(u,v)}$ where $D_c(u,v)$ is a cosine distance between u and v. In our experiment, the inverse of $D_c(u,v)$ distinguishes spatial nodes better than normal cosine similarity.

This study applied the metrics to characterize disturbances in human mobility system and found the following pros and cons.
Pros of the method:

1. Interpretability: each metric has own social-geological implication that researchers can understand in a straightforward way.

2. Robustness: even though a dynamic mobility graph is complex, noisy, sparse, or has seasonality and trend, one can describe an abnormal event with these metrics.

Cons of the method:

1. It requires exogeneous information: one needs additional information about travelers' origin to calculate the entropy and homogeneity.

3.1.3. Dynamic graph prediction

To recognize an anomaly as an observation that differs from the prediction(Type B anomaly), we utilize the dynamic graph prediction. However, forecasting a future graph requires to capture the complicated spatio-temporal dependency between nodes in graphs. Thanks to the deep learning, it is possible to model the structure of the graphs. The fundamental architecture is Graph Convolutional Network(GCN) suggested by Kipf and Welling[8]. They proposed a neural network for an undirected graph. Its core operation is as follows:

$$H^{(l+1)} = GCN(A, H^{(l)}, W^{(l)}) := \sigma(\hat{A}H^{(l)}W^{(l)}) \tag{3.5}$$

where $\hat{A} = \widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}$. It initializes $H^{(0)} = X$ and $H^{(l)}$ is a hidden representation by node embedding. $H^{(l+(1)}$ is recursively defined with $\hat{A}H_t^{(l)}$. Because $\hat{A}$ is an normalized symmetric adjacency matrix, one can interpret the multiplication of $\hat{A}$ to $H^{(l)}$ as an operation that gathers information from all neighbors. Besides, $W^{(l)}$ multiplication transforms the former layer, $H^{(l)}$.

On top of the original GCN, Pareja and others(2019)[10] extended the concept so that a graph can vary by time and be either directed or weighted. The researchers published the code at https://github.com/IBM/EvolveGCN. According to their adaptation, the input graph can be expressed as A_t and X_t where the former is an adjacency matrix, and the latter is a d-dimensional feature vector of the corresponding node at time t. With this t notation, the key layer structure of GCN becomes as follows:

$$H_t^{(l+1)} = GCN(A_t, H_t^{(l)}, W_t^{(l)}) := \sigma(\hat{A}_t H_t^{(l)} W_t^{(l)}) \tag{3.6}$$

where $\hat{A} = \widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}$, $\widetilde{A} = A + I$ and $\widetilde{D} = diag(\sum_j \widetilde{A}_{ij})$.

Furthermore, they created EvolveGCN(EGCN) by adjusting the transformation, $H_t^{(l)}W_t^{(l)}$ where $H_t^{(l)}$ is a hidden representation by node embedding at time t. The EGCN uses

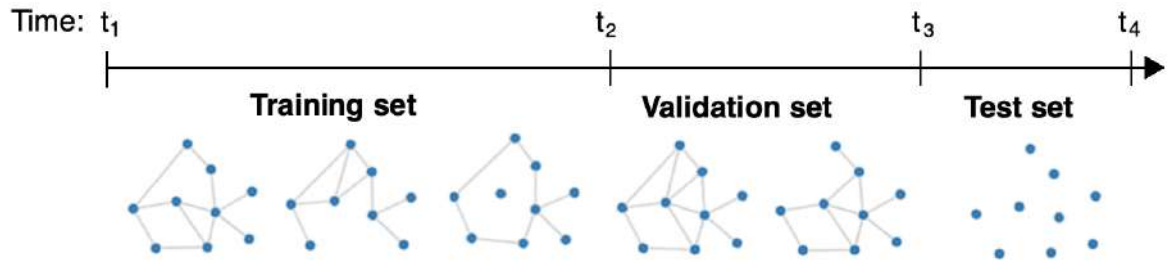

Figure 3.3.: Train, validation, test graph sets for link prediction

gated recurrent unit(GRU) or Long short-term memory(LSTM) instead of the simple matrix multiplication to capture a more dynamic graph sequence. By selecting one of the methods, EGCN forecasts $W_{t+1}^{(l)}$. Specifically, GRU updates the weight in this way:

$$W_{t+1}^{(l)} = GRU(H_{t+1}^{(l)}, W_t^{(l)}) \tag{3.7}$$

However, LSTM updates $W_{t+1}^{(l)}$ as follows:

$$W_{t+1}^{(l)} = LSTM(W_t^{(l)}) \tag{3.8}$$

If EGCN selects the former strategy, it is called as EGCN-H, while the other is called EGCN-O. They argued that the adaptations outperform the original GCN with experiments. This thesis also checked their superiority, but the best F1 score is only bigger by 0.015. The detailed experimental results can be found in the Twitter data analysis part. Besides, EGCN supports three types of tasks: (1) link prediction, (2) edge classification, and (3) node classification. The task of link prediction is to leverage information up to time t and predict the existence of an edge (u, v) at time $t + 1$. The second one is to predict the label of an edge (u, v) at time t. The last one is the label prediction of a node u at time t.

To train the model, we need to divide the spatio-temporal graph into three sets: train, validation, test graph sets. The figure 3.3 shows this separation for the link prediction. The algorithm has to estimate if there is a link between nodes.

After making predictions, we will compare the result to the historical graph by a scaled graph density. The formula for the density is as follows:

$$density'(G_t) = \frac{density(G_t) - min(D_T)}{max(D_T) - min(D_T)} \tag{3.9}$$

where $D_T = \{density(G_t) | t = 1, ..., T\}$. In addition, $density(G_t) = \frac{1}{N(N-1)} \sum_{u,v \in G_t} e_t(u, v)$ where G_t refers to a graph at time t whose number of nodes is N, Also, $e_t(u, v)$ becomes 1 if there is a directed edge from u to v within G_t. The reason we scale the measure is to prevent abnormal difference between the predicted and actual values.

Because we implement link prediction, we will also compare a predicted single node to the real one by their degrees. Again, we scale the measure as follows:

$$degree'(v)_t = \frac{degree(v)_t - min(D'_T)}{max(D'_T) - min(D'_T)} \tag{3.10}$$

where $D'_T = \{degree_t(v)|t = 1, ..., T\}$ and $degree(v)_t = \sum_{u \in G_t}(e_t(u, v) + e_t(v, u))$ is a single node.

Finally, we will visualize the graph and its community structure to compare the predicted graph and historical one . For visualization, we apply the spring layout by Fruchterman-Reingold force-directed algorithm[28]. The method is a classical and popular force-based methods, cited 5,830 times as of October 5, 2020. It embeds nodes and edges into two-dimensional space so that all the edges are of more or less equal length, and there are as fewer crossing edges as possible[29]. We used a python library, NetworkX, for the visualization.

For the community clustering, the Leiden algorithm is used[30]. The method is designed for a weighted graph, and its community structure is similar to the nations' geological distribution. Newman's eigenvector algorithm is not originally designed for a weighted network, although it can be extended[31][32]. Besides, we tested Girvan–Newman algorithm[33], Reichardt and Bornholdt's one[34], and Traag et al.'s methodology using asymptotical surprise[35]. However, their communities do not distinguish European and North American communities from the Twitter dataset, having the USA and UK in the same cluster according to our test. That is why we chose the Leiden way. At last, we used a python library named CDlib to calculated the clustering[36].

In practice, the graph prediction approach has the following pros and cons.
Pros of the method:

1. Explanatory power: it can capture the complex spatio-temporal dependency so that one can predicts a future graph. Thus, it is also possible to compare various kinds of graph information, including community clusters and link density.

2. Robustness to noise: they successfully has predicted the global travel network in spite of noisy observations in the Twitter data.

3. Robustness to non-stationarity: the algorithms perform regardless of seasonality and trend.

Cons of the method:

1. It requires a large train dataset: small data does not train the model correctly. For example, patient networks from COVID dataset could not train the model which only has 54 nodes and 21-time points.

2. Task limitation: even though it is a state-of-art method, EGCN only supports the three types of tasks: link prediction, edge classification, node classification. It does not tell exact edges with weights.

3. Low interpretability: like a typical deep learning model, it is not easy to understand why the algorithm delivers the results.

3.1.4. Neighborhood feature analysis

We propose a method to detect and characterize the type B anomaly with graph features. In detail, this method detects a drastic change of graph features such as in-weight compared to the past time steps. It is the application of the method of Akoglu and Faloutsosto[6]. They designed the original way to detect an anomaly from a graph. The mathematical algorithm is as follows.

1. Formulate a dynamic graph as a tensor, that is $h(G) = \chi \in \mathbb{R}^{T \times N \times F}$ whose axes express time(T), node number(N), and features(F). $\chi_{t,n,f}$ represents nth node's the fth feature such as in-degree and in-weight of the node at time t.

2. Choose one feature f in the F dimension, which leads to a matrix $\chi_f \in \mathbb{R}^{T \times N}$.

3. From χ_f, slide a window whose size is W x N to generate a series of covariance matrix Cov_t where $t \in \{W...T\}$.

4. Calculate the left singular vector $u(t)$ of Cov_t for all t.

5. Calculate $r(t-1) = \frac{1}{W} \sum_{i=t-W}^{t-1} u(i)$.

6. Measure abnormality scores $z_t = 1 - u(t)^T r(t-1)$.

7. (Optional) compare the elements of $u(t)$ and $r(t-1)$ to find a node to attributing an event.

We can summarize the steps from one to six as $F_f(G) = v_T$ where $v_T = [z_1, z_2, ..., z_T] \in \mathbb{R}^T$. Also, we computed the twelve features of every single node: (1) total out-degree, (2) total out-weight, (3) total in-degree, (4) total in-weight, (5) the number of neighbors, (6) the number of reciprocal neighbors, (7) the number of triangles including the node, (8) average in-weight over the edges (9) average out-weight over the edges, (10) maximum in-weight of a node, (11) maximum out-weight of a node, (12) maximum out- and in-weight ratio. A researcher can designs any features. The figure 3.4 visualized the above steps except the seventh.

We found the following pros and cons of the method from the perspective of its application. Pros of the method:

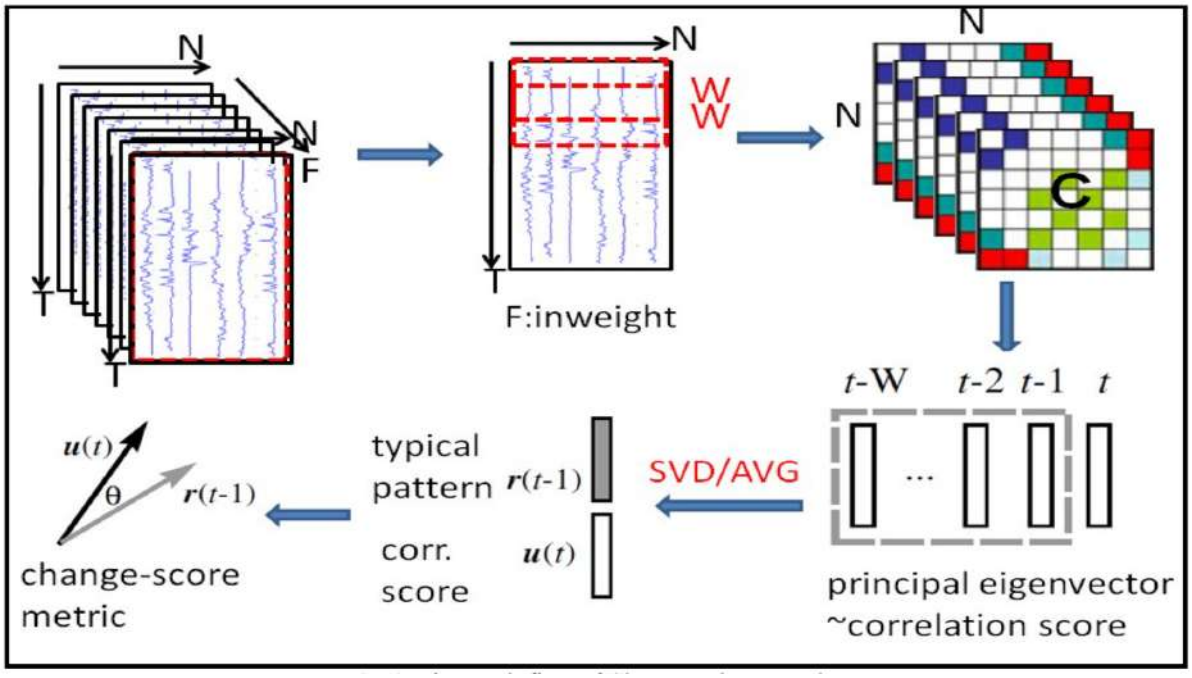

Figure 3.4.: The anomaly analysis with graph feature[6]

1. Simplicity: the computation is naive and straightforward.
2. Feature attribution: one can recognize which graph features represent an anomaly while others do not.
3. Node attribution: one can also attribute a node to an anomaly considering spatiotemporal dependency.
4. Feature engineering: one can create and observe any features for analysis.

Cons of the method:

1. Parameter sensitivity: the result highly depend on the window size.
2. Weak Interpretability: it is hard to understand why features cause the change of anomaly score.
3. Seasonality sensitivity: it regards a seasonal change as anomaly.
4. Low robustness: A noise influences the anomaly scores.

We apply their method in two ways. First, we use it to find when a single node shows anomaly due to an abnormal event. To find the time point, we use the node's neighborhood

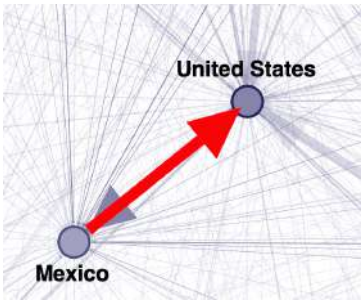

Figure 3.5.: Example of bilateral traffic from Mexico to USA

network, not the whole graph. It is because causality between a single location's event and a neighborhood network' anomaly is straightforward. Second, we use it to find representative features of an event from the neighborhood network. In the first and second case, a researcher can set the G as a dynamic neighborhood network of the target node.

Due to the cons, it is hard to use this method for the first purpose from Twitter users' mobility. In other words, the global human travel network is a complex system with more noises. As a result, the algorithm considers various time points abnormal. Still, it is meaningful to find representative features from major anomaly. On the other hand, the COVID-19 dataset is less complicated. Thus it is more proper for both goals.

3.1.5. Bilateral traffic model

To analyze the type C anomaly, we model a directed flow between two nodes and analyze outliers. We generalize a gravity model, a representative model to predict bilateral flows based on the attributes of origin and destination, and their distance. One example of the flow is in the figure 3.5. The red arrow means directional flow from Mexico to the United States. A basic gravity model is suggested by Fotheringham and Brunsdon[37], and Oshan introduced it as follows [38]:

$$T_{ij} = k\frac{E_i^{\alpha} A_j^{\beta}}{d_{ij}^{\gamma}} \tag{3.11}$$

where T_{ij} is the traffic volume from node i to node j, E_i is the origin node's emission(outgoing traffic), A_j is the attraction(incoming traffic) of the destination node, d_{ij} is the distance between i and j. With logarithm transformation, the equation becomes the following equation:

$$log(T_{ij}) = log(k) + \alpha \cdot log(E_i) + \beta \cdot log(A_j) - \gamma \cdot log(d_{ij}) \tag{3.12}$$

Then, we generalize the equation3.12 with a proper link function $g(x)$ so that the model can describe the mobility pattern even better.

$$g(log(T_{ij})) = log(k) + \alpha \cdot log(E_i) + \beta \cdot log(A_j) - \gamma \cdot log(d_{ij}) \tag{3.13}$$

A different version of the gravity model can use different attributes. For example, Hawelkaa et al. also described the global mobility patterns by using the geo-located Twitter messages and showed their geolocation data is a proxy for global mobility[39]. The difference is that they used populations of origin and destination countries as attributes, so the place needs only one variable, population.

It is also possible to model a causal effect by setting exogenous attributes as explanatory variables and endogenous results as the target variable. For example, we build the model whose explanatory variables are the normal floating population in a city, and the target is the traffic volume of the COVID-19 patients.

One issue is that this research handles spatio-temporal mobility where the attribute E_i and A_i vary by time. To solve this problem, this study averages attributes over the whole time interval. For instance, E_i becomes the average emission of region i and T_{ij} is the average traffic from area i to j. By aggregating, noises are reduced so that the final model can better explain the volatility .

In practice, anomaly characterization with the bilateral traffic model has the following pros and cons.
Pros of the method:

1. Simplicity: the model is straightforward. Therefore it is more interpretable and does not require a complex computation. In this thesis, a simple regression with logarithm transformation sufficiently explains human mobility without more generalized linear model.

2. Exogenous information adoption: compared to the other methods, the gravity model is easier to plug in information from outside of system. For example, this research set floating populations as explanatory variables to estimate COVID carriers' traffic where the population data does not exist in the patients' mobility network.

Cons of the method:

1. It requires manual modeling: like a traditional statistics, a researcher should design the model manually. For instance, the goodness of fit highly depends on the scale of explanatory variables. Also, the underlying distribution should be chosen.

