Dynamic Network Anomaly Modeling of Cell-Phone Call Detail Records for Infectious Disease Surveillance

Carl Yang[*,†]
Department of Computer Science, Emory University
Atlanta, GA, United States
j.carlyang@emory.edy

Hongwen Song[*]
Department of Computer Science, Emory University
Atlanta, GA, United States
hongwen.song@emory.edu

Mingyue Tang
Department of Engineering Systems and Environment, UVA
Charlottesville, VA, United States
utd8hj@virginia.edu

Leon Danon
Department of Engineering Mathematics and Bristol Vaccine Centre, University of Bristol
Bristol, United Kingdom
l.danon@bristol.ac.uk

Ymir Vigfusson[†]
Department of Computer Science, Emory University
Atlanta, GA, United States
ymir.vigfusson@emory.edu

ABSTRACT

Global monitoring of novel diseases and outbreaks is crucial for pandemic prevention. To this end, movement data from cell-phones is already used to augment epidemiological models. Recent work has posed individual cell-phone metadata as a universal data source for syndromic surveillance for two key reasons: (1) these records are already collected for billing purposes in virtually every country and (2) they could allow deviations from people's routine behaviors during symptomatic illness to be detected, both in terms of mobility and social interactions. In this paper, we develop the necessary models to conduct population-level infectious disease surveillance by using cell-phone metadata individually linked with health outcomes. Specifically, we propose GraphDNA—a model that builds Graph neural networks (GNNs) into Dynamic Network Anomaly detection. Using cell-phone call records (CDR) linked with diagnostic information from Iceland during the H1N1v influenza outbreak, we show that GraphDNA outperforms state-of-the-art baselines on individual Date-of-Diagnosis (DoD) prediction, while tracking the epidemic signal in the overall population. Our results suggest that proper modeling of the universal CDR data could inform public health officials and bolster epidemic preparedness measures.

CCS CONCEPTS

• **Applied computing** → **Health informatics**; • **Information systems** → *Mobile information processing systems.*

KEYWORDS

disease surveillance, cell-phone call detail records, temporal networks, anomaly analysis, graph neural networks

KDD '22, August 14–18, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9385-0/22/08…$15.00
https://doi.org/10.1145/3534678.3542678

ACM Reference Format:
Carl Yang, Hongwen Song, Mingyue Tang, Leon Danon and Ymir Vigfusson. 2022. Dynamic Network Anomaly Modeling of Cell-Phone Call Detail Records for Infectious Disease Surveillance. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3534678.3542678

1 INTRODUCTION

The COVID-19 pandemic underscores the need for early outbreak detection and infectious disease surveillance. In normal times, public health officials continuously monitor emerging pathogens and smaller epidemics to mitigate the chances for any of these turning into a global pandemic. These efforts include *syndromic surveillance* where multiple data sources, such as hospital records, cross-sectional surveys, or even search-engine queries are searched for clusters of symptoms that warrant further scrutiny. For diseases where symptoms coincide with the infectious period, such as most influenza variants, such symptomatic surveillance can further track the progression of an epidemic and provide direct feedback for mitigation strategies, such as quarantines, lock-downs, or vaccinations.

Recent efforts have advanced cell-phone metadata, such as the *call-detail records* (CDR), as a potential universal data source to augment symptomatic surveillance [7, 29, 45]. First, CDR data include the (anonymized) caller and recipient numbers, a timestamp of the call or text, and the GPS-coordinates of the cellular tower through which the call was routed. They thus provide time-series for individual mobility and social interactions—behaviors that may differ when the person is ill (cf. studies such as [46] on the connection between cell phone calls and physical contacts). Second, in contrast with aggregated mobility models [12], CDR data may be linked with health data at the *individual* level while accommodating privacy concerns [45], allowing deviations from individual routines—such as staying home when ill—to be detected. Third, CDR data are already recorded by virtually every mobile-network provider for billing purposes within an established regulatory and privacy framework. Disease monitoring using an existing data source, such

[*]Equal contribution.
[†]Corresponding.

as CDR logs, is easier and cheaper than alternatives. Important ethical and privacy concerns can be addressed using data deidentification protocols between health officials and mobile operators (cf. Appendix A). Finally, cell-phone use is ubiquitous (105 mobile subscriptions per 100 inhabitants; 97% of the world population covered by a mobile network) whereas Internet access is less pervasive (57% of the world population) and heavily skewed towards affluent regions (19% of individuals in the least developed countries (LDC) have Internet access), according to 2020 estimates [26]. Many lower and middle-income countries lack resources for direct public health monitoring, standing to benefit most from inexpensive disease monitors.