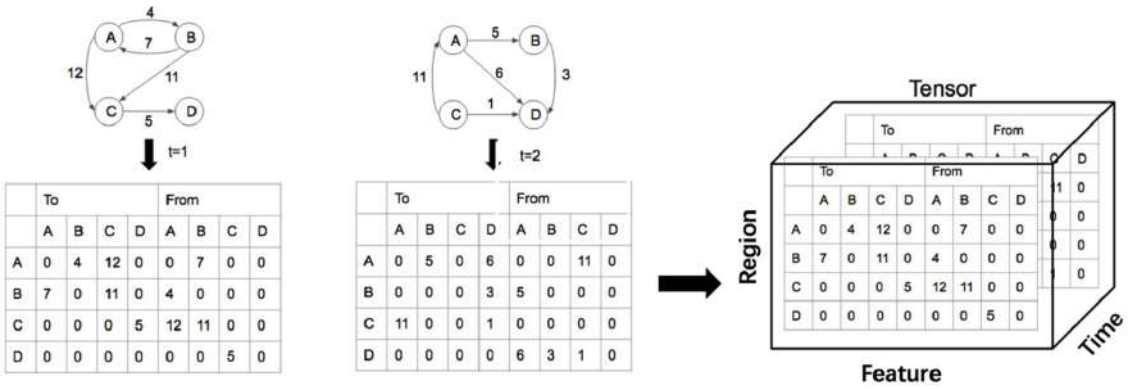

Figure 3.6.: Tensor formulation[7]

2. It only estimates the bilateral flows: the method only computes the traffic between two areas. Thus, if we need to observe and characterize an anomaly of a whole graph, it is not the appropriate way.

3. Non time-series model: the model does not consider time dependency.

3.1.6. Standard mobility graph comparison

We can detect an anomaly mobility pattern by comparing a traffic pattern to a standard state. To compute the optimal one, we use the algorithm proposed by Lin et al. (2018)[7]. Their algorithm localizes abnormal events in spatio-temporal graphs based on tensor decomposition. Their suggestion follows an unsupervised learning procedure in which a model is trained based on samples showing standard traffic patterns.

To understand the method, one has to know how they define a target tensor and decompose it. The tensor $\chi \in \mathbb{R}^{N \times K \times M}$ indicates spatio-temporal graphs, which represents the feature variances of different regions over time. Specifically, $\chi_{i,j,k}$ denotes the number of trips from region i to region j in kth time slice and $\chi_{i,j+N,k}$ denotes the number of trips from region j to region i in kth time. The figure 3.6 suggests a sample for this tensor.
Furthermore, a 3-dimensional tensor can be decomposed, like SVD for a 2-dimensional matrix, as follows:

$$\chi \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r = [[A, B, C]] \tag{3.14}$$

where $\circ$ indicates outer product of vectors, R is positive number, and $a_r \in \mathbb{R}^I, b_r \in \mathbb{R}^J, c_r \in$

$\mathbb{R}^K$. We set R as 30 in the first case study(chapter 4), and 15 in chapter 5. Also, the operation $[[A, B, C]]$ is called tensor rank or Canonical Polyadic(CP) decomposition where $A = (a_1, a_2, ..., a_R), B = (b_1, b_2, ..., b_R)$, and $C = (c_1, c_2, ..., c_R)$. By using the formula, the anomaly detects algorithm aims to optimize the following

$$B^* = \underset{A_i, C_i, B^* \geqq 0}{\operatorname{argmin}} \sum_{i=1}^{n} (\|\chi_i - [[A_i, B^*, C_i]]\|^2 + \alpha \|A_i - A_{i-1}\|^2 + \beta \|C_i - C_{i-1}\|^2 \tag{3.15}$$

where B^* is a standard factor matrix of all the training tensors. Therefore, it is the standard state, storing the most common relationship between features and basic traffic patterns. A_i, C_i matrices respectively capture the dynamic spatial and temporal distribution of basic patterns. The details of the optimization process is written in section 4.5 of the original paper. The paper does not explain how to define A_1 and C_1, but we set them in two ways. In our first case study, we initialize Twitter users' mobility using the left singular vectors after unfolding the input tensor. However, we use uniform distribution between 0 and 1 to initialize in the later case study, COVID-19 patients' mobility in that the uniform way delivers a better result. Finally, the α, β control smoothness on factor matrices A_i and C_i , which aims to eliminate the anomaly in the training set. We set their values 0.05.

After updating the A_i, C_i, B^*, one can detect which node shows abnormal behavior regarding standard pattern, B^*. In particular, the row vector $A_i(k, :)$ in each factor matrix A_i represents the patterns' historical co-occurrence of B^* on k-th region. We apply LOF algorithm to $A_i(k, :)$ $(i = 1, ..., n)$ to detect how much the co-occurrence has changed at ith time interval.

The standard mobility pattern comparison based on CP-decomposition has the following pros and cons:
Pros of the method:

1. Improved anomaly detection: the original paper argued this performs better than normal LOF method without the CP decomposition.
2. It provides an optimal graph state: B^* stands for the smoothed state where we reduced the anomalies by regularization.
3. Interpretablity: one can visualize the optimal state and interpret the latent pattern $A_t(k, :)$.
4. Seasonality underestimation: it hardly considers seasonal pattern as an anomaly.

Cons of the method:

1. One-stable state assumption: it assumes the dynamic graphs has only one standard state. However, if an event causes ever-lasting effect, system might have more than two stable states.

2. Relative anomaly scores: It only measures abnormality within one node. Thus, it is not reasonable to compare a region's abnormality score to another region's score.

3. Trend exaggeration: the algorithm considers the trend changing point as a strong abnormality. Even though we can partially release this problem by differencing time-series adjacency graphs, the solution is imperfect.

3.2. Methods for causal inference

We analyze causality between events outside the system and disturbance within the human mobility system using statistical methods. Granger causality is the core method. On the other hand, we have other hypothesis test methodologies which helps the causal inference.

3.2.1. Granger causality for panel data

We measure the Granger causality on multiple locations to infer the causal relation between the event and the traffic anomaly. The Granger causality is a classical and intuitive causality concept for the pair of time series variable[12]. In detail, the following model:

$$Y_t = \alpha + \sum_{k=1}^{L} \beta_k Y_{t-k} + \sum_{k=1}^{L} \gamma_k X_{t-k} + \epsilon_t \tag{3.16}$$

can be used to test whether$X$ causes Y. The basic idea is, if discarding X_t reduces the predictive power regarding Y_t, then X_t Granger-causes Y_t. To define the causality from the time series, we need the following concepts as pre-requisites :

1. Set X and Y as stationary stochastic processes
2. $U_t = (U_{t-1}, ..., U_{t-L})$ is all the information in the universe between time t and $t - L$.
3. $X_t = (X_{t-1}, ..., X_{t-L})$ is all information in X between time t and $t - L$.
4. $\sigma^2(Y_t|I_t)$ is a variance of the residual of predicting Y_t given an information $I_t \in \{U_t, U_t \setminus X_t\}$.

where L is called time lag. On top of the ideas, the definition of Granger causality is as follows

If $\sigma^2(Y_t|U_t) < \sigma^2(Y_t|U_t \setminus X_t)$ then we say that X Granger-causes Y, and write X $\Rightarrow$ Y.

One can evaluate this causality by hypothesis test based on the concept of the Vector Autoregression Model whose null hypothesis is:

$$H_0 \colon \gamma_1 = ... = \gamma_L = 0 \tag{3.17}$$

The H_0 imples X does not Granger-cause Y. Granger causality is different from the causal relation in common sense. It is solely about the ability to predict the future values of a time series using prior values of another time series. Finally, the time lag value, L, has a strong influence on the test result, so it should be carefully chosen based on domain knowledge.

Hurlin and Dumitrescu extended the concept to measure the causality from the panel data[14]. The panel data means a typical structure derived from several observations over a time interval on many cross-sectional units like countries. The underlying regression writes as follows:

$$Y_{i,t} = \alpha + \sum_{k=1}^{L} \beta_k Y_{i,t-k} + \sum_{k=1}^{L} \gamma_k X_{i,t-k} + \epsilon_{i,t} \tag{3.18}$$

where i stands for the individual subject unit. Then, we can also extend the null hypothesis as

$$H_0 \colon \gamma_{i,1} = ... = \gamma_{i,L} = 0 \quad \forall i = 1, ..., N \tag{3.19}$$

where N is the number of individual units. If reject the H_0, then we accept alternative hypothesis: there is Granger causality for at least one individual. Lopez and Weber[40] implemented as a Stata command which is also available in R's plm package.

This thesis applies the above causal methods to quantify the cause and effect of abnormal events and disturbance in the mobility network. Specifically, we figure out causality between the urban floating population, COVID-19 carriers' mobility, and the number of confirmed patients.

3.2.2. Augmented Dickey–Fuller test

This hypothesis test is to evaluate whether a given time series is stationary which is crucial assumption of Granger causality. Given an auto-regressive processes y_t, its first order differencing becomes as follows:

$$\Delta y_t = c + \beta t + \gamma y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t, \tag{3.20}$$

The null hypothesis is $\gamma \neq 0$, meaning the given time series is non-stationary.

3.2.3. Benjamini-Hochberg procedure

If Granger causality is tested multiple times, one should consider Benjamini-Hochberg procedure to correct a problem of the multiple hypothesis tests. For example, if 100 tests are conducted, given all corresponding null hypotheses are true, the expected number of incorrect rejections is still five with 5% of a significant level(α). The expected proportion of false discoveries is named as False Discovery Rate (FDR). One can correct this error by adjusting the individual α by Benjamini-Hochberg procedure[41]. The detailed algorithm is as follows:

1. Assume there are null hypotheses $H_1, \ldots, H_m$ whose p-values are $p_1, ..., p_m$.
2. Sort the p-values in an ascending order: $p_{(1)}, ..., p_{(m)}$.
3. For a given α, find the largest k such that $p_{(k)} \leq \frac{k}{m}\alpha$
4. Reject the null hypothesis for all $H_{(i)}$ for $i = 1, \ldots, k$ where $H_{(i)}$'s corresponding p-value is p_i.

By the Benjamini-Hochberg process, we can reduce the error within the multiple hypothesis tests.

4. Case Study 1: International Travel of Twitter Users

We would like to observe how social events/accidents affect the inter-country travel pattern. Hawelka et al. suggested that one could estimate international human mobility by geolocation tags from Twitter users[39]. Also, Sen and Dietz collected the geolocation tags and built their database[19]. They consider the user's trajectory as a continuous stream of check-ins: A check-in is a tuple of the unique identifier of a user (UUID), a location, and a timestamp. They adopted a heuristic assumption that a user's home country is the one with most check-ins. On top of their study, Widmer[42] proved that the intercountry travel from the Sen and Dietz's data represents the actual travels by referring to official European travel statistics. Utilizing the qualified dataset, we apply the anomaly analytics introduced in the former chapter and characterize the mobility.

Note that this chapter assumes one directional causal relation: exogenous events affect a travel pattern in a dynamic network. Because the global mobility is a complex system, it is hard to interpret how a specific travel pattern causes an abnormal event. We describe the opposite directional causality in chapter 5, where an mobility anomaly causes a social event.

4.1. Trip mining

To represent international travels as spatio-temporal graphs, we also utilize the data of Sen and Dietz. We extracted global trips from 10.Jan.2016 to 10.Nov.2019, in a total of 200 weeks. Then, we created a series of adjacency matrixes, where each matrix has weekly travel information. The individual matrix has 226 columns and rows whose index represents a country in the dataset. Also, each matrix is weighted and directed form. For example, if $A_{1(i,j)} = 3$, then it means there were three mobility records from country i to j at the first week. Finally, we formulate them into a tensor, $\chi \in \mathbb{R}^{200\times226\times226}$.

We numbered countries according to the amount of incoming traffic. The USA has the most visitors; thus, its index is 0, followed by United Kingdom(1), and France(2), Spain(3) and Italy(4). The table in the appendixA.1 provides the list of the top 30 countries with abbreviation code.

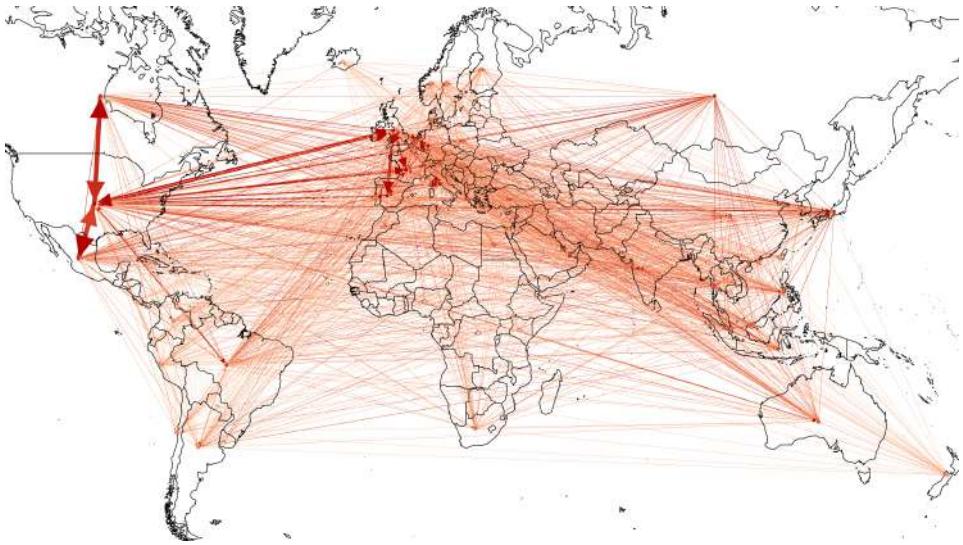

Figure 4.1.: The geological distribution of Twitter users' trips

4.2. Exploratory data analysis

Before analyzing anomaly from the data, we need to understand data through Exploratory Data Analysis (EDA). The data has four properties: (1) skewed traffic distribution, (2) trend and seasonality, (3) complexity (4) noise reduction by time aggregation

4.2.1. Skewness of traffic distribution

First of all, the distribution of travel records are skewed, and the US is a hub node. The figure 4.1 describes the average geological distribution of the travels. The more traffic volumes a country has, the darker and thicker the directed edges. As mentioned, US has the massive traffic volume. Specifically, the first plot of figure 4.2 illustrates the total incoming and outcoming traffic over the entire time interval by country, and it shows the distribution skewed to one side. The skewness is due to the demographic distribution of Twitter users. According to Widmer's study, specific demographics such as white males [43] are over-represented on Twitter. Further, demographic differences have been observed between those Twitter users who geotag their tweets and those that do not[44]. Moreover, the outgoing traffic also has a comparable distribution to the incoming one, even though we originally designed directed graphs.

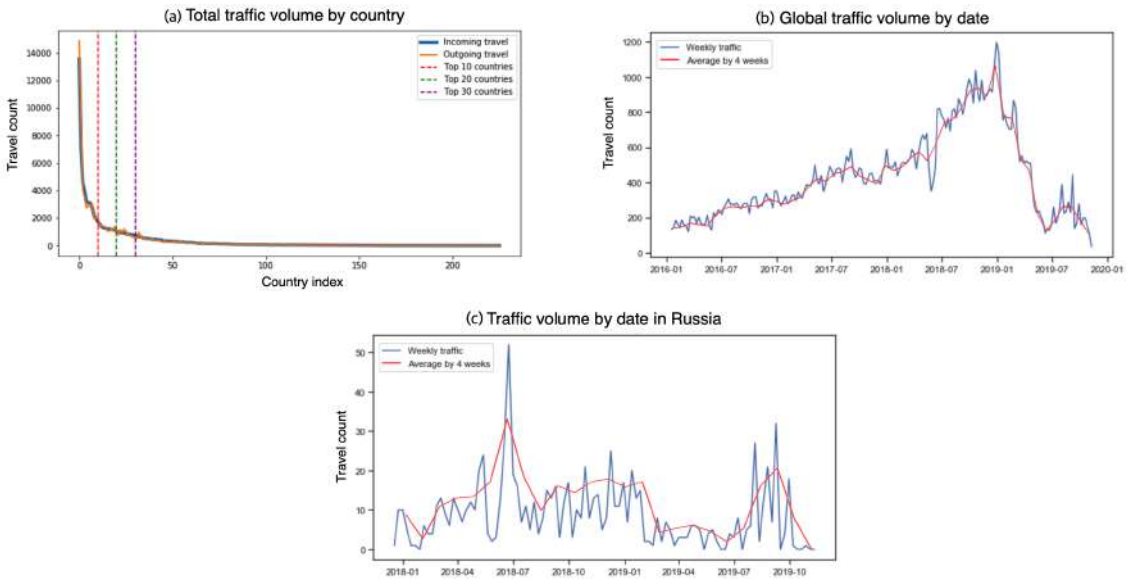

Figure 4.2.: Travel pattern from Twitter geotags.

4.2.2. Trend and seasonality

Next, the traffic has a trend and seasonality; thus, it is not stationary. The second plot of figure 4.2 shows the trend. It is due to the amount of data collection, not because of the actual trend of international trips. Moreover, the third plot explains the seasonality. One can observe the high peaks during specific seasons depending on a country. One of our challenge is distinguishing major event from the seasonality and trend

4.2.3. Complex system

The global travel data represents a complex system with multiple interactions between many different components. In such a system, it is never straightforward to infer causality of every volatility. For example, the traffic amount varies at any time point in the second plot of figure 3.1, but it is impossible to figure out particular reasons for the variance. If readers are interested in causal inference between disturbance in the system, and exogenous events, refer to the case study of COVID-19 patients.

Mobility scale	The number of nodes	Time length	Density1	Density2	Seasonality	Trend
International travel	226 countries	200 weeks	0.1137	0.6233	Yes	Yes

Table 4.1.: Properties of mobility network from Twitter

4.2.4. Noise reduction by time aggregation

Despite the complexity, we would like to characterize how a major exogenous event affects the human mobility system, thus consolidated the weekly graph into 4-week to reduce noise. The second and third plots of the figure 4.2 illustrate the smoothed the traffic volume. Furthermore, it reduces the sparsity to use the most frequent nodes and the larger time granularity. The table1 shows two densities. The density1 is from the weekly dynamic graph, while the density2 comes from average graphs by 4-week units. Due to the skewed traffic distribution, we used only the top 30 nodes concerning incoming traffic to compute them. The specific formula for graph density is as follows:

$$D = \frac{\sum_{t,u,v} |E_{t(u,v)}|}{|T||V|(|V|-1)} \tag{4.1}$$

where $E_{t(u,v)}$ is 1 if there is a directed edge from u to v at time t, otherwise 0. $|V|$ is the number of nodes that we set as 30. Finally, $|T|$ is time length whose value is 200 in the weekly case and 50 in the other one.