Key technical challenges must be resolved to make individual CDR-based methods practical for epidemiological surveillance. Using linked health and CDR data from the H1N1v epidemic in Iceland in 2009, Vigfusson *et al.* [45] showed that individual mobility is reduced around the day of influenza-like illness diagnosis. While it is interesting that infection produces measurable behavioral changes from sparse CDR data, the key question is whether measurable behavioral changes imply infection symptoms, which would permit estimation of the number of people that are taken ill at a point in time (symptomatic prevalence). This direction is challenging for several reasons, including networked signals (involving individuals' behaviors regarding both themselves and their social contacts), temporal routines (requiring the capture of dynamic behavioral patterns), and weak supervision (because disease labels are sparse and only weakly correlated with behavioral anomalies).

Here, we work towards the goal of estimating population-level disease prevalence. Formulating the problem as an individual disease prediction task, we augment the existing individual-level features [45] with a social context to capture regular contacts and group interactions to better distill routine social interaction patterns. Central to our approach are graph neural networks (GNNs) [21, 28, 39] that have recently been adapted to model dynamic and temporal networks. Existing research into dynamic GNNs has predominantly been focused on modeling network formation and evolution in the context of link prediction [13, 35, 51], but such GNNs are not yet suited to tracking dynamic social behaviors of individuals and their routines. On the other hand, several traditional (non-GNN-based) dynamic and temporal network models have been designed to capture emergent patterns during network evolution, and to identify abnormal individuals or subgraphs [3, 34, 47]. Yet these approaches were also less suitable, since they are not designed to incorporate node features or be trained for specific tasks, such as disease prediction.

In this paper, we propose a novel integrated GNN for Dynamic Network Anomaly modeling (GraphDNA) to meet the goal of detecting deviations from an individual's routine social behaviors for predicting disease onset. Broadly, GraphDNA combines two key modules concerning dynamic social behavior prediction and anomaly-based disease prediction. The former module employs a graph convolutional neural network (GCN) model [28] to capture individuals' social behaviors and builds it into a long-short term memory (LSTM) model [23] to record the dynamic patterns of such behaviors. The latter module then combines a data-driven learnable logistic regression (LR) model [24] and a temporal-pattern-oriented statistical Gaussian tail probability (GTP) model [1] to predict disease diagnosis from anomalies in the social behavior dynamics.

In our experiments on the same labeled dataset from the H1N1v epidemic in Iceland [45], we evaluate GraphDNA by comparison to the most relevant baselines from state-of-the-art including dynamic GNNs and other temporal or networked anomaly detection models. With a focus on estimating the Date-of-Diagnosis (DoD) of diagnosed individuals, we demonstrate the advantages of our GraphDNA method on the generic task of supervised dynamic network anomaly detection. We also apply the individual inference of GraphDNA to the larger population, tracking the epidemic curve within the diagnosed population and, further, finding an illness-associated behavioral change signal in the whole population. Finally, we analyze key design decisions, hyper-parameter settings, and provide an efficiency study of GraphDNA.

2 BACKGROUND

2.1 Syndromic Surveillance

Keeping with technological developments and new data sources, syndromic surveillance systems emerged in the 2000s to seek to use existing health data in real time to provide immediate analysis and feedback to those charged with investigation and follow-up of potential outbreaks [22]. In 2009, Google Flu ushered in the era of big data syndromic surveillance through passively collected data sources by using aggregated search engine queries for flu-like symptoms to estimate regional influenza levels with a lag of only one day [19]. Google Flu's approach, however, was later found to have been flawed, missing non-seasonal influenza outbreaks and overestimating disease burden, and was shut down in 2015. Prominent researchers characterized the project's indifference to supplementing the existing body of science and instead seeking to replace it with black-box models as an example of big data hubris [31]. Research into other data sources for use in syndromic surveillance, such as social media, has followed [36, 38], built around technologies used primarily in high-income countries.