4.3. Event characterization

In this section, we analyze how an abnormal exogenous event affects the spatio-temporal mobility network of Twitter users. We aim to deliver a valuable illustration of anomalies beyond the simple data exploration by the five methods in the previous section. We explain the details of each topic.

4.3.1. Events information

This case study mainly focuses on three major events: the demonstration in Hong Kong in 2019, the huge earthquake of Indonesia in 2018 and the 2018 World Cup in Russia. They respectively stand for a political protest, natural disaster, and sport event in a global scale.

1. Demonstration in Hong Kong(2019): it was a massive demonstration where more than one million Hong Kong civilians had gathered since June of 2019. The protest was

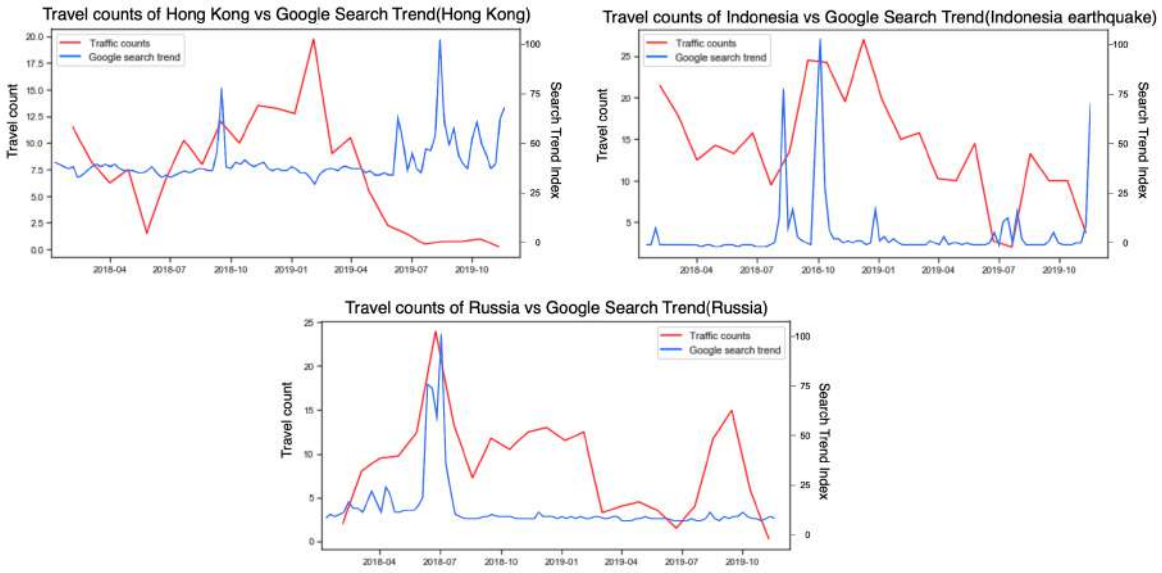

Figure 4.3.: Travel pattern and Google search trend of Hong Kong

against plans to allow extradition to mainland China. This event decreased the travel counts from/to Hong Kong. The first plot in the figure 4.3 describes the total traffic counts in Hong Kong, a sum of outgoing and incoming travel records. One challenge is that the decrease in April is not because of the demonstration, but due to the dataset trend. Google's global search trend supports this. The trend index is a scaled metric on a range of 0 to 100 to represent relative popularity given a specific time period and location. According to the index, global people had become interested in the issue in Hong Kong since June of 2019. The protests lasted longer than the Twitter data was compiled. This chapter figures out the impact of this exogenous event.

2. Indonesia's earthquake (2018): it was a 7.5-moment earthquake in Sulawesi, Indonesia, on September 28, 2018. According to the Straits Times in Singapore, more than 1,200 people had died. We assumed this accident would reduce the traffic volume. However, the decrease is not clearly visible through the EDA(figure 4.3). It did decrease around November, but it is not confirming how the event disturbed the mobility network. In the next section, we illustrate that our method can characterize it more clearly.

3. The World Cup in Russia(2018): the 2018 FIFA World Cup was an international football tournament contested by national teams between 14 June and 15 July 2018 in Russia. We can recognize the peak of traffic volume in July. 2018(Figure 4.2). We aim to characterize the impact of this event on the network beyond the normal EDA.

4.3.2. Social-geological metrics

We applied the four measures to analyze the impacts of the three types of exogenous events: a political demonstration, sport event, and natural disaster. We considered the top 30 countries to draw Hong Kong and Indonesia cases while used 40 countries for Russia.

1. The change of entropy and homogeneity in Hong Kong is in the first plot of figure 4.4. One can easily recognize that entropy has dropdown after the middle of June. The drop means less diversity of travelers during the period. In other words, the protest blocked visitors to Hong Kong from various origin countries. Plus, there were not enough travel records to measure homogeneity during the demonstration. The denominator in the homogeneity formula becomes zero due to the little visit. As a result, there is no homogeneity since June. Recall that the demonstration lasted longer than the data collection period. Therefore we highlighted the anomaly period as a light red color.

2. The second plot shows Indonesia's cases, where entropy and the brokerage dropped at the beginning of October. Presumably, the earthquake decreased the diversity of travelers' home countries. Also, it decreased the centrality of Indonesia. Recall that the simple travel counts do not show this correctly. Thus, it is a representative example that the graph metrics better reveal an anomaly by an exogenous event.

3. The World Cup event drastically raised brokerage and popularity by attracting people from worldwide. Noticeably, the popularity had increased eight weeks before the start but drastically dropped during the actual event period. Thus we can conclude that people tend to visit the World Cup host country even two months before, but not the actual festival period. On the other hand, the brokerage peaked accurately during the event. Again, we highlighted the World Cup period with the red light.

4.3.3. Dynamic graph prediction

By the graph prediction, we detect the mobility network's anomaly and describe how the graph structure change with visualization. Recall that the EvolveGCN supports only three types of forecast: link prediction, edge, and node classification. We will make two link prediction models; one is for 2018 9, and the other is for only 2019, then compare the historical links to the prediction. Regarding the first model, the prediction period becomes from 6.Jan.2018 to 19.Nov.2019. We need to split the spatio-temporal graphs of 4 weeks unit into train, validation, and test dataset. Specifically, $\chi_{train} \in \mathbb{R}^{24\times N\times N}$, $\chi_{validation} \in \mathbb{R}^{5\times N\times N}$, $\chi_{test} \in \mathbb{R}^{20\times N\times N}$, where N is the number of countries. In the application, we have checked the N is a crucial hyper-parameter for the forecasting performance. As already explained in

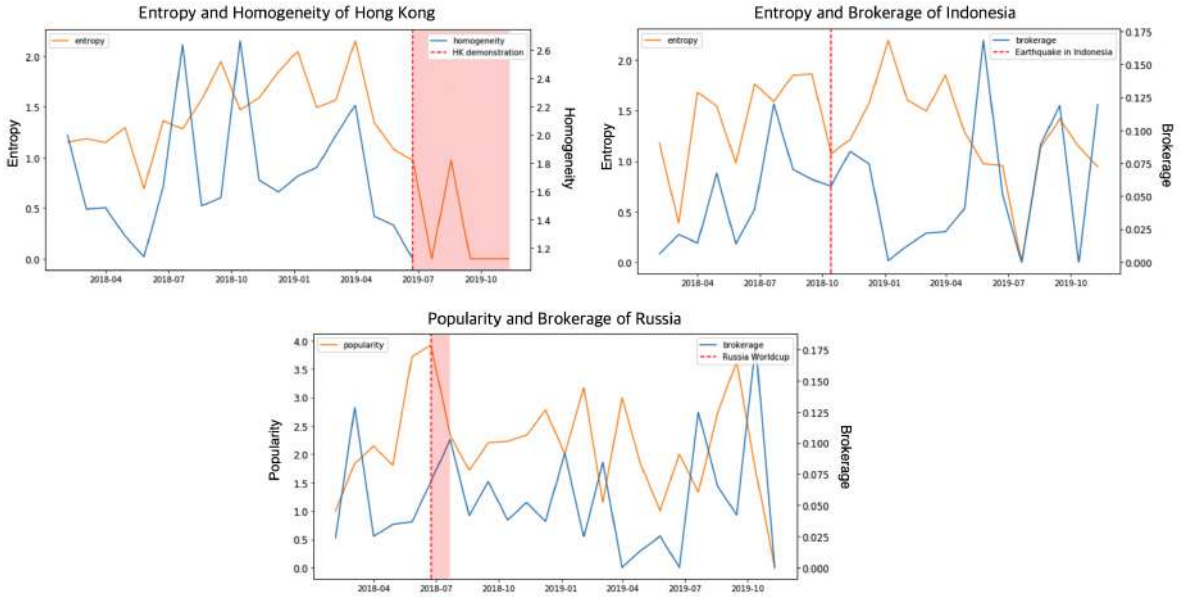

Figure 4.4.: Social-geological metrics from the Twitter user mobility. The highlighted red areas indicate the duration of the events.

Model	Best F1 for the link presence	Epoch	MAP	Absence of link			Presence of link		
				Precision	Recall	F1	Precision	Recall	F1
GCN	0.6067	46	0.6610	0.7586	0.5214	0.6180	0.5095	0.7497	0.6067
EGCN-O	0.6175	59	0.6656	0.7764	0.5030	0.6105	0.5104	0.7815	0.6175
EGCN-H	**0.6188**	50	0.6658	0.7590	0.6286	0.6877	0.5551	0.6990	**0.6188**

Table 4.2.: Evaluation of the graph prediction for 2018/9.

the EDA section, tighter N reduces the less informative nodes that play a noise. In this first modeling, we set N as 25, because both our target countries, Indonesia and Russia, belong to the top 25 countries.

The table 4.2 shows the validation result. EGCN-H shows the highest F1 score of 0.6188, but the difference is subtle. Remarkably, the normal GCN shows the lowest grade, 0.6067, which is only lower by 0.0121.

After making the model, we compare the prediction result to actual values by a scaled graph density, $density'(G_t)$ and $degree'(v)_t$ whose formulas are the equations 3.9, 3.10 respectively. The figure 4.2 shows the forecasting result. The plot (a) is the global density of prediction and historical record. It seems the prediction generally matches the ground truth. Particularly, The red line refers to the trend change point, and the local peak in July and August of 2019 stands for the seasonality. Therefore, the model can capture the not stationary properties. The

Model	Best F1 for the link presence	Epoch	MAP	Absence of link			Presence of link		
				Precision	Recall	F1	Precision	Recall	F1
GCN	0.6844	148	0.7460	0.6984	0.5520	0.6166	0.6242	0.7574	0.6844
EGCN-O	0.6882	87	0.7465	0.7277	0.3837	0.5025	0.5764	0.8538	0.6882
EGCN-H	**0.6965**	131	0.7468	0.7569	0.3608	0.4887	0.5755	0.8821	**0.6965**

Table 4.3.: Evaluation of the graph prediction for 2019.

plot (b) and (c) illustrate the link density of egonets for Russia and Indonesia. The Russian case shows more confirming anomaly. In other words, during the World Cup season(June and July 2018), the historical traffic showed an apparent anomaly.

On the other hand, we do not have such an obvious observation concerning the earthquake. Specifically, the red dotted line in the third plot indicates the time interval from 16.Sep.2018 to 14.Oct.2018 when the prediction and historical records do not differ.

The graph visualization (d) and (e) shows the temporal graph state during the time of the red-dotted line(24.June.2018 21.July.2018). The 25 countries are distributed according to the Force-directed graph drawing method[29]. It assigns forces among the set of edges and the set of nodes, based on their relative positions. Roughly speaking, the layout locates major nodes to the center. As a result, we can observe that the actual location of Russia(RU) is closer to the United States(US). Recall that the US is a hub node in this mobility network. Regarding centrality, the predicted between centrality of RU is 0.0045, while the real one is 0.0195. Thus, the special event raised the centrality of Russia around four times.

Next, we build the other model to forecast the mobility network in 2019. Because the prediction period is different, the size of the train, validation, and test sets also vary; $\chi_{train} \in \mathbb{R}^{33\times N\times N}$, $\chi_{validation} \in \mathbb{R}^{5\times N\times N}$, $\chi_{test} \in \mathbb{R}^{10\times N\times N}$. table 4.3 presents the model evaluation. In this training, we used the top 30 nodes to include Hong Kong which is the 28th node, so $N = 30$. Again, EGCN-H shows the best F1 score with the difference of 0.121. Therefore, the following prediction is based on the EGCN-H.

The figure 4.6 describes the forecast. Again, the global prediction regarding *scaled* d_t captures the historical pattern. Notably, there is little difference between the two values when the demonstration happened. However, since September, the actual record has stayed lower. It explains the disturbance by the exogenous event lasted until November, the end of the time interval.

Also, the plot (c) and (d) illustrate the anomaly in more detail. According to the prediction, Hong Kong should have interacted with other countries, but historically there was only one traveler from Hong Kong to Japan without any incoming traffic.

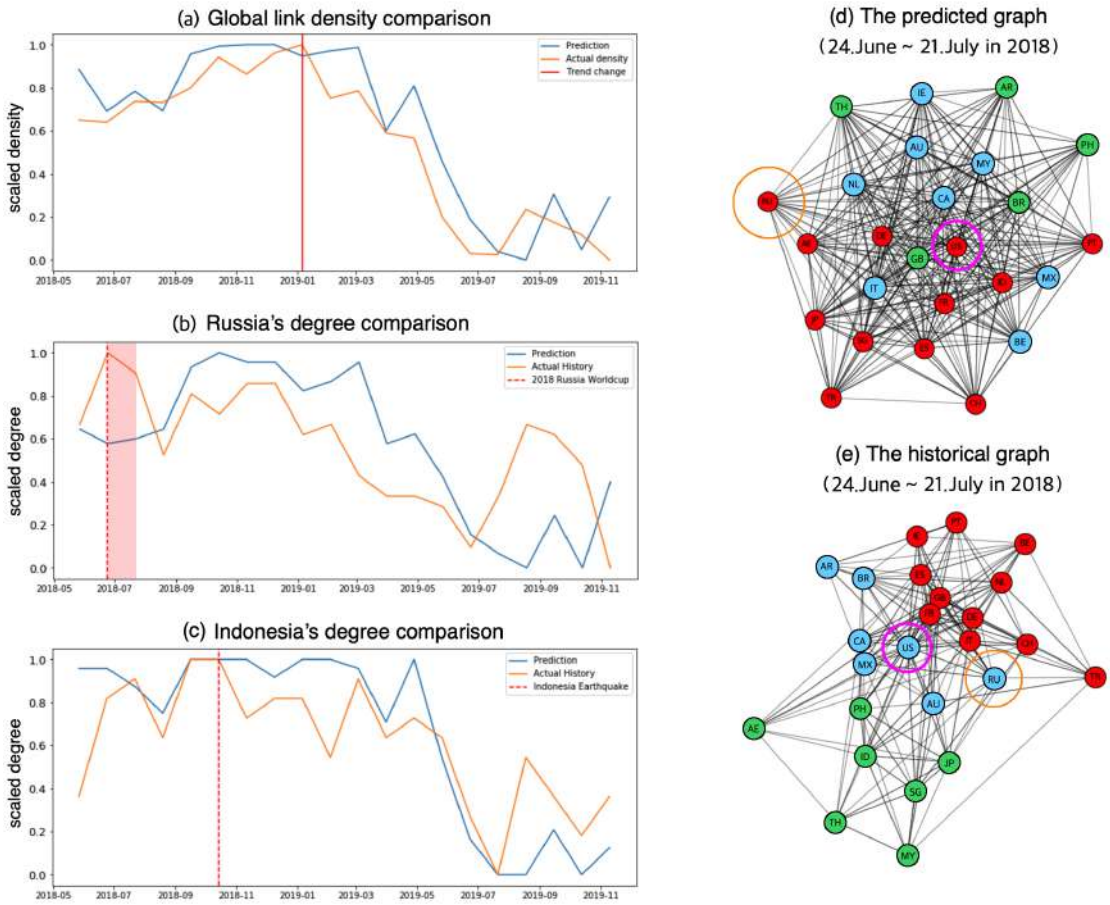

Figure 4.5.: Graph prediction for the 2018. The highlighted red area in the plot (b) indicates the duration of the World Cup. (d) and (e) refer to the predicted and historical graphs at the time of Russia World Cup. 'US' and 'RU' respectively stands for the United States and Russia. In addition, the node colors mean communities according to the Leiden clustering[30]. The red cluster acquires the most countries, followed by the light blue and green.