Aggregated CDR data, such as rates of population movement between cell-phone towers [29], have informed epidemiological models for cholera [6], dengue fever [49], malaria [8, 48], Ebola [30], influenza [44], and recently SARS-CoV-2 the pathogen that causes COVID-19 [12, 14]. Because these models lack linkage at the individual level, they rely on correlations between the aggregated data and other datasets, thereby limiting their statistical power and generality [17]. Individual CDR data were used during COVID-19 to infer likely contacts of infection in Israel, with staunch privacy objections [20] (cf. Appendix A).

2.2 Dynamic Network Anomaly Modeling

Anomaly detection refers to the data mining process that measures the deviations of objects of interest from the majority group [2, 10]. One of the most common scenarios of anomaly detection is on sequential data (*e.g.*, time-series), where the algorithm is often composed by a sequence modeling part and a deviation scoring part [11]. For instance, [1, 16, 25, 33, 55] employ sequential neural networks such as LSTM and HTM (hierarchical temporal memory) to model sequential records and then access the likelihood of anomalies based on the models' predictions. Recent studies for many emerging real-world applications concern the more complicated problem of anomaly detection on graph data [32]. For example, [3, 34, 47] detect abnormal nodes in graphs based on their deviations from normal node clusters without supervision. [47] combines one-class classi cation with GNNs for graph anomaly detection in a supervised manner, whereas [43] models node and edge features using time-series. However, these methods are designed only for graphs with xed structure.

Real-world networks can be modeled as dynamic graphs to represent evolving objects and relationships among them [32, 50]. Extensive research has been done into dynamic network modeling, including tasks such as temporal link prediction [13, 35, 51] and e cient graph streaming [5, 15, 18] none of which encompass anomaly detection. AddGraph [54] and NetWalk [53] are two methods that are closest to our setting of dynamic network anomaly modeling. AddGraph employs temporal GCN to detect anomalous edges but cannot trivially detect anomalous nodes, whereas NetWalk leverages a DeepWalk-based framework to detect both anomalous nodes and edges, but cannot readily incorporate node attributes or task-speci c supervision.

3 THE GRAPHDNA FRAMEWORK

3.1 Dataset Analysis

Description. The data set from Iceland contains CDR data for 93,409 people (about a quarter of the Icelandic population) over a 3-year period beginning in February 2009, with 87,773 individuals making calls during the 1-year period beginning in February 2009 when the H1N1v epidemic occurred. The CDR records are linked with in uenza-like illness (ILI) diagnosis data for 1,434 individuals who provide a spatially representative sample (A_i 0.86) of the homogeneous Icelandic population [45]; we focus only on an individual's rst ILI diagnosis. Each record contains the encrypted source and destination numbers for a call placed over a cell-phone tower, the GPS coordinates of the cell-phone tower, a timestamp, and the duration of call; similar metadata for text messages (SMS) are also included in the CDR data. No content of calls or text messages are included. The linked health dataset includes the encrypted number and the Date of Diagnosis (DoD) of ILI by healthcare providers in Iceland for the owner of that number.

The CDR data reveals rich movement and social patterns. Common contacts and their own interactions give a proxy for daily communication networks. The GPS location of a call gives a proxy for a person's location; a series of such locations provides a proxy for movement; and a series of movements can act as proxy for routine patterns, such as weekday commute to and from work. Existing studies identi ed that the movement patterns were di erent on the day before the DoD and up to three days after were signi cantly di erent from regular days, speci cally that 1.1 1.4 fewer unique tower locations were visited on average [45]. They also found that signi cantly fewer calls were placed but that calls were longer on the day following diagnosis. Prior work did not consider more advanced movement, social features, or dynamics.

Node features. We conducted principled analysis of the many node features that can be constructed from the CDR data, including location_num (number of unique tower recorded), avg_len (average call length), tot_len (total call length), call_cnt (call count), degree (number of contacts), clus_coe (cluster coe cient), abg_lon (average longitude), avg_lat (average latitude), all of which are varying by day. Intuitively, multiple features may indicate disease onset or diagnosis. We studied feature correlations based on the days from DoD to quantify such potential. Speci cally, since these predominantly ordinal attributes usually did not follow normal distributions, we measured feature correlations using the Spearman's Correlation Coe cient (SCC) [42].

Link features. To account for people's connections in the phone call network, we conduct the social behaviors of every individual, that includes their own behaviors (node features) together with those of their neighbors in the phone call network. For simplicity, we reduce social behaviors to features of a node together with the aggregate of the node features of its direct neighbors. In addition to binary indicators of whether two people were in contact during a day, the CDR data further allow us to extract various link features, such as call counts and (total) call durations. Before designing more complicated models beyond elementary GCN [28], we extend our data analysis over the correlations with days from DoD to study the potential impact of such link features in disease prediction.

Diagnostic features. Unlike during COVID-19, no large-scale control interventions (such as lock-downs or restaurant closures [12]) were imposed during the H1N1 epidemic in Iceland [41].

Figure 1 demonstrates the results of our node and link feature analysis based on their correlations with days from diagnosis based on the training data. Although the absolute correlation values are small, they are statistically signi cant with P-value 0.01, and are good indicators towards the utilities of these single features (as concluded from a similar analysis in [45]). Based on the correlation scores, we set an empirical threshold to select the top ve node features as a trade-o between model capacity and simplicity. For the link features, we found that unweighted links already encompass the strongest signal towards the DoD, obviating the need for more complicated GCN designs to model link features.

Combining nodes and links, in Figure 2, we visualize the dynamic social behaviors of individuals via three prominent node features aggregated through the (weighted/unweighted) links in the direct neighborhoods, where deviations are clearly observed around the DoD. Such observations motivate our goal of predicting symptomatic but unreported disease infections based on dynamic network anomalies in the CDR data.

Other features. While we rely on analyzing real data, both to identify the node and link features and to justify the design of our models, we underscore the greedy nature of such analysis and the potential over-simpli cation of the problem. However, the focus of our work is to provide the rst fundamental framework of symptomatic disease prediction based on dynamic network anomalies in CDR data, and believe that model simplicity is crucial.

3.2 Problem Formulation

Input. From CDR, we construct daily snapshots of the cell-phone call network as graphs $G = \{G^t = (V, E^t, \cdot^t)\}_{t=1}^{T}$. Here, V is the set of all vertices (individuals) who have at least one call record, E^t is the set of unweighted directed links at timestamp (day) t

Figure 1: Node and link feature analysis : Spearman's CC between social behaviors and days from diagnosis. We set an empirical threshold (dashed line) to choose relevant node features for inclusion. Unweighted links links without additional features were found to be the most useful.

Figure 2: Dynamic social behaviors of diagnosed people vs. days from DoD: We observe clear deviations of social behaviors around the DoD. The shaded interval marks the period between days -1 to +3 days from DoD when the largest deviations are observed.

i.e, $4^{1C^{\circ}}_{89}$ = 1 if there is at least one call from E8 to E9 on day C, and 0 otherwise, and $^{1C^{\circ}}$ denotes the behavioral features at timestamp C, i.e, $5^{1C^{\circ}}_{8}$ 2 R denotes the individual behavioral features of E8 on day C We model the complete year, between 02/01/2009 to 02/01/2010, to capture the entire 2009 H1N1v outbreak in Iceland, and use C2 f0•1• " " " •)= 364g to denote the relative days within that time frame.

Within + , we pay special attention to the subset $+^0$ + who had a record for an in uenza-like illness (ILI) diagnosis during the one year period. 2 R$^{j+^0 j \ j) \ j}$ stores the day of ILI diagnosis (DoD) labels of people in + ^{0} ($\sim^{1C^{\circ}}_{8}$ = 1 if E8 has a positive diagnosis on day C, and 0 otherwise). Recovery from in uenza may take several days and anomalous behavior is often observed in the several days surrounding the DoD [45]. We thus follow common practice [9] and prior work to de ne the extended DoD labels, where $\sim^{1C^{\circ}}_{8}$ = 1 if $\sim^{1C^{\circ}}_{8}$ = 1 and C2 ›C^0 1• C^0¸ 3¼ and 0 otherwise.

Output. The primary goal of our work is to predict the DoD of E8 2 + ^{0}, through modeling the connection between people's dynamic social behaviors and disease diagnoses based on the phone call graphs G given above. Beyond + ^{0}, the model should also generalize to the larger population + , where much of the diagnosis labels are unavailable, and yet provide disease prediction whether and when an individual gets infected and shows symptoms consistent with behavioral anomalies in the labeled input. Such estimates could be used to monitor the e ective disease burden of a population during an epidemic, as long as some data are available about how symptoms a ect behavior. Estimates could be further broken down by, e.g., age, region, sub-populations, as needed to inform policy and intervention strategy [45].