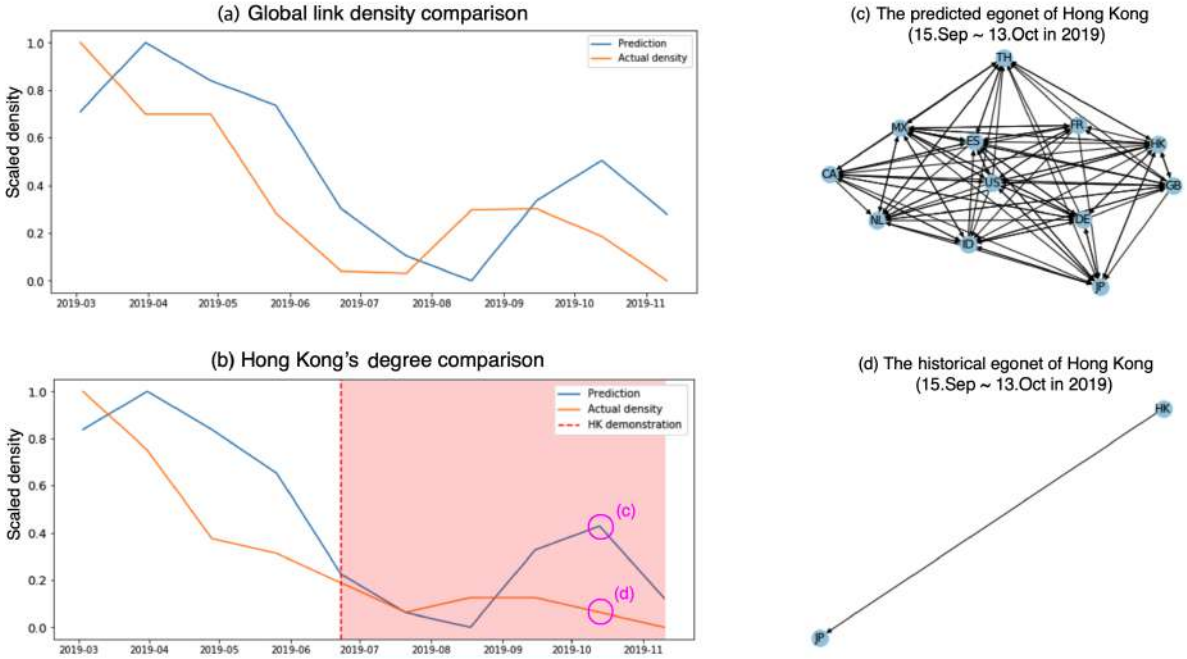

Figure 4.6.: Graph prediction for the 2019. The highlighted red area in plot (b) indicates the duration of the demonstration. (c) and (d) refer to the predicted and historical egonet of Hong Kong.

4.3.4. Neighborhood feature analysis

We characterize a representative feature of an event from the neighbors by the method introduced in the subsection3.1.4. The algorithm measures the abnormality compared to the past W weeks. We need to use a node's dynamic neighbors as an input. For example, $F_f(G_{HK}) = v_T$ where G_{HK} is the Hong Kong's dynamic neighbors. A different choice of graph feature f leads different v_T. We tested and observed v_T, given all twelve features previously mentioned in the subsection. As a result, we comprehended that some features stand for an anomaly by an exogenous event, while the others do not.

The plot (b) of the figure shows that the anomaly score surged during the protest period when we select the feature as in-degree. However, it has not showon any anomaly since July with the in-weight as the feature. In other words, the protest mainly affects the number of connections to other countries, rather than weight on the edges. We made the analysis, given the hyper-parameter W is 9. In the 2018 World Cup case, out-weight represents the anomaly in Russia's neighborhood network, while the number of connected nodes(neighbors) are not. Thus, we can conclude that the World Cup cause mainly change the traffic volume between Russia and other foreign countries, rather than Russia's degree. We did the experiment, given the W is 8 for the case of out-weight feature, and 9 for the number of neighbors.

4.3.5. Bilateral traffic model

With the generalized gravity model 3.13, we do modeling and characterize traffic between two spatial points. Surprisingly, this simple model accurately captures the complex global mobility with proper data transformation. To build the model, we need to prepare the four variables: an emission(E_i), an attraction of a country(A_j), a distance between countries d_{ij}, and a traffic volume from one country to another T_{ij}. We can extract their values from the mobility graph except for d_{ij}. Specifically, $T^t_{(i,j)} = \chi_{(t,i,j)}$, $\sum_i T^t_{(i,j)} = A^t_j$, $\sum_j T^t_{(i,j)} = E^t_i$. Recall that the model does not consider time. Thus we need to aggregate all variables into average. For example, if $i = 0$ and $j = 1$, then $T_{ij} = E_t(T^t_{ij})$, which indicates the average traffic from node 0(USA) to node 1(UK). We observed that this aggregation increase the model accuracy by reducing noise. Additionally, we measure d_{ij} as the one between their largest cities by using the geological database of Mayer and Zignago[45].

After preparing the initial dataset for the four variables, we need to check their distribution for the generalized linear modeling(GLM). The figure 4.8 shows that the variance of $log(T_{ij})$ increase as emission and attraction rise. It violates the underlying assumption of the simple regression. Moreover, the variance is asymmetric. To model such volatility, we need to set another distribution of the target variable, $log(T)$, and another relation between the target and explanatory variables if necessary. Therefore, we tested Gamma, Inverse Gaussian, and

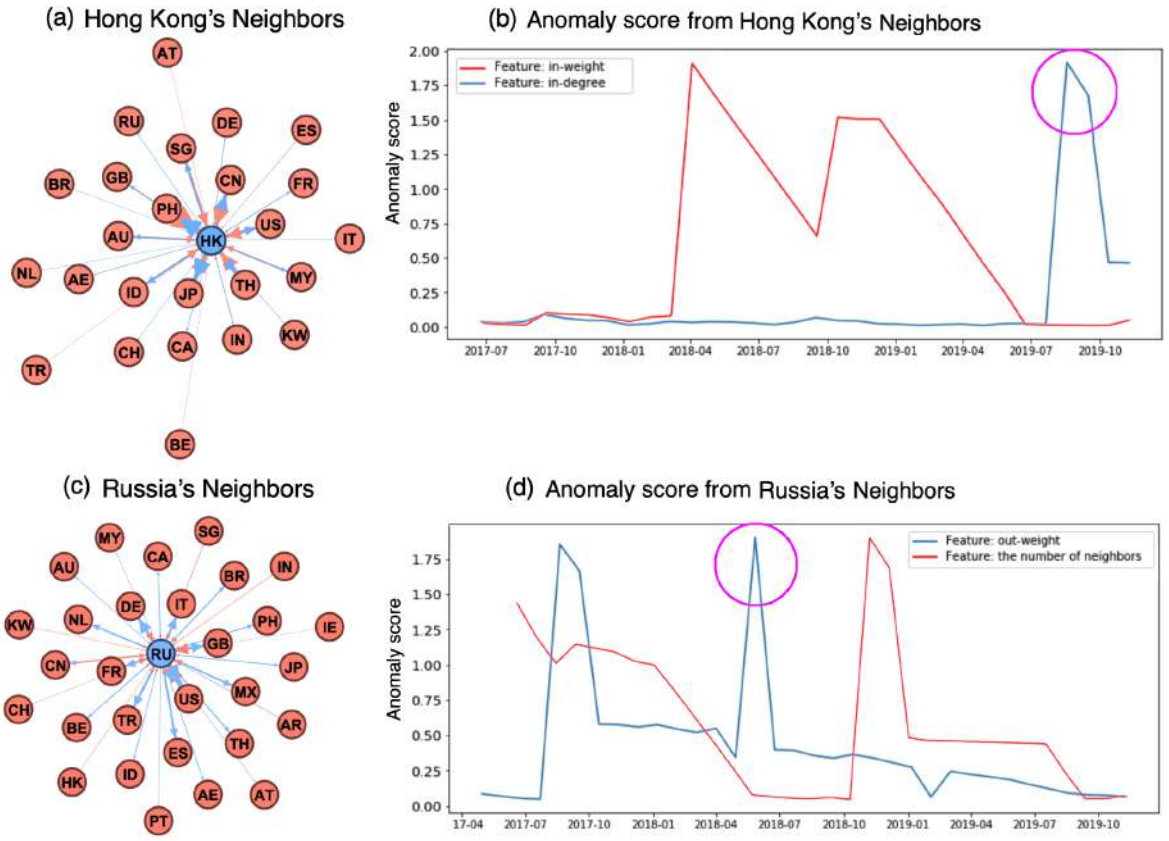

Figure 4.7.: (a) the Hong Kong's neighborhood network whose edge weight is the average over the whole time. The blue arrow means out-going traffic from Hong Kong and the red traffic is in-coming one. Thickness stands for the traffic volume (b) anomaly scores from the neighborhood network's in-degree. (c) the Russia's average neighborhood network (d) anomaly scores from the neighborhood network's out-weight

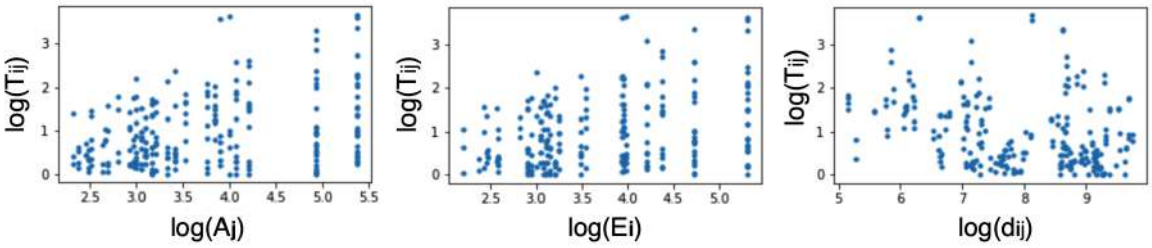

Figure 4.8.: EDA for gravity modeling

Distribution	**Gaussian**		**Gamma**			**Inverse Gaussian**		
Link function	μ	$log\mu$	μ	$log\mu$	μ^{-1}	μ	$log\mu$	μ^{-2}
Pearson χ^2	28.45	25.05	7.27	6.74	6.16	3.87	3.64	**3.19**
Deviance	28.45	25.05	7.82	7.13	6.46	4.26	3.93	**3.41**

Table 4.4.: Evaluation table for the generalized gravity model

normal Gaussian one. Moreover, we need to add e on the original data to transform the data with logarithm and inverse as link functions. The table 4 presents the result of the GLM. The model with Inverse Gaussian distribution and inverse squared link performed the best. The table 4.4 presents the result of the GLM. The model with Inverse Gaussian distribution and inverse squared link performed the best. As a result, the generalized gravity model is

$$(log(T'_{ij}))^{-2} = 0.8628 - 0.0897 \cdot log(E'_i) - 0.0876 \cdot log(A'_j) + 0.0506 \cdot log(d'_{ij}) \tag{4.2}$$

where $X^{'} = X + e$, and all the coefficients are estimated with p-values less than 10^{-20}. The figure 4.9 proposes the comparison between the actual traffic and estimation by the model. The model expresses almost every traffic volume.

Because the model describes bilateral traffic, we characterize anomaly regarding two

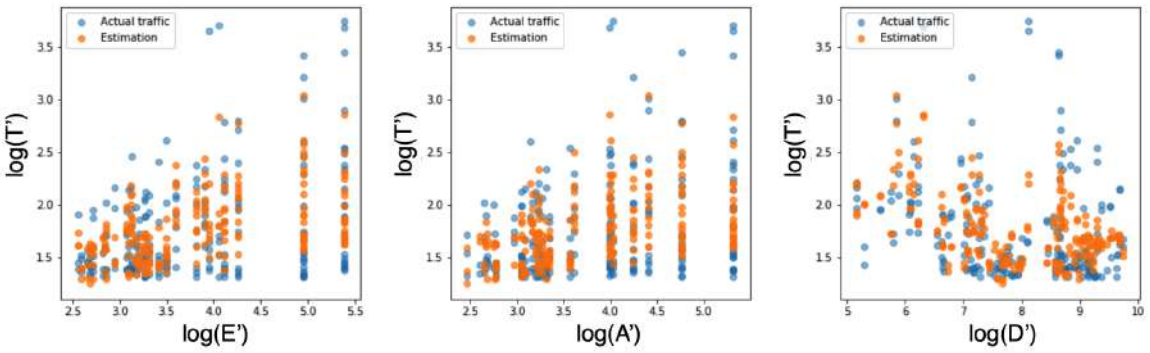

Figure 4.9.: The actual traffic vs estimation by the generalized gravity model

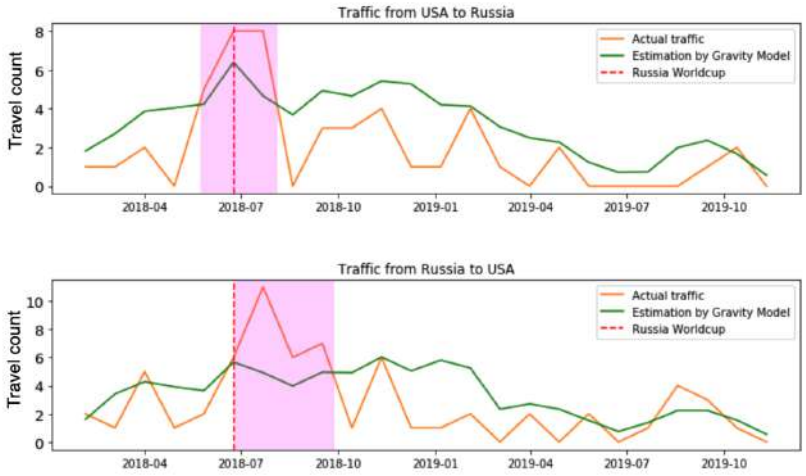

Figure 4.10.: The anomaly caused by the World Cup. The highlighted pink area indicates anomaly period regarding the bilateral traffic.

directed edges between nodes. For example, figure 4.10 proposes two traffic flows between the USA and Russia. While the USA's actual traffic to Russia had exceeded since June (the pink box in the upper plot), the opposite direction's traffic started to exceed near July(the pink box in the lower plot). The first exceed finishes around six weeks after the event, but the other one lasted until near October. Presumably, the pattern is because the travelers arrived in Russia before the sporting event and leave after it finished.

On the other hand, we do not characterize the earthquake and protest. It is because the three variables, E_i, A_j,d_{ij}, already include sufficient information to explain the traffic change by the two events, so the actual traffic is not significantly different from the estimation.

4.3.6. Standard mobility graph comparison

This subsection describes how to recognize an anomaly pattern by comparing a travel pattern to the ideal mobility state. To compute the ideal one, we use the tensor rank decomposition introduced in the previous section. First, we formulate the tensor and decompose it into A_i, C_i, and B^*, where i is a variable meaning time step. Because we use the dynamic graph of 4 weeks of the time unit, the time index i is between 1 to 50. The B^* means the regularized mobility pattern where less abnormality exists. The figure 4.11 shows this pattern. The more traffic a country emits, the more the outgoing edge's color becomes magenta. As we have already explored, the United States has the most significant outgoing traffic volume. The volume generally decreases as it goes to the east. After acquiring A_i, C_i, and B^*, we apply the LOF algorithm to measure the anomaly scores from $A_i(k:)$. Because k is a constant that indicates a specific node, our interest lies in finding the temporal point i where the anomaly

Figure 4.11.: The regularized mobility network(B^*)

score is notable, given a specific country.

We focus on the three nodes: Hong Kong, Indonesia, and Russia. The three plots in the figure 4.12 illustrate our analysis. The blue line means abnormality score by the method based on tensor rank decomposition, and the orange line is the total traffic volume, which is a summation of incoming and outgoing traffics. The purple dotted line indicates the trend change point when the data collection started to decrease. Finally, the red dotted line is the time when the events occurred.

The first plot shows Hong Kong's case, and it does not show the demonstration's time as an abnormal. Instead, it recognizes the trend change point in Jan.2019 as the most significant anomaly. We have observed this problem from 23 of the primary 30 nodes.

To reduce the trend, we did differencing the time series of adjacency matrixes. As a result, the algorithm reported the trend changing as the strongest abnormality in 11 of the 30 nodes; thus, it worked partially. However, the algorithm still did not detect the anomaly at the event time. Instead, it pointed out the traffic decrease in April.2019. One presumable reason for this malfunction is that Hong Kong's traffic pattern violates the model's assumption: there is only one standard mobility state. However, the protest decreased the traffic volume for more than five months. Even though the interval is a lasting abnormality, the model considers it as the usual situation.

On the other hand, the method successfully identifies the mobility change by the Worldcup. The result is because the trend in Russia is not significant. In other words, the orange line in the plot (a) shows the evident increase before the purple line and decrease after the time point, while the traffic volume in the plot (c) does not. Plus, it is noticeable that the method

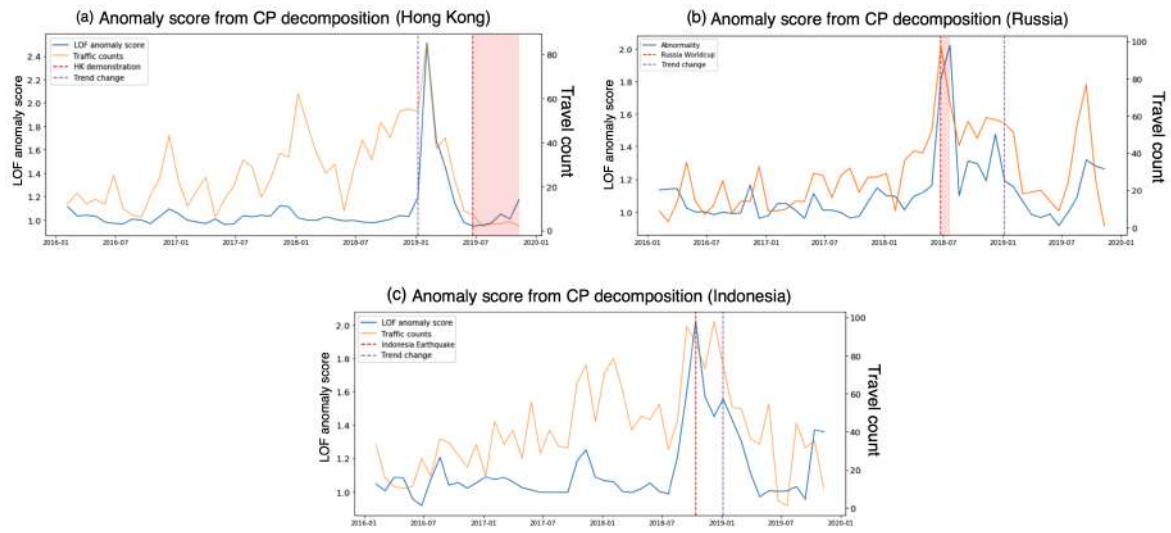

Figure 4.12.: Anomaly scores by the tensor rank decomposition. The highlighted red areas indicate the duration of the events.

assigns more significant abnormality to the World Cup than the seasonal traffic increase in 2019. It is a case that shows the method's robustness to seasonality.