3.3 Model Overview

The main aim of our work is to model people's behaviors in G from CDR data, and measure deviations from routine to facilitate symptomatic surveillance. To meet this goal, we design a two-stage framework: (1) dynamic network behavior prediction, and (2) anomaly-based disease prediction, which can be further integrated through iterative training.

We survey our proposed GraphDNA framework in Figure 3. In the rst stage, a sequential graph representation learning module is designed to capture people's daily behaviors in phone call graphs G and then make consecutive predictions on their next-day behaviors. For people with diagnosis labels, only data on healthy days are used in this stage. In the second stage, an anomaly detection module is designed to compare the predicted behaviors with the true behaviors on each day and make predictions about whether a person might have fallen ill and show symptoms of H1N1v on that day.

We use a subset of people with diagnosis labels $+^1_{\text{train}}$ + ^{0} and the entire set of non-diagnosed people + + ^{0} to train the dynamic social behavior prediction module in stage one. We then use a disjoint subset of people with diagnosis labels $+^2_{\text{train}}$ + ^{0} to train the anomaly-based disease prediction module in stage two. Another disjoint set $+_{\text{val}}$ + ^{0} is used to iteratively validate and improve the model design as well as tune the model hyper-parameters, and the nal disjoint set $+_{\text{test}}$ + ^{0} is held out until the nal testing and reporting of the results.

3.4 Dynamic Social Behavior Prediction Module

To model people's routine behavior over time in cell-phone call graphs, we design a dynamic graph model to predict people's behaviors at each day (i.e., $5^{1C^{\circ}}_{8}$ •8E8 2 + • C2 f1•2• " " " •)g) based on their own past behaviors (i.e., f $5^{1C^{\circ}}_{8}$ j C^0 = 0• " " " • C 1g) and the past behaviors of their neighbors (i.e., f $5^{1C^{\circ}}_{9}$ j C^0 = 0• " " " • C 1; E9 2

Figure 3: An overview of our Dynamic Network Anomaly modeling (GraphDNA).

$\mathcal{N}(v_i, t, k)\}$, where $\mathcal{N}(v_i, t, k)$ denotes the k-hop neighborhood of v_i in graph $G^{(t)}$). To efficiently encode such dynamic social behaviors, we design an integrated model of GCN [28] and LSTM [37] that we train on the node set $\mathcal{V} = \mathcal{V}^1_{\text{train}} \cup \mathcal{V} \cup \mathcal{V}^0$.

Social behavior modeling. Motivated by recent advances in GCNs for node representation learning in content-rich networks [28], we employ GCN for modeling of static social behaviors of individuals based on the neighborhood of each node on each day in the cell-phone call graphs (i.e., $x_i^{(t)}, x_j^{(t)} \mid j \in \mathcal{N}(v_i, t, k), \forall v_i \in \mathcal{V}, t \in \{0, 1, \ldots, T\}$). We encode this information into representation vectors $h_i^{(t,L)}$ through recursive operations

$$H^{(t,l+1)} = \sigma\left(\hat{A}^{(t)} H^{(t,l)} W^{(l)} + b^{(l)}\right), \tag{1}$$

where $\hat{A}^{(t)}$ is the normalized adjacency matrix with self-loop on day t, $W^{(l)}$ and $b^{(l)}$ are the learnable parameters of the GCN model, σ is a non-linear activation function such as LeakyReLU, and $t \in \{1, 2, \ldots, T\}$. $H^{(t,0)} = X^{(t)}$ is the feature matrix on day t. Based on our data analysis in Section 3.1, we used the binary directed adjacency matrix $A^{(t)} \in \{0,1\}^{N \times N}$ and real-valued feature matrix $X^{(t)} \in \mathbb{R}^{N \times D}$ of selected node features. The number of GCN layers L (also denoted as L_1) is a tunable hyper-parameter. To capture a distillation of common patterns, we share and train the same GCN model across all nodes $v_i \in \mathcal{V}$ and all days $t \in \{0, 1, \ldots, T\}$.