The analysis becomes complicated concerning the Indonesia case. It seems that the algorithm does not solely recognize the earthquake as a severe anomaly. Instead, because the event caused the local minimum of traffic volume, the trend seems to change at the earthquake, not the purple line. Presumably, that is why the algorithm assigns the highest anomaly score to the event time.

To summarize, we can (1) characterize the ideal state of a mobility network where no anomaly exists and (2) detects an abnormal traffic pattern compared to the ideal one, and the method is (3) robust to a mobility seasonality, not to a trend.

4.4. Summary

We have detected and analyzed anomaly mobility pattern caused by the abnormal events. First, we extract travel patterns from Twitter users' geotags and explored the data. The trip trajectories have skewed distribution whose hub node is the USA. Also, it involves seasonality and trend. At last, the data includes many noises as the global human mobility is a complex system. To reduce the noises, we decided to use a four-week dynamic graph. In particular, one graph represents the movement patterns of Twitter users around the world for four weeks. By stacking 50 of these graphs, we built our spatiotemporal graphs.

Next, we applied our five methods to characterize the impact on the mobility network by

Characterization Methods	Anomaly type	Demonstration	Earthquake	World Cup	Sesoanlity	Trend	Comment
Naive EDA	Type A	No	No	Yes	Yes	Yes	The most naive traffic volume exploration
social-geological metrics	Type A	Yes	Yes	Yes	Yes	Yes	Measure social diversity in a node
Dynamic graph prediction	Type B	Yes	No	Yes	Yes	Yes	Predicting whole graph structure
Neighborhood feature analysis	Type B	Yes	Yes	No	Yes	Yes	Recognizing an anomaly's representative features
Generalized graivty model	Type C	No	No	Yes	Yes	Yes	Bilateral mobility modeling with the three variable
Graph tensor decomposition	Type C	No	Yes	Yes	Yes	No	The standard mobility graph by regularization

Table 4.5.: Summary of the method application

global events. Specifically, we chose three international issues: the protest in Hong Kong, the earthquake in Indonesia, and the World Cup in Russia. We applied our five methods, and the table suggests our analysis result. The first 4.5 column describes the methodology names. The second column means the anomaly type that the techniques can recognize. Specifically, type A, B, C respectively stand for our definitions of an anomaly in a spatio-temporal graph: (1) outlier concerning a graph metric(type A), (2) an observation differing from the past pattern(type B), (3) an observation differing from standard mobility pattern(type C). The third to the fifth columns say if the method could detect and portray the individual event. The seasonality and trend column means if each method can underestimate the time-series properties. For example, the dynamic graph prediction successfully identifies the anomaly by the demonstration and World Cup, but not the earthquake. Plus, the algorithm is robust to the trend and seasonality in the data. Thus it is reasonable to apply the method to such not stationary dynamic graph data with trend and seasonality. Finally, the comment column briefly explains the properties of the methods.

The table 4.6 summarizes our analysis from the perspective of the events. They have diverse spatio-temporal information as well as exogenous causes. The impact period means the duration when the event affects the mobility network. If a method failed to characterize specific details, we omitted the cell($-$).

In this Twitter users' mobility analysis, the causal relation is one-directed. In other words, an exogenous event unilaterally affects the mobility pattern, not inverse direction. On the other hand, the next chapter, the case study for the COVID patients' mobility in South Korea, explains their bidirectional interaction, where an exogenous factor affects the mobility network, and the abnormal mobility pattern influences the outside of the system.

	Demonstration	**Earthquake**	**World Cup**
Exogenous event type	Political protest	Natural disaster	Sport event
Location	Hong Kong	Indonesia	Russia
Date	9.June.2019	28.Sep.2018	15.June.2018
Impact period on mobility	Jun−Nov.2019	Oct.2018	Jun−Oct.2018
Resulting anomaly type	Type A, B	Type A, C	Type A, B, C
Traffic volume	Decrease	Decrease	Increase
Social diversity	Decrease	Decrease	–
Between centrality	–	Decrease	Increase
Difference to the prediction	Fewer neighbors	–	Closer to the hub node(USA)
Traffic between a hub node	–	–	Bilateral overflows
Representative feature of the local anomaly	Connections to neighbors	–	Traffic volume to neighbors

Table 4.6.: Summary of the events

5. Case Study 2: Mobility of COVID-19 Carriers in South Korea

This chapter analyzes anomalies from the COVID-19 virus careers' dynamics that leads to the epidemic outbreak. We use the patients' dataset from South Korea and create a spatio-temporal mobility graph. Then, we explore the graph. We quantify the causal relations between carriers' traffic volume and the newly confirmed cases as well as the normal floating population. We also analyze anomalies to interpret the virus carriers' mobility within the network. Plus, we figure out pre- and post-signals when the infected population surges in an urban area.

The following three points distinguish this chapter from the former: (1) event causality, (2) network property, (3) anomaly types. Regarding the causality, the former mainly explains how exogenous events affect the mobility network. However, this chapter describes how virus carriers' traffic surge causes the outbreak and quantifies the relationship.

Next, they have different network structures. We briefly introduced it in section 1.4, and suggest further details in the section 5.2. Because of the different property, anomaly analysis methods operate differently.

Finally, this chapter does not specify anomaly types. It is because we already handled the issue enough in the previous chapter, and our methods internally distinguish them. In other words, an outlier by EDA and Social-geological metrics is a type A anomaly. Neighborhood feature analysis and dynamic graph prediction detect a type B anomaly. Moreover, the bilateral traffic model and standard mobility comparison methods do type C.

5.1. Virus carriers' mobility mining

We utilize South Korea's epidemic dataset and formulating patients' trajectory as a spatio-temporal mobility graph. Here, the patients refer to the infected people as a synonym to carriers. Global researchers and engineers cooperated with the Korea Disease Control and Prevention Agency(KDCA) and published the data source online: https://github.com/ThisIsIsaac/Data-Science-for-COVID-19. Among the dataset, PatientRoute.csv includes location data by date and individual patient ID, which is no more available because of the privacy issue. KDCA tracked the carriers' roots after they got confirmed. For example, if a carrier 'A' got confirmed

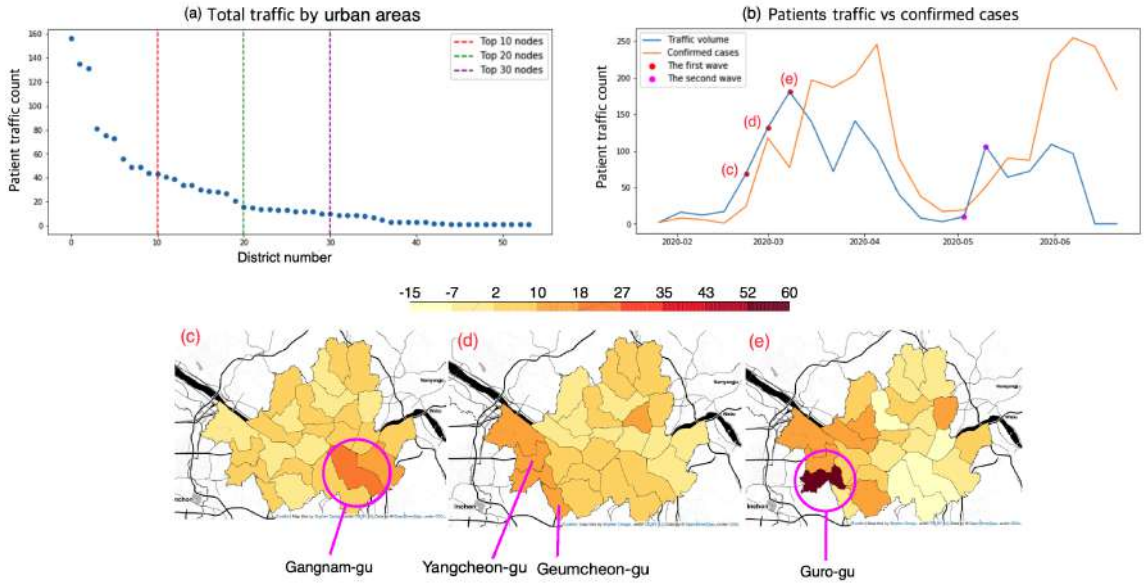

Figure 5.1.: COVID-19 patients mobility pattern

on 6. June, he is no more allowed to the outside, and KDCA traced his past routes. We transform the raw data into a tensor $\chi \in \mathbb{R}^{22\times54\times54}$ that consist of 22-weeks graphs with 54 nodes. The 22 weeks from 19.Jan.2020 to 21.June.2020. The former date is when the first COVID-19 carrier arrived in South Korea from Wuhan, China, and she got confirmed on 20.Jan. The 54 nodes are urban districts in and surrounding Seoul. Even though the original data includes other cities information, we only concentrate on Seoul and its suburbs to reduce noisy pattern.

Moreover, the dataset involves the floating population of Seoul based on one telecommunication company(SK Telecom)'s signal information. The original data reflects only their customers' floating information. Thus, we scaled them by multiplying 100/41.8, where 41.8 is the SK telecom's market share as of July 2020. One can refer to the data at https://www.bigdatahub.co.kr/

5.2. Exploratory data analysis

Before investigating anomaly, we should understand data through Exploratory Data Analysis (EDA). The data has four properties: (1) skewed traffic distribution, (2) waves, (3) Limited complexity (4) Sparsity (5) Causality. However, it does not involve a trend.

5.2.1. Skewness of traffic distribution

The plot (a) in the figure 5.1 illustrates the skewness of traffic distribution. Among 54 nodes, the prime 30 nodes occupy 93.8 percent, and the top 20 ones do 84.4 percent. The reason for the unbalance is each district has its role. Some regions are crowded downtown, while the others are residential areas. From now, we keep using the 30 locations to include all districts in Seoul. We attach an appendix to map the index numbers to geological names.

5.2.2. Waves

There were two epidemic waves. The first wave started at the end of February, while the second does in May. To observe them, we consider two metrics: (1) the patient traffic volume and (2) the number of confirmation. The plot (b) indicates that the last metric follows the first(Figure 5.1).

5.2.3. Limited complexity

Because the mobility system only includes urban COVID-19 carriers, the system is less complicated than the former case study. It means that we can understand causal inference between floating population, COVID-19 carriers' mobility, and the number of confirmation. Due to the simplicity, the blue line in the plot (b) has smaller fluctuation than the Twitter users' traffic in the former chapter. We can consider the low volatility as little noise.

5.2.4. Sparsity

The dynamic graph only describes the virus carriers, but South Korea has little patients records because of their prevention. Therefore the traffic volume is low, and the final tensor(the stack of the dynamic graphs) becomes less dense. The table 5.1 explains this. We calculated the density1 with top 30 nodes, and the density2 with top 10 nodes. Recall the Twitter case where the density2 was 0.6233(table 4.1). Because of the sparsity, we failed to apply graph prediction. We would explain the detail later.

5.2.5. Causality

The plot (b) in the figure 5.1 indicates that the number of confirmation and carriers' traffic volume have a positive relation. Their correlation is 0.6616, with one week of time-shifting. We further analyze the causality in this section 5.3.2

Mobility scale	The number of nodes	Time length	Density1	Density2	Seasonality	Trend
Urban mobility	54 countries	22 weeks	0.0400	0.1009	Yes	No

Table 5.1.: Properties of COVID-19 patients mobility network

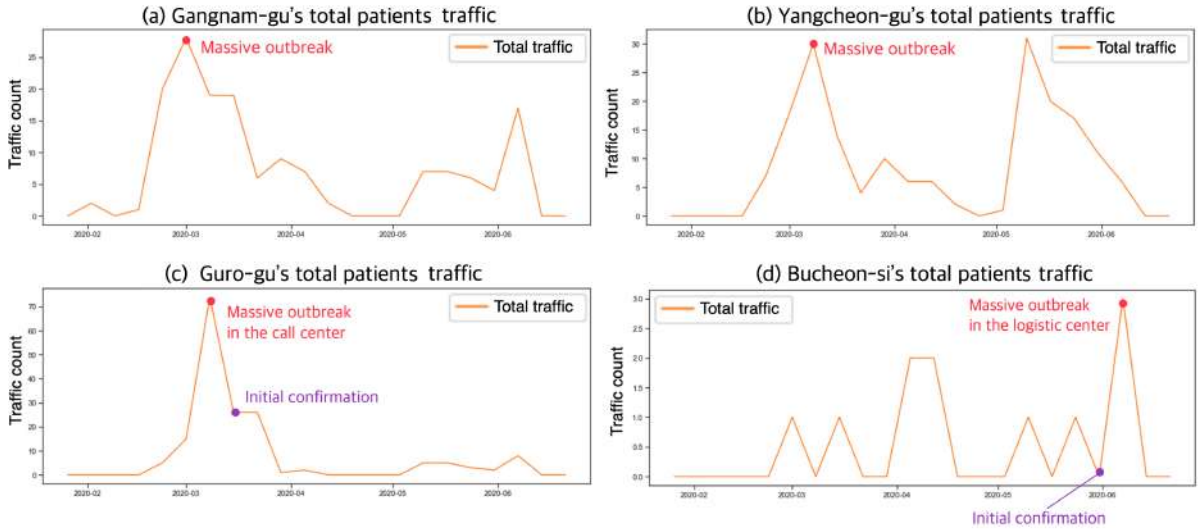

Figure 5.2.: The patient traffic volumes in the four cities.

5.3. Event characterization

In this case study, events are massive COVID-19 outbreaks in different urban areas. We define the epidemic outbreak as occurrence of the newly infected. We mainly focus on the spreads in the first wave. We analyze the details of the wave by sites. One notable analysis is our causal inference described in detail later section. On top of that, we assume the carriers' traffic surges directly leads to the outbreak. Then, we apply our anomaly characterization methods to the dynamic mobility network.

5.3.1. Event information

This subsection explains notable outbreaks in South Korea with social information, concentrating on the first wave. The plots (c), (d), and (e) in the Figure 5.1 indicate whether the traffic increased or decreased during the wave.

1. The first period (23.Feb - 29.Feb): One can see patients traffic on the city map. The darkest orange indicates **Gangnam-gu**, one of the biggest downtowns in South Korea. The traffic has increased by 19 during the period.

2. The second period (01.Mar - 06.Mar): No champion place. The traffic changes have more

even distribution than the other periods. Still, **Yangcheon-gu**(+11) and Geumcheon-gu(+13) show great increase. Yangcheon-gu has more incoming traffic than the outgoing one.

3. The third period(07.Mar - 13.Mar): The massive spread happened in **Guro-gu**. Traffic increased by 59, which is the most significant surge during the first wave. The core outbreak was reported at the local call center on March 9. Although there was a great social impact, the KDCA's response was prompt. They closed the building immediately and conducted an epidemiological investigation. They also sent a total of 16,628 text messages to citizens who stayed in the vicinity of the building for more than 5 minutes. Finally, they even tracked the patients' movements using mobile phone location data.[46]

Besides the first wave, there is an abnormal outbreak on May 25 at **Bucheon-si**, a western suburb city of Seoul. The town has a logistic center where 81 confirmed cases occurred, which is the core outbreak in this town. The reason for the massive spread was because they missed the golden time of the epidemiological investigation. This slow reaction generates an interesting, abnormal mobility pattern, and we will cover this later. The first verified person in the logistic center was confirmed on May 24, but it was after the eleven days since his initial symptom. One social issue is that the city has a different privacy policy from Seoul's, so most of the patients' routes are not published. Still, we found that the limited data provides notable facts about patient movement patterns in this area. Finally, we have attached South Korea's map to the appendix A.3 and highlighted the Bucheon city with magenta color.

The four plots in figure 5.2 explain the traffic of the virus carriers(patients) in the four places. As mentioned above, Guro-gu and Bucheon-si have 'core' outbreaks, where a super-influencer spread the virus in one building to many people. In the cases, we can specify the initial confirmation dates and highlighted them as purple dots. In Guro-gu's case, the patient traffic dropped drastically in that the Korean government rapidly reacted and prevented additional spread(plot (c)). In contrast, the plot (d) shows such a traffic peak after the confirmation. It is because the response was late; the government did not successfully control it.

Finally, although Gangnam-gu and Yangcheon-gu have two peaks in both waves, our analysis focus on the accidents in the first one.

5.3.2. Causal inference

This section quantifies the causal relation between the floating population(FP), the virus carriers' traffic(CT), and the number of confirmed cases(NC). We figure out that FP influences CT and NC, while CT does NC. However, the opposite causality does not hold. The figure 5.3 is a diagram of the cause and effect relationship.

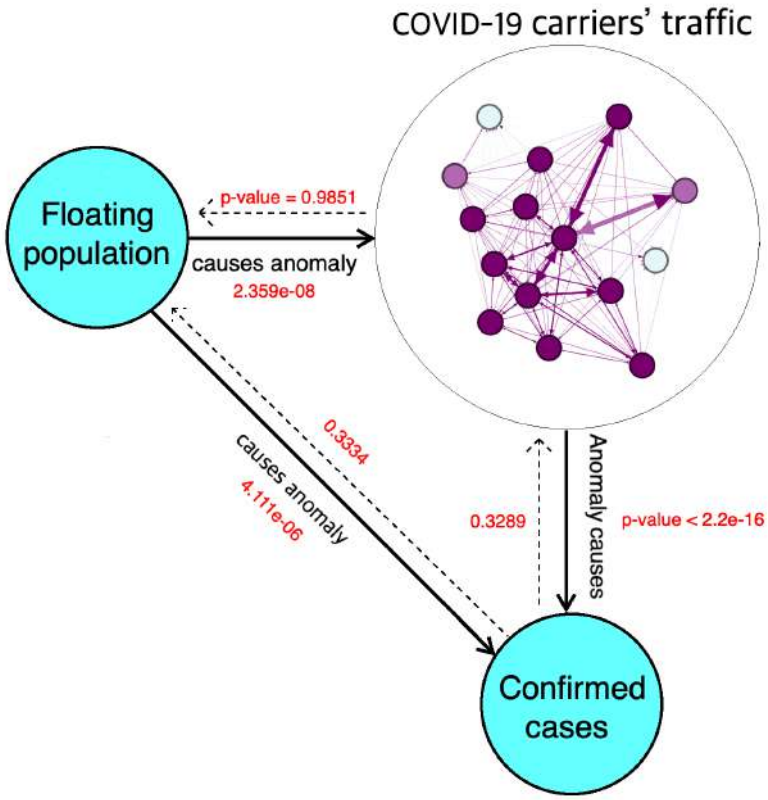

Figure 5.3.: Causality of COVID-19 patients mobility. The smaller p-values(the red numeric) indicate the more confident Granger causality.

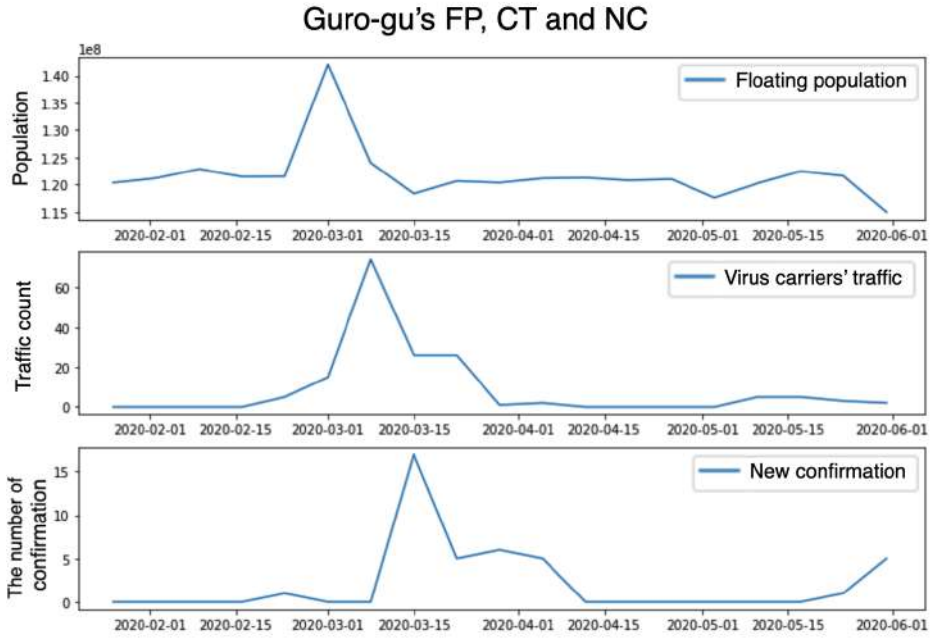

Figure 5.4.: Causal relations in Guro-gu

Furthermore, we analyze their relation in a single node. In particular, the figure 5.4 includes three plots for FP, CT, and NC in Guro-gu, which shows a time-shifting pattern. Specifically, the correlation between FP and CT is 0.8538, and the one between CT and NC 0.9408. Even though it is one of the most obvious examples, 22 of 25 Seoul districts still display similar time-shifting patterns. Also, this analysis focuses on only Seoul's towns because we only have data for Seoul's FP.

Additionally, such shifting patterns are an excellent example to measure Granger causality. We first did first-order differencing for every time series to make them stationary. According to the Augmented Dickey-Fuller test, all 25 FP, 24 CT, and 24 NC time-series become stationary, given the total 25 Seoul districts.

Next, we applied the Hurlin and Dumitrescu's hypothesis test to measure the causality from the panel data[14]. First of all, we build a panel dataset by aggregating the stationary dataset. The table 5.2 explains how the dataset appears. The first column, index is the district number. The time column means the time points from 0 to 17, where 0 stands for the earliest time, while the 17 does for the latest.

Then, we applied the hypothesis test to the panel data. The null hypothesis of the test is there is no Granger causality between any two variables at any place. The alternative one is Granger causality for at least one individual. Given a 0.01 of significant level, we concluded

index	time point	differenced FP	differenced PM	differenced CC
0	0	457368.	4	1
0	1	-1326818.	-2	-1
...	...	...	...	...
0	17	-5528325.36	9	0
1	0	1046770.33	0	0
...	...	...	...	...
28	16	-657775.	0	-1
28	17	-6239713.	3	-1

Table 5.2.: Panel data for COVID-19 patients' mobility

that the FP Granger-causes the PM and CC, and PM Granger-causes CC. In contrast, any opposite causality does not hold (i.e., the p-values for the inverse cause is even greater than 0.1). The red numbers in the figure 5.3 mean the p-values from the causality tests. We used only stationary time-series for each test and set the time lag parameter as 1.

Because the alternative statement only indicates Granger causality at least one place, we further measure concerning every node. With the time lag parameter L as 5, we implemented the Granger causality test for every 25 nodes whose differenced FP and CT are stationary. Because it is multiple hypothesis tests, we adjusted the significant level with Benjamini-Hochberg's methodology. The result shows 24/25 nodes show the Granger causality by FP to CT. In the same way, 22/23 nodes show that FP Granger causes NC. Finally, 23/23 nodes show the causality from CT to NC.

Therefore, it is a fair statement that an abnormal increase in the floating population leads to COVID-19 patients' mobility anomaly. Also, if the patients' traffic surges, the number of confirmation does too. Such rise is the epidemic outbreak, the main event of interest in this chapter.

5.3.3. Social-geological metrics

We use the social-geological metrics to achieve two goals: (1) detecting anomaly or pre- and post-signals when the actual patient traffic increases, (2) describing how social diversity changes in a region. Also, We introduce three case regions where the metrics present the influential pre-signals: Guro-gu, Yangcheon-gu, and Bucheon-si. We first explain the Guro-gu's representative case, and followed by Yangcheon-gu and Bucheon-si. Plus, the abnormalities detected by the metrics belongs to type A. The details are as follows: According to the figure 5.5, the traffic and homogeneity surge simultaneously in Guro-gu. We highlighted the time as red points in the figure. We computed the metrics by the weekly mobility data from 01.March to 07.March. Also, the green dot means when the first patient got confirmed.

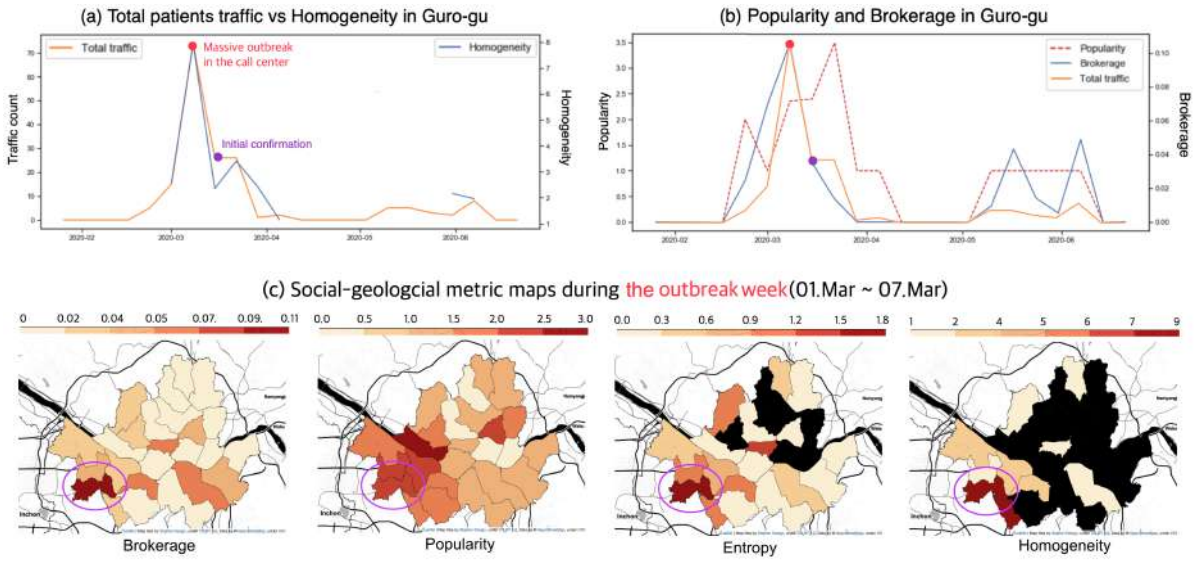

Figure 5.5.: Social-geological metrics in Guro-gu

As soon as the initial confirmation, the traffic rapidly decrease due to the early prevention by the government. The homogeneity is an interpretation that the patients come from similar home socially regions. Technically, a high homogeneity means that patients' home cities have similar outgoing traffic patterns.

Before actual traffic surges, we can observe pre-signals by brokerage and popularity. The plot (a) shows the brokerage more rapidly increases before the total traffic, the sum of incoming and outgoing traffic. Conceptually, the brokerage is equal to between centrality. The higher it is, the more patients will pass through that node to go their destination. Note that the traffic volume and brokerage have different patterns. The chart shows three brokerages peaks, while only one peak for the traffic. The popularity shows a local peak two weeks before the outbreak, and it reaches the maximum after two weeks. It plays as the pre- and post-signals of the local outbreak.

How were the four metrics of other districts during the event time? We compared the towns by city maps. The Guro-gu's brokerage, entropy, and homogeneity are exceptionally high during the traffic peak time. Its popularity is also high, but relatively similar to other regions'. It is because, as the plot indicates, the popularity did not surge during the time.

The figure 5.6 presents the Bucheon-si's case. It is interesting in that the rise in entropy precedes the initial confirmation. It indicates the heterogeneous patients arrived, which presumably causes the mass spread. However, because of the failed spread prevention, it finally shows the massive outbreak in June.

We also applied the method to the Gangnam- and Yangcheon-gu's mobility. The Yangcheon-

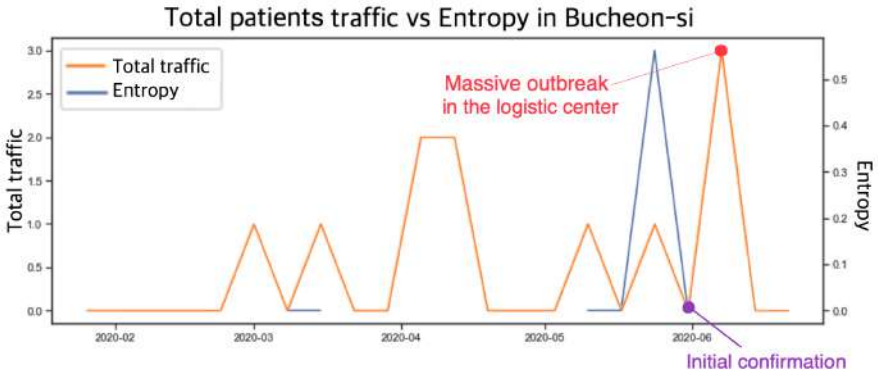

Figure 5.6.: Bucheon-si's Social-geological metrics

gu's popularity lasted as a post-signal and brokerage specify the two waves in the region. In Gangnam-gu, all four metrics change according to the two waves. However, in both regions, there is no pre-signal.

5.3.4. Neighborhood feature analysis

We analyze the representative graph feature of an anomaly to describe which feature stands for a local epidemic outbreak and to find pre-signal before a massive influence. The method identifies anomaly by comparing the current graph's features to the past pattern(type B anomaly). Because the methodology analyzes the relation between one node and its neighbors, it can more sensitively detect its local abnormality than the naive EDA. Note that neighborhood does not mean geologically close places, but connected nodes at a specific time in a dynamic mobility network.

The figure 5.7 explains Guro- and Gangnam-gu's cases regarding the methods. According to plot (a), The anomaly score already peaked when March started. Plus, it shows the second anomaly at the beginning of May because the algorithm detects an irregular traffic pattern between Guro-gu and other districts despite the similar traffic volume. Also, both degree and weight features recognize both two waves. It means the waves change the number of connections and traffic volume of the linked nodes to Gangnam-gu.

Another feature, the number of neighbors, has the same pattern as the traffic volume's, providing a pre-signal. Finally, we set the parameter W as 5 to analyze Guro-gu's case.

In the Gangnam-gu, we can observe the surging anomaly score from all 12 graph features mentioned in the previous section 3.1.4, given W is 3. It is because the algorithm sensitively responded to the initial patient in the district regardless of feature types. Still, the anomaly measure from the in-weight describes the traffic pattern best. It could also successfully monitor the second wave.

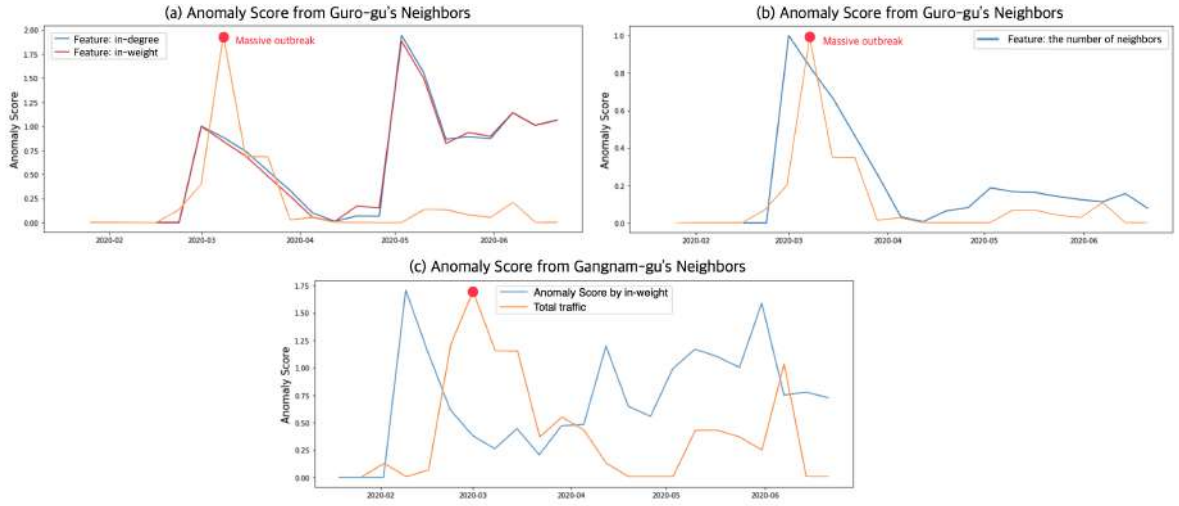

Figure 5.7.: Feature analysis from Guro- and Gangnam-gu's Neighborhoods

In case of Yancheon-gu and Bucheon-si, we could only figure out the representative features, not the pre-signals. The Yangcheon-gu's main features are In-weight, in-degree, and the number of neighbors. Bucheon-si's out-degree and out-weight well descrbies the traffic fluctuation. In both analyses, we set parameter W as 4.

5.3.5. Bilateral traffic model

In this subsection, we model the bilateral traffic for the two goals: (1) describing virus carriers' mobility by emission, attraction, and distance of two nodes, (2) providing a model whose explanatory variables are floating population. Regarding the first goal, we show that mobility is also explainable by the three factors similar to the classical gravity model. Considering the second point, we develop such a model to provide it to other researchers to estimate an urban epidemic carriers' traffic. In other words, the carriers' routes data is rare, usually due to privacy. Still, the floating population is more generally available, so we provide a model that can be used more generally.

We start from the first modeling. The blue dots in the figure show that the data violates the simple linear regression's homoscedasticity assumption: the residual variance is the same for any value of X. That is why we take the generalized modeling approach, as we did in the former chapter. Plus, we aggregate variables over the whole time to reduce noise, so we are thinking of a static graph. For example, the average E_i in the plot is the mean emission of individual node.