Dynamic social behavior modeling. To integrate the history of past behaviors and model the dynamics of social behaviors, we further employ an LSTM model [37] based on the outputs of the GCN model. Specifically, given the sequence of representation vectors as the outputs of the GCN model (i.e., $\{h_i^{(t)} \mid t = 0, \ldots, T-1\}, \forall v_i \in \mathcal{V}$), the LSTM model seeks to predict the node features of the next days (i.e., $\{\hat{x}_i^{(t)} \mid t = 1, \ldots, T\}, \forall v_i \in \mathcal{V}$), which is computed through the standard recursive operations of LSTM following [37]. The number of LSTM layers L_2 is a tunable hyper-parameter. Given an input behavior representation of a node v_i on day t (i.e., $h_i^{(t)}$), the final output of the LSTM model is the predicted behavior (node feature) of v_i on day $t+1$ (i.e., $\hat{x}_i^{(t+1)}$).

To capture the common patterns, we share and train the same LSTM model across the representation and feature sequences of all nodes $v_i \in \mathcal{V}$, which we do in an end-to-end fashion jointly with the GCN model through the following objective function:

$$\min_{\Theta_1, \Theta_2} \sum_{v_i \in \mathcal{V}} \sum_{t=1}^{T} \mathcal{L}_1\left(x_i^{(t)}, \hat{x}_i^{(t)}\right), \tag{2}$$

where Θ_1 and Θ_2 denote the parameters of the GCN model and LSTM model, respectively. Here, $\mathcal{L}_1$ is a loss function such as MSE. We detail the training process in Algorithm 1.

3.5 Anomaly-based Disease Prediction Module

We focus on the task of DoD prediction not only because we only have positive labels of diagnosed people in the dataset but also due to the crucial impact of accurate detection of patient DoD on disease transmission control. Following past studies [45] and our data analysis in Section 3.1, our central hypothesis is that the DoD labels may be predicted to an extent based on people's deviations from their routine behaviors (i.e., anomalies) as captured in the cell-phone call graphs.

To detect anomalies, we first compute the deviation scores between the predicted behaviors and real behaviors for all people in another training set $\mathcal{V}^2_{\text{train}}$ that is disjoint from $\mathcal{V}$:

$$e_i^{(t)} = x_i^{(t)} - \hat{x}_i^{(t)}, \quad \forall v_i \in \mathcal{V}^2_{\text{train}}, t \in \{1, 2, \ldots, T\}. \tag{3}$$

Every individual $v_i \in \mathcal{V}^2_{\text{train}}$ is associated with a sequence of D-dimensional vectors $\{e_i^{(t)} \mid t = 1, 2, \ldots, T\}$, from which we will seek to predict the extended DoD labels $\{y_i^{(t)} \mid t = 1, 2, \ldots, T\}$.

We design and experiment with three representative types of anomaly detection models based on the output of our dynamic social behavior prediction stage: (1) a deep learning model based on logistic regression (LR) [24], (2) a statistical model based on Gaussian tail probabilities (GTP) [1], and (3) a hybrid model that integrates the first two.

Deep learning model. Since the DoD labels y are binary, we devise a LR model for binary classification [24]. To mitigate noise and asynchronous anomalies across different features, we smooth the input sequences over a rolling window. We have

$$\tilde{y}_i^{(t)} = \sigma\left(\mathrm{MLP}\left(\bar{e}_i^{(t)}\right)\right), \quad \forall v_i \in \mathcal{V}^2_{\text{train}}, t \in \{1, 2, \ldots, T\}, \tag{4}$$

where $\bar{e}_i^{(t)} = \mathrm{mean}\left(\ldots, e_i^{(t-1)}, e_i^{(t)}, e_i^{(t+1)}, \ldots\right)$. We pad both ends of the sequence with zeroes. Here, the window size w is a tunable hyper-parameter, σ is the sigmoid function, MLP is the multilayer perceptron with LeakyReLU activation, and the number of layers L_3 is another tunable hyper-parameter.

The LR model is trained with the following objective function

$$\min_{\Theta_3} \sum_{v_i \in \mathcal{V}^2_{\text{train}}} \sum_{t=1}^{T} \mathcal{L}_2\left(\tilde{y}_i^{(t)}, y_i^{(t)}\right), \tag{5}$$

where $\mathcal{L}_2$ is a loss function such as cross-entropy. To counter the propensity of LR to simply predict the majority class when the class labels are imbalanced, we employ a top-k selection mechanism during testing where we predict the top-k $\tilde{y}_i^{(t)}$s as 1 (illness) for each $v_i \in \mathcal{V}_{\text{val}} \cup \mathcal{V}_{\text{test}}$, and then set k to 5 since the largest interval of concern around the DoD is 5 days ([t-1, t+3]).