We build model whose formulation is as follows:

$$g(T_{ij}) = c + \alpha \cdot E_i + \beta \cdot A_j + \gamma \cdot D_{i,j} \tag{5.1}$$

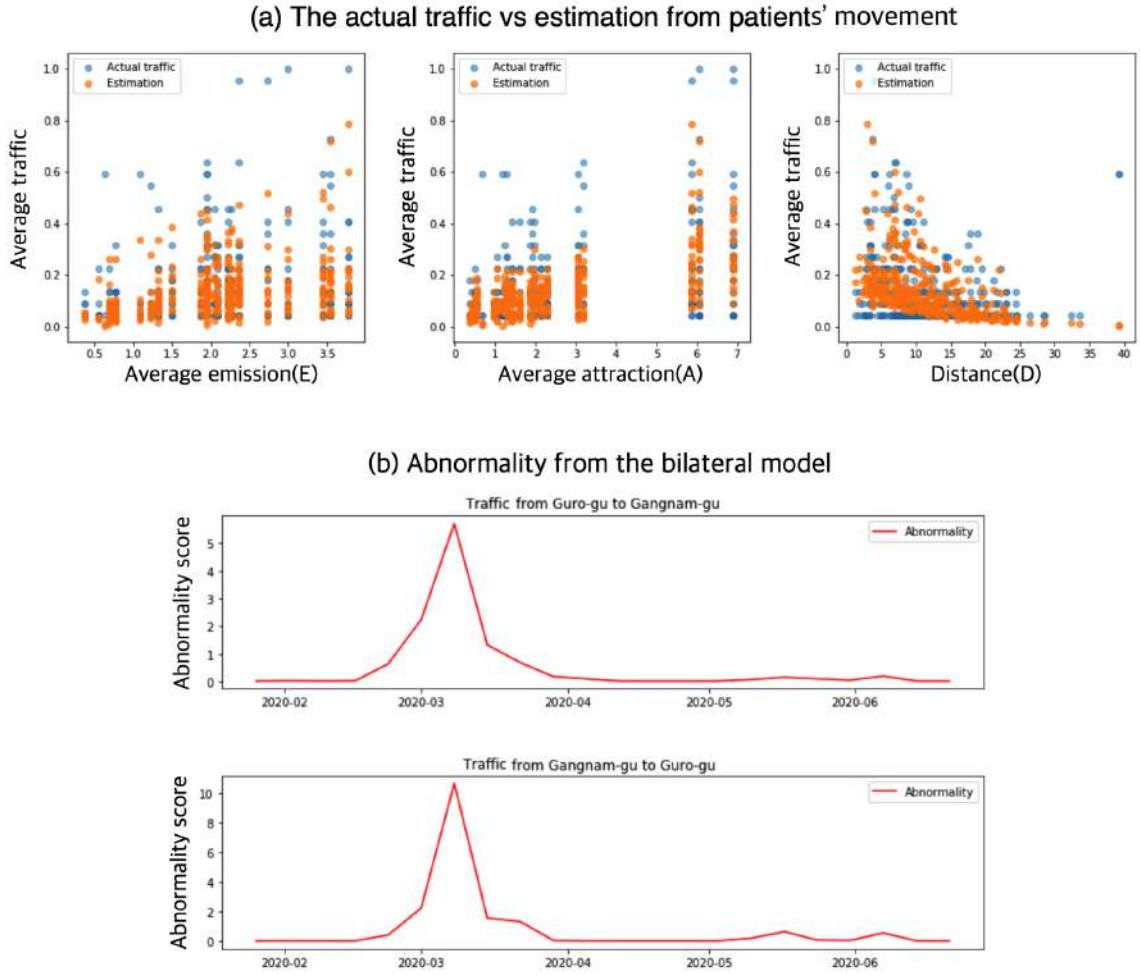

Figure 5.8.: Bilateral traffic modeling from COVID-19 patients' mobility

Distribution	Gaussian		Gamma			Inverse Gaussian		
Link function	μ	$log\mu$	μ	$log\mu$	μ^{-1}	μ	$log\mu$	μ^{-2}
Pearson χ^2	6.6834	**5.2673**	230.01	201.32	187.13	5727.7	-	1765.5
Deviance	6.6834	**5.2673**	565	339	357	6.88E+03	-	3.27E+03

Table 5.3.: Evaluation table for the bilateral patients mobility

where $g(\cdot)$ is a link function, and T_{ij} is our target variable. We evaluated different types of models, and the results are in the table where the distribution means T_{ij}'s probabilistic behavior. The final model has the Gaussian distribution with a logarithm link as follows:

$$log(T_{ij}) = -2.5895 + 0.3168 \cdot E_i + 0.2381 \cdot A_j - 0.0800 \cdot D_{i,j} \tag{5.2}$$

Note that the coefficient of E_i is bigger than A_j. It means out-going traffic more heavily affects the traffic between two nodes. The figure indicates, the model captures almost every mean pattern. Thus, we have shown that the bilateral traffic of the epidemic carriers can also be explained by adjusting the classical gravity model.

One issue is that when patients traffic abnormally surges, it significantly exceeds the estimated traffic. Thus we can measure the difference as an abnormality metric.

$$Abnormality(i,j) \coloneqq log(|T_{ij} - \hat{T_{ij}}| + 1) \tag{5.3}$$

The plot (b) in the figure 5.8 explains that the traffic from Gangnam- to Guro-gu is higher than the opposite direction. It means the former flow is more different from the estimation than the last one. Thus, the former is a more substantial abnormality regarding the standard traffic pattern. With the abnormality measure, we could specify anomalies of all four regions: Gangnam-gu, Yangcheon-gu, Guro-gu, and Bucheon-si.

The next model consists of exogenous explanatory variables: the floating populations. We build a simple linear model with log transformation as its R-squared is already 0.998, so more sophisticated modeling is unnecessary. The regression model is as follows:

$$log(T'_{ij}) = 0.2905 + 0.0298 \cdot log(F'_i) + 0.0220 \cdot log(F'_j) - 0.0415 \cdot log(D'_{ij}) \tag{5.4}$$

where $X' = X + e$ and F_i stands for the average floating population in node i. Note that T'_{ij} is also average value over time.

Then, when does the model 5.4 perform poorly? The figure 5.9 suggests the original observations(blue dots) and the model's estimation(orange dots). The blue dots display little overdispersed variances. According to the plots, the high traffic volume causes outliers. It turns out that the outliers with $log(T'_{ij}) > 1.25$ are the nodes number 0, 1, 2. They have the noticeably highest incoming traffic, but their outgoing ones are natural. Presumably, it is a characteristic of an epidemic patients' mobility that the virus more actively and recursively spread when the population size increases, leading to more patient movement.

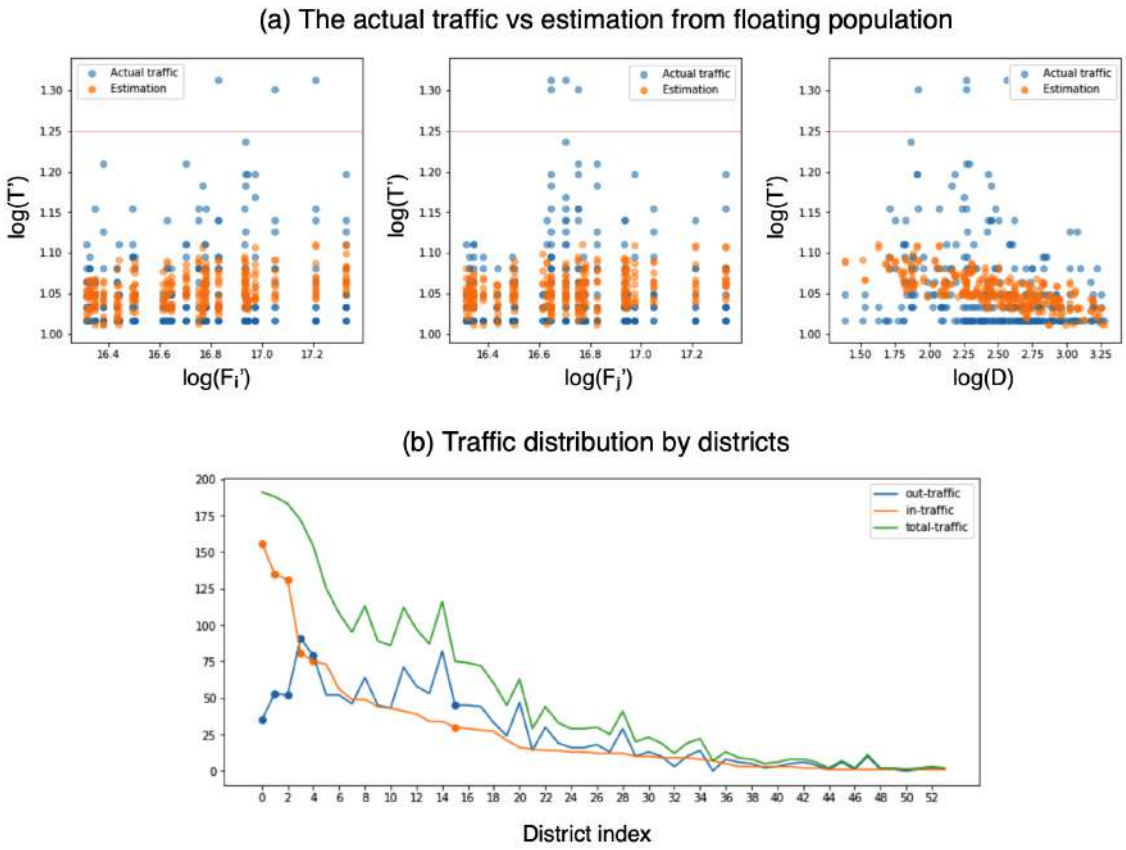

Figure 5.9.: Bilateral traffic modeling from Seoul floating population

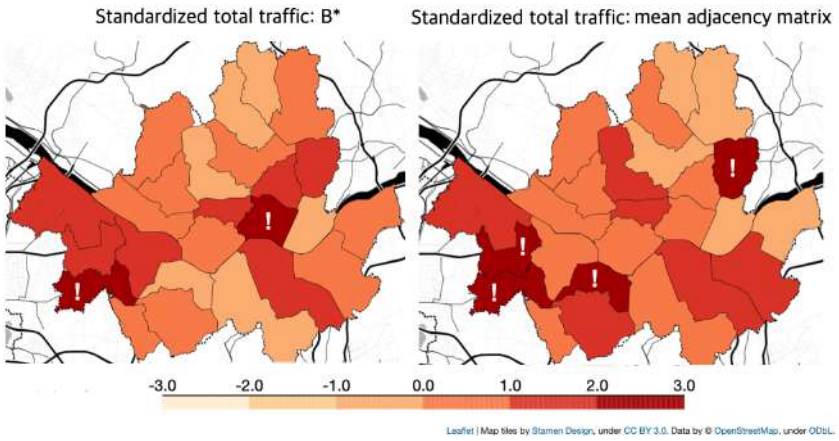

Figure 5.10.: Seoul traffic comparison: B* vs average pattern

5.3.6. Standard mobility graph comparison

In this section, we compare the virus carriers' traffic pattern to the standard mobility graph for the following goals: (1) to measure an anomaly's strength, beyond the naive EDA, and (2) to compute the regular mobility graph by reducing abnormal observations. Note that the algorithm needs to set initial matrixes, which we did randomly in this analysis. Therefore, the below statement holds in general, but specific numbers vary every time. One may initialize with constant matrixes by SVD, but it underperforms for this carrier mobility analysis.

Recall the algorithm computes B* which stands for the most general basic mobility pattern, which has less anomaly. Specifically, the figure's left map shows the B*. In contrast, the right side indicates the traffic from the mean adjacency matrix. The color means standardized total traffic. Here the total traffic is the sum of incoming and outgoing flows, and standardization means transform average to zero and variance to one. Plus, the darkest districts with exclamation marks means the outliers regarding traffic volume. The right one displays four such regions, while the left one does only two. It means the regularized graph has fewer outliers.

After acquiring B*, we computed anomaly scores by LOF. The figure describes that the algorithm successfully recognizes anomalies in that the anomaly scores surge whenever outbreaks occur. The orange line means the total traffic, while the blue one is the algorithm's anomaly score.

The method measures local abnormality degrees. For example, the maximum score in Guro-gu exceeds 4.0. The value represents how one node's pattern is different from the B*. Often, abnormality degrees differ even though traffic volume is profoundly similar. For example, the maximum total traffic in Yangcheon-gu during the first wave was 30. The maximum value becomes 31 during the second wave period. Thus the raw traffic volume

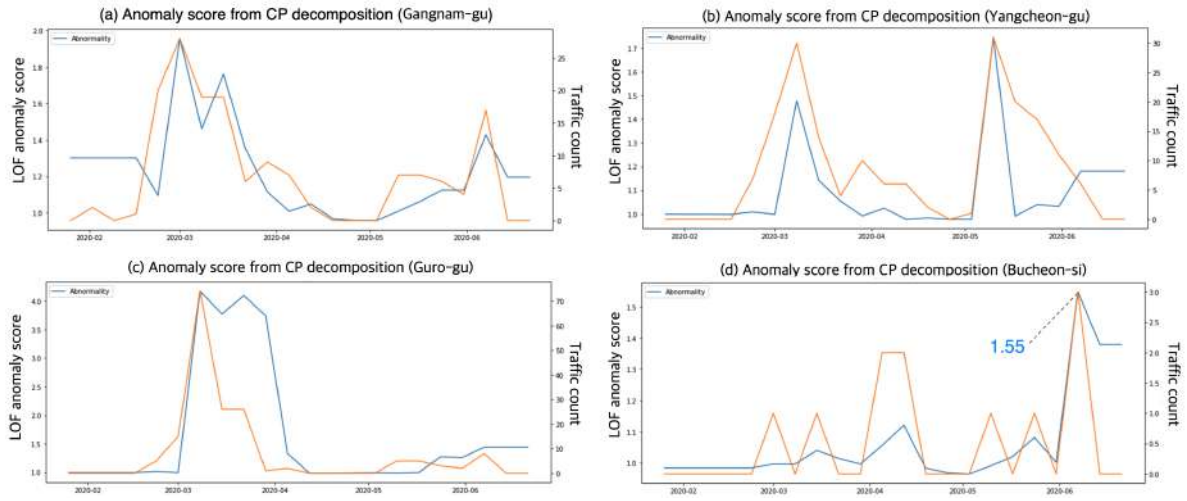

Figure 5.11.: Anomaly scores by standard mobility graph comparison

surges by 3%. On the other hand, the anomaly score increases by 18%. It is because the algorithm does consider not only the mobility volume but also graph structure.

Note that we cannot compare the anomaly scores between different regions. For example, the maximum score in Bucheon-si is 1.55, but it does not mean the absolute degree of abnormality. Therefore it is not reasonable to compare the value to other city's anomaly scores. It only measures relative severity on a local scale.

We can also observe the lasting impact of outbreaks. For example, the anomaly score lasted bigger than 3.5 during March in the plot (c), and 1.3 during June in the plot (d). It indicates the impact of the outbreak did last for a month in the dynamic network. We did not see such post-signal from the naive EDA.

Recall that the analysis does not correctly detect all the three events in the former chapter. The Twitter geotags data involves a confusing trend and more fluctuates due to its complexity(i.e., any minor event in the world influences the inter-country mobility). In contrast, patients' mobility is less complicated and noisy. That is why we could detect all the outbreaks in the four regions.

5.3.7. Dynamic graph prediction

Here we report one failure case: characterizing anomaly of patient mobility by graph prediction. The given patients' network is so sparse that we do not have sufficient training datasets to build the prediction model. The figure 5.12 explains the training process. It shows that the maximum validation F1 score is less than 0.25. The score only increases when the epoch is below 50 but starts to decrease after it exceeds 100, over-fitting problem. The validation

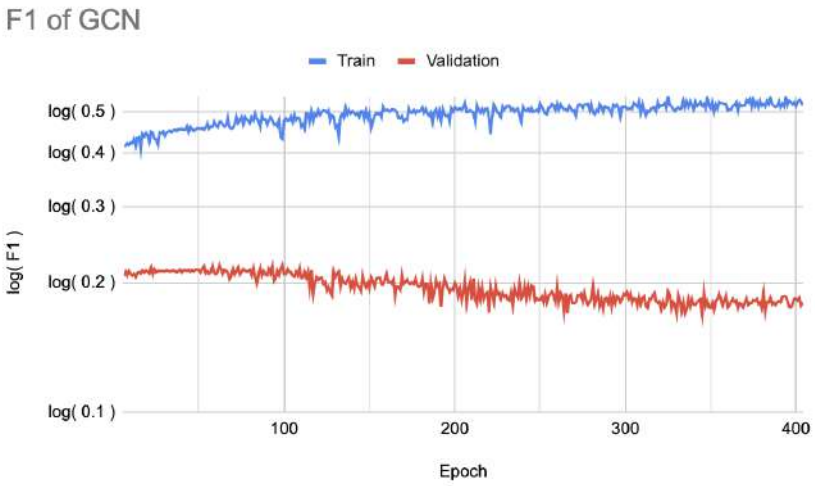

Figure 5.12.: Train and Validation F1 scores

period is from 26.Apr to 23.May in 2020. The COVID-19 mobility data has only 54 nodes and 21 weeks, while the Twitter dataset has 226 nodes and 220 weekly time intervals.

To solve the issue, we tried a bootstrap method, which improved the performance slightly. Specifically, the original dataset has a sparse matrix format, including triples (origin, destination, weight of edges). We resampled some triplets to acquire 100 sub-trainsets and calculated the average of 100 predictions. This bootstrap increased test F1 by 0.025, from 0.1409 to 0.1662. However, it is still not accurate, considering that the F1 scores for Twitter are over 0.60.

To summarize, the graph prediction is not appropriate for small mobility datasets. The performance can be partially complemented by bootstrapping, but still not enough.

5.4. Summary

So far, we have analyzed anomalies from the COVID-19 virus careers' mobility network that leads to the epidemic outbreak. In detail, we used COVID-19 patients dataset in South Korea and built spatio-temporal graph analysis to unveil the human dynamics. The main events in this chapter were the epidemic outbreaks in different regions. We defined the outbreak as the occurrence of new influences. We specifically focus on the outbreaks in the four regions: Gangnam-gu, Yangcheon-gu, Guro-gu, and Bucheon-si.

We also quantified the causal links from the population size to that of virus carriers, and finally to the newly infected volume. To do this, we tested Granger-causality for the panel data and individual nodes. The table 5.4 summarizes the result. The third column means how many districts in Seoul show such causality by the multiple hypothesis tests. Thus we

assumed the virus carriers' large traffic directly leads to the newly infected.

By our five methods, we understood the various properties of the urban outbreaks and the mobility network. First approach was social-geological metrics. We interpreted how social diversity and human mobility changeby the metrics. Besides, they suggested pre- and post-signals when the carriers' traffic increases. In Guro-gu and Bucheon-si, Popularity, Brokerage, and Entropy act as pre-signal before actual traffic surge. Popularity does as post-signals showing how the outbreak' impact lasted in Guro- and Yancheon-gu.