Statistical model. While LR provides an effective way of searching the feature space and finding the inductive bias with the help of training data, it ignores dynamic contexts and is not designed to capture temporal anomalies. On the other hand, anomaly detection has been explored in temporal settings through statistical models such as the Gaussian Tail Probability (GTP) model [1]. Following their design, to effectively detect temporal anomalies from the T-dimensional time-series data of the deviation scores of each individual v_i (i.e., $\{s_i^{(t)} \mid t = 1, 2, \ldots, T\}$), we first apply two rolling windows, W_1 and W_2 of sizes l_1 and l_2 as follows

$$W_1 = [\max(0, t - l_1 + 1), \max(0, t)], \quad l_1 \geq 1$$
$$W_2 = [\max(0, t - l_2 + 1), \max(0, t)], \quad l_2 \geq 1 \tag{6}$$

where $l_1 > l_2$ are two tunable hyper-parameters. We then model the values in W_1 as normal distributions, and use values in W_2 to compute the recent short-term average. An anomaly likelihood of $s_i^{(t)}$ based on the GTP is computed as

$$p_i^{(t)} = 1 - Q\left(\frac{\text{mean}(\{s_i^{(t)} \mid t \in W_2\}) - \text{mean}(\{s_i^{(t)} \mid t \in W_1\})}{\text{std}(\{s_i^{(t)} \mid t \in W_1\})}\right), \tag{7}$$

where Q represents the Gaussian tail probability approximation function [27]. The total anomaly probability of v_i on day t is computed as $p_i^{(t)} = \prod_{j=1}^{t} p_i^{(t,j)}$, which is directly used for the prediction of $\tilde{y}_i$ with the same top-K selection mechanism.

Hybrid model. The GTP model adds temporal context to the deviation scores and is thus more suitable for anomaly detection in the dynamic social behavior data. However, the multi-dimensional behavioral features are not parameterized for the task of symptom (DoD) prediction. To this end, we propose a novel hybrid model that combines the power of both worlds by simply replacing the $s_i^{(t)}$ in Eq. (4) with $p_i^{(t)}$ in Eq. (7). Sequence smoothing with w_0 is no longer needed due to the sliding windows W_1 and W_2.

Algorithm 1: Dynamic Social Behavior Prediction

Input: $\{\mathcal{G}^{(t)} \mid t = 0, \ldots, T\}$, $\mathcal{V} = \mathcal{V}^1_{train} \cup \mathcal{V}^0$, # GCN layers L_1, # LSTM layers L_2, hidden layer sizes

Output: $\{\hat{s}_i^{(t)}, \forall v_i \in \mathcal{V}, t \in \{1, 2, \ldots, T\}\}$

- while not converged do
  - for $t \in (0$ to $T)$ do
    - $h^{(t)} \leftarrow \text{GCN}(\mathcal{G}^{(t)}; \Theta_1)$
  - for $v_i \in \mathcal{V}$ do
    - for $t \in (0$ to $T)$ do
      - $\hat{s}_i^{(t+1)} \leftarrow \text{LSTM}(h^{(t)}; \Theta_2)$
      - loss $\leftarrow \mathcal{L}_1(\{s_i^{(t)}\}, \{\hat{s}_i^{(t)}\})$
      - Update the GCN and LSTM model parameters Θ_1 and Θ_2 according to the loss

Algorithm 2: Anomaly-based Disease Prediction (Hybrid)

Input: $\{\mathcal{G}^{(t)} \mid t = 0, \ldots, T\}$, $\{\hat{s}_i^{(t)} \mid v_i \in \mathcal{V}, t = 1, 2, \ldots, T\}$, $\mathcal{V}^2_{train}$, # LR layers L_3, hidden layer sizes, GTP window sizes l_1 and l_2

Output: $\{\hat{y}_i^{(t)}, \forall v_i \in \mathcal{V}, t \in \{1, 2, \ldots, T\}\}$

- while not converged do
  - for $v_i \in \mathcal{V}^2_{train}$ do
    - for $t \in (0$ to $T)$ do
      - $p_i^{(t)} \leftarrow \text{GTP}(\hat{s}_i^{(t)}; l_1, l_2)$
      - $\hat{y}_i \leftarrow \text{LR}(p_i^{(t)}; \Theta_3)$
      - loss $\leftarrow \mathcal{L}_2(\tilde{y}_i^{(t)}, \hat{y}_i^{(t)})$
      - Update the LR model parameters Θ_3 according to the loss

3.6 Training Algorithms

The detailed training algorithms of the two modules are outlined in Algorithms 1 and 2. We note that our GraphDNA framework does not rely on more hyper-parameters than the basic ones for classic GCN, LSTM, LR, and GTP models. In this work, we train the two stages separately and achieve promising results for symptom prediction in the end. Potentially, the two stages can also be trained jointly (iterative or end-to-end), which we leave as an interesting direction for future work.