Next, we applied the methods of neighborhood feature analysis. We could find pre-signals in Guro- and Gangnam-gu cases by in-degree, in-weight, and the number of connected nodes. Again, the method worked to describe mobility in the other regions, without pre-signals.

By bilateral traffic model, we figure out the carriers' mobility also can be modeled by traffic emission, attraction, and distance similar to the classical gravity model. We detected all four regions' bilateral traffic anomalies and measured their strength. In addition, we suggest another bilateral model to estimate the virus carriers by only using floating populations.

Next, we computed and visualized the regular mobility graph of the carriers. By comparing the standard mobility graph, we could find post-signals not observable by the naive EDA. The tables 5.5 and 5.6 summarize the contents.

This case study is distinctive because we analyzed virus carriers' dynamics(not the public mobility) and figure out the anomaly regarding the outbreak. Also, we found pre- and post-signals of the anomalies and interpret/unveil the anomalies within the spatio-temporal network by using other studies. Especially, the early signal indicates a chance to prevent the outbreak if we could track weekly mobility of the patients.

Cause	Effect	Proportion in Seoul
Floating population	Virus carriers' traffic volume	24/25
Floating population	The number of confirmation	22/23
Virus carriers' traffic volume	The number of confirmation	23/23

Table 5.4.: Summary of causal inference

Characterization Methods	Anomaly detection	Pre-signal	Post-signal	Result
Naive EDA	Yes	No	No	The most naive traffic volume exploration
Socio-geological metrics	Yes	Yes	Yes	Change of social diversity for each node
Neighborhood feature analysis	Yes	Yes	No	Anomaly's representative graph features
Bilateral traffic model	Yes	No	No	Traffic estimated by attraction, emission, and distance, or floating population.
Standard mobility graph comparison	Yes	No	Yes	Regular mobility graph, post-signals, anomaly strength.
Dynamic graph prediction	No	-	-	Not trained due to the sparsity

Table 5.5.: Summary of the methods application

	Events: COVID-19 massive outbreaks			
Location	Gangnam-gu	Yangcheon-gu	Guro-gu	Bucheon-si
Traffic peak time	23.Feb−29.Feb	01.Mar−07.Mar	01.Mar−07.Mar	31.May−06.June
Accident by a super influencer?	No	No	Yes	Yes
Initial confirmation of the influencer	-	-	08.May	24.May
Social-geological metrics	entropy, homogeneity, brokerage, popularity	Brokerage, Popularity	Brokerage, Homogeneity, Popularity	brokerage, entropy, popularity
Pre-signal	-	-	Popularity, Brokerage	Entropy
Post-signal	-	Popularity	Popularity	-
Representative feature of the local anomaly	In-weight	In-weight, in-degree, neighbors	In-weight, in-degree, neighbors	Out-degree, out-weight
Pre-signal	In-weight	-	In-weight, in-degree, neighbors	-
Post-signal by $B*$ comparison	No	No	Yes	Yes

Table 5.6.: Summary of the local events

6. Conclusion

6.1. Summary

As was discussed in the introduction, we have answered our first research question, 'How do social events affect and are affected by anomalies in a spatio-temporal human mobility network?'. To answer the question, we set two sub-targets: (1) to offer methodologies that describe an abnormal event and mobility anomaly, and (2) to practice the techniques supplying an understanding through actual case studies. We achieved the two sub-goals and answered the main research questions at the end of each case study. In this final chapter, we compile this thesis and recommend potential future works.

Initially, we presented related recent articles. Some researchers verified the impacts of specific social issues on human mobility. Others developed algorithms to identify anomalies in a dynamic network. Building on the previous works, we suggested general methodologies that describe the causality between events and anomalies by a dynamic mobility network analysis.

Then we proposed the details of our methodologies. We split them into two categories: (1) anomaly characterization methods, (2) causal inference ones. The methods in the first category detect and analyze the three types of anomalies. Type A is an outlier regarding a metric, and Type B is an observation differing from the past mobility pattern. At last, type C is an observation differing from the standard mobility pattern. Next, the causal inference methods include statistical concepts. They are Granger-causality for panel data, hypothesis tests for stationarity, and multiple hypothesis experiments.

After proposing methodologies, we applied them to the two heterogeneous mobility systems. The first one is about Twitter users' global travel patterns, and the other is about COVID-19 patients' urban mobility.

In the first case, we assume a one-directional causal relationship: exogenous event affects the mobility system. We figured out that our methods unveil anomalies that are not visible via simple EDA. Plus, we describe how three social events(Hong Kong's demonstration, Russia's World Cup, Indonesia's earthquake) change the dynamic mobility network structure.

All methods operated successfully in this first case study except causal inference. The methods found representative features and social-geological metrics of anomalies. Also, we

could compare the predicted graph state to the disturbed one by exogenous events. The bilateral model measured the different anomaly period by traffic directions. Finally, we measured their abnormality scores by comparing one normal mobility state. We described the anomalies' characteristics through the multiple applications and summarized the result in the tables 4.5 and 4.6.

The next chapter characterizes anomalies from the COVID-19 virus carriers' mobility network. One core difference is that anomaly within the mobility system causes abnormal events. We suggested and quantified causal relationships of three factors: floating population, the virus carriers' traffic, and the number of confirmed cases. On top of the causality, we assume the carrier's traffic surge leads to the epidemic outbreak.

Then we applied the characterization methods and drew interpretations. Representative graph feature analysis and social-geological metrics found pre-and post-signals when the traffic volume increases. The standard graph comparison way found only post-signals. The result indicates a possibility to prevent massive epidemic spread by detecting the early signals.

The bilateral model can measure two degrees of bilateral anomalies. We also show that we can model the carriers' traffic volume by the floating population. However, graph prediction does not work due to the network's sparsity. We presented the result in the table 5.5 and 5.6.

The second case study is a unique analysis of virus carriers' dynamics(not the public mobility) and causality between the anomalies and the outbreak. In particular, the early signal indicates that we will be able to prevent the outbreak, given the virus carriers' weekly mobility.

The two case studies suggest our anomaly characterizing methods are available given heterogeneous properties of a mobility system. They operate regardless of mobility scale(inter- or intra-country), complexity, trend, seasonality, a hub node's existence. However, the graph sparsity heavily affects the performance of methodology based on a neural network.

In contrast, we could only infer causality from the less complex system, second case. Because the first system covers global travels, every minor factor affects the mobility pattern. However, the second only handles urban virus carriers' mobility, so less complicated than the first system; there is little noisy traffic variance. Besides, our target event is only an epidemic outbreak in different districts. Thus we could quantify that their traffic leads to an epidemic outbreak.

Understanding anomaly patterns from human mobility is crucial and applicable in various fields, including epidemic prevention, social marketing, transport system design. We hope our study help researchers in such domains.

6.2. Future works

We propose several lines of research arising from this work that needs further examination. Firstly, our work only characterizes events causing drastic change. For example, we did not detect Brexit's progressive impact on social mobility for the long term. Presumably, such an event changes a long term trend. Even though the anomaly is not readily observable, it is meaningful to measure such a major social issue's impact.

The second point is causal analysis from a complex system. As we observed in the previous chapter, global traffic varies for unknown reasons. For example, in the figure 4.3, Indonesia's travel counts dropped in July.2019. Also, the figure 4.2 shows varying traffic in Russia, but we do not understand the causes. Such noise makes researchers hard to interpret an impact by an exogenous abnormal event.

The final work is to track the infected people. To capture the pre-signal before the epidemic outbreak, one needs to track patients. However, it is not reasonable to track them in real-time. Hopefully, we found that the floating population is heavily correlated with, and does granger-cause the virus carriers' mobility. Therefore, by urban crowd estimation, one can capture the signal and prevent the massive outbreak.

A. Appendix

A.1. Country index and abbreviation

Table A.1.: Index of top 30 districts in Seoul metropolitan regarding total traffic

Index	Country name	Abbreviation
0	the United States	US
1	the United Kingdom	GB
2	France	FR
3	Spain	ES
4	Italy	IT
5	Germany	DE
6	Mexico	MX
7	Canada	CA
8	Japan	JP
9	Netherlands	NL
10	Brazil	BR
11	Thailand	TH
12	Indonesia	ID
13	Malaysia	MY
14	Singapore	SG
15	Belgium	BE
16	United Arab Emirates	AE
17	Australia	AU
18	Switzerland	CH
19	Philippines	PH
20	Ireland	IE
21	Portugal	PT
22	Argentina	AR
23	Turkey	TR
24	Russia	RU
25	Kuwait	KW
26	China	CN
27	India	IN
28	Hong Kong	HK
29	Austria	AT

A.2. Index of Seoul metropolitan districts

Table A.2.: The index of the top 30 districts in Seoul metropolitan regarding total traffic

Index	Area name
0	Jungnang-gu
1	Dongjak-gu
2	Yangcheon-gu
3	Guro-gu
4	Gangnam-gu
5	Jung-gu
6	Jongno-gu
7	Eunpyeong-gu
8	Gangseo-gu
9	Yeongdeungpo-gu
10	Dongdaemun-gu
11	Songpa-gu
12	Seongdong-gu
13	Geumcheon-gu
14	Gwanak-gu
15	Seocho-gu
16	Seodaemun-gu
17	Gangbuk-gu
18	Yongsan-gu
19	Nowon-gu
20	Mapo-gu
21	Ansan-si
22	Seongbuk-gu
23	Goyang-si
24	Seongnam-si
25	Gwangjin-gu
26	Dobong-gu
27	Namyangju-si
28	Gangdong-gu
29	Anyang-si

A.3. Map of Gyeonggi-do

Figure A.1.: Map of Gyeonggi-do, South Korea
Image source: https://english.gg.go.kr/

List of Figures

List of Tables

Bibliography

[1] F. Harary and G. Gupta. "Dynamic graph models". In: *Mathematical and Computer Modelling* 25.7 (1997), pp. 79–87.

[2] D. Rind. "Complexity and climate". In: *science* 284.5411 (1999), pp. 105–107.

[3] J. Ladyman, J. Lambert, and K. Wiesner. "What is a complex system?" In: *European Journal for Philosophy of Science* 3.1 (2013), pp. 33–67.

[4] S. und Weltwirtschaft. "No man is an island". In: (2009).

[5] S. Ranshous, S. Shen, D. Koutra, S. Harenberg, C. Faloutsos, and N. F. Samatova. "Anomaly detection in dynamic networks: a survey". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 7.3 (2015), pp. 223–247.

[6] L. Akoglu and C. Faloutsos. "Event detection in time series of mobile communication graphs". In: *Army science conference*. Vol. 1. 2010.

[7] C. Lin, Q. Zhu, S. Guo, Z. Jin, Y.-R. Lin, and N. Cao. "Anomaly detection in spatiotemporal data via regularized non-negative tensor analysis". In: *Data Mining and Knowledge Discovery* 32.4 (2018), pp. 1056–1073.

[8] T. N. Kipf and M. Welling. "Semi-supervised classification with graph convolutional networks". In: *arXiv preprint arXiv:1609.02907* (2016).

[9] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li. "T-gcn: A temporal graph convolutional network for traffic prediction". In: *IEEE Transactions on Intelligent Transportation Systems* (2019).

[10] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. B. Schardl, and C. E. Leiserson. "EvolveGCN: Evolving Graph Convolutional Networks for Dynamic Graphs." In: *AAAI*. 2020, pp. 5363–5370.

[11] K. Lei, M. Qin, B. Bai, G. Zhang, and M. Yang. "GCN-GAN: A non-linear temporal link prediction model for weighted dynamic networks". In: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE. 2019, pp. 388–396.

[12] C. W. Granger. "Investigating causal relations by econometric models and cross-spectral methods". In: *Econometrica: journal of the Econometric Society* (1969), pp. 424–438.

[13] J. Runge. "Detecting and quantifying causality from time series of complex systems". PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2014. DOI: http://dx.doi.org/10.18452/17017.

[14] E.-I. Dumitrescu and C. Hurlin. "Testing for Granger non-causality in heterogeneous panels". In: *Economic modelling* 29.4 (2012), pp. 1450–1460.

[15] M. Eichler. *Causal inference in time series analysis*. Wiley Online Library, 2012.

[16] J. Pearl et al. "Causal inference in statistics: An overview". In: *Statistics surveys* 3 (2009), pp. 96–146.

[17] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth. "Understanding human mobility from Twitter". In: *PloS one* 10.7 (2015), e0131469.

[18] Y. Martıén, Z. Li, and S. L. Cutter. "Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew". In: *PLoS one* 12.7 (2017), e0181701.

[19] A. Sen and L. W. Dietz. "Identifying travel regions using location-based social network check-in data". In: *Frontiers in Big Data* 2 (2019), p. 12.

[20] F. Schlosser, B. F. Maier, D. Hinrichs, A. Zachariae, and D. Brockmann. "COVID-19 lockdown induces structural changes in mobility networks–Implication for mitigating disease dynamics". In: *arXiv preprint arXiv:2007.01583* (2020).

[21] A. Galeazzi, M. Cinelli, G. Bonaccorsi, F. Pierri, A. L. Schmidt, A. Scala, F. Pammolli, and W. Quattrociocchi. "Human Mobility in Response to COVID-19 in France, Italy and UK". In: *arXiv preprint arXiv:2005.06341* (2020).

[22] S. M. Iacus, C. Santamaria, F. Sermi, S. Spyratos, D. Tarchi, and M. Vespe. "Human mobility and COVID-19 initial dynamics". In: *Nonlinear Dynamics* (2020), pp. 1–19.

[23] M. U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, N. R. Faria, R. Li, W. P. Hanage, et al. "The effect of human mobility and control measures on the COVID-19 epidemic in China". In: *Science* 368.6490 (2020), pp. 493–497.

[24] A. M. Chu, A. Tiwari, and M. K. So. "Detecting early signals of COVID-19 global pandemic from network density". In: *Journal of travel medicine* 27.5 (2020), taaa084.

[25] A. Kapoor, X. Ben, L. Liu, B. Perozzi, M. Barnes, M. Blais, and S. O'Banion. "Examining COVID-19 Forecasting using Spatio-Temporal Graph Neural Networks". In: *arXiv preprint arXiv:2007.03113* (2020).

[26] D. Hristova, M. J. Williams, M. Musolesi, P. Panzarasa, and C. Mascolo. "Measuring urban social diversity using interconnected geo-social networks". In: *Proceedings of the 25th international conference on world wide web*. 2016, pp. 21–30.

[27] U. Brandes. "A faster algorithm for betweenness centrality". In: *Journal of mathematical sociology* 25.2 (2001), pp. 163–177.

[28] T. M. Fruchterman and E. M. Reingold. "Graph drawing by force-directed placement". In: *Software: Practice and experience* 21.11 (1991), pp. 1129–1164.

[29] S. G. Kobourov. "Spring embedders and force directed graph drawing algorithms". In: *arXiv preprint arXiv:1201.3011* (2012).

[30] V. A. Traag, L. Waltman, and N. J. van Eck. "From Louvain to Leiden: guaranteeing well-connected communities". In: *Scientific reports* 9.1 (2019), pp. 1–12.

[31] M. E. Newman. "Finding community structure in networks using the eigenvectors of matrices". In: *Physical review E* 74.3 (2006), p. 036104.

[32] M. Chen, K. Kuzmin, and B. K. Szymanski. "Community detection via maximization of modularity and its variants". In: *IEEE Transactions on Computational Social Systems* 1.1 (2014), pp. 46–65.

[33] M. Girvan and M. E. Newman. "Community structure in social and biological networks". In: *Proceedings of the national academy of sciences* 99.12 (2002), pp. 7821–7826.

[34] J. Reichardt and S. Bornholdt. "Statistical mechanics of community detection". In: *Physical review E* 74.1 (2006), p. 016110.

[35] V. A. Traag, R. Aldecoa, and J.-C. Delvenne. "Detecting communities using asymptotical surprise". In: *Physical Review E* 92.2 (2015), p. 022816.

[36] G. Rossetti, L. Milli, and R. Cazabet. "CDLIB: a python library to extract, compare and evaluate communities from complex networks". In: *Applied Network Science* 4.1 (2019), p. 52.

[37] A. S. Fotheringham and C. Brunsdon. "Local forms of spatial analysis". In: *Geographical analysis* 31.4 (1999), pp. 340–358.

[38] T. M. Oshan. "A primer for working with the Spatial Interaction modeling (SpInt) module in the python spatial analysis library (PySAL)". In: *Region* 3.2 (2016), R11–R23.

[39] B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti. "Geo-located Twitter as proxy for global mobility patterns". In: *Cartography and Geographic Information Science* 41.3 (2014), pp. 260–271.

[40] L. Lopez and S. Weber. "Testing for Granger causality in panel data". In: *The Stata Journal* 17.4 (2017), pp. 972–984.

[41] Y. Benjamini and Y. Hochberg. "Controlling the false discovery rate: a practical and powerful approach to multiple testing". In: *Journal of the Royal statistical society: series B (Methodological)* 57.1 (1995), pp. 289–300.

[42] C. Widmer. "Network dynamics in spatial networks inferred from social media interactions". In: (2019).

[43] A. Mislove, S. Lehmann, Y.-Y. Ahn, J.-P. Onnela, and J. N. Rosenquist. "Understanding the demographics of twitter users." In: *Icwsm* 11.5th (2011), p. 25.

[44] L. Sloan and J. Morgan. "Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter". In: *PloS one* 10.11 (2015), e0142209.

[45] T. Mayer and S. Zignago. "GeoDist: the CEPII's distances and geographical database". In: (2006).

[46] S. Y. Park, Y.-M. Kim, S. Yi, S. Lee, B.-J. Na, C. B. Kim, J.-i. Kim, H. S. Kim, Y. B. Kim, Y. Park, et al. "Early Release-Coronavirus Disease Outbreak in Call Center, South Korea". In: (2020).