Complexity analysis. The training of the GCN model in stage one takes $O(N_1^2 L_1)$ time; the training of the LSTM model takes $O(N_1 L_2)$ time in each epoch, where $N_1 = |\mathcal{V}| + |\mathcal{E}|$, $\mathcal{E} = |\mathcal{E}|$. In stage two, $O(N_2 T(l_1 + l_2))$ time is taken to calculate the GTP, and $O(N_2 L_3^2)$ time is taken to train the LR model, where $N_2 = |\mathcal{V}^2_{train}|$, $\mathcal{E} = |\mathcal{E}|$. $T, L_1, L_2, L_3, l_1, l_2$ are all constant numbers: T is 364, and all others are smaller than 100.

4 EXPERIMENTS

In this section, we evaluate GraphDNA by conducting extensive experiments on the CDR dataset, with a focus on the following research questions (RQs).

- RQ1 How does GraphDNA perform compared to closest baselines from state-of-the-art on DoD prediction?
- RQ2 Does GraphDNA have the potential to be generalized for disease prediction in the larger population?
- RQ3 How does each major component of GraphDNA contribute to the overall performance?
- RQ4 What are the effects of different tunable model hyper-parameters on GraphDNA?
- RQ5 Is the running time of GraphDNA comparable to existing methods?

4.1 Experimental Settings

Dataset. The Iceland CDR dataset has a total of 87,773 distinct nodes, and an average of 54,867 nodes and 30,451 links across the 365 graph snapshots. The nodes comprise two types: the 1,414 diagnosed nodes $\mathcal{V}^0$ and the remaining non-diagnosed nodes $\mathcal{V} \setminus \mathcal{V}^0$. There are DoD labels for diagnosed nodes, but we do not know if any individuals in the non-diagnosed set were infected or not. We divide the diagnosed nodes $\mathcal{V}^0$ into $\mathcal{V}^1_{train}$, $\mathcal{V}^2_{train}$, $\mathcal{V}_{val}$, and $\mathcal{V}_{test}$ as discussed in Section 3.3 with a ratio of 3:3:2:2. We use $\mathcal{V}^1_{train}$ and $\mathcal{V}^2_{train}$ to train the two stages of our model, respectively.

Table 1: Anomaly detection performance comparison. All results are averaged from 5 random data splits, passing signi cance test with ? = 0.01

Model	Metrics				
	Micro Precision	Micro Recall	Micro AUC	Micro F1	Macro Accuracy
NetWalk	0.0529 0.0019	0.1599 0.0028	0.5025 0.0005	0.0773 0.0004	0.1672 0.0007
LSTM-AD	0.0386 0.0035	0.2836 0.0047	0.4995 0.0003	0.0667 0.0003	0.3016 0.0014
OddBall	0.0362 0.0001	0.3530 0.0001	0.4988 0.0001	0.0648 0.0001	0.3578 0.0001
OCGNN	0.1754 0.0009	0.5491 0.0073	0.5043 0.0033	0.2586 0.0043	0.5749 0.0046
GraphDNA-w/o-GCN	0.0490 0.0018	0.1356 0.0006	0.0694 0.0005	0.0693 0.0006	0.1441 0.0013
GraphDNA-w/o-LSTM	0.2326 0.0063	0.6792 0.0034	0.5855 0.0040	0.3333 0.0017	0.6871 0.0061
GraphDNA-w/o-LR	0.2138 0.0036	0.4652 0.0063	0.5728 0.0037	0.2807 0.0012	0.4723 0.0029
GraphDNA-w/o-GTP	0.0871 0.0005	0.2356 0.0016	0.5167 0.0022	0.1222 0.0015	0.2372 0.0018
GraphDNA	0.2344 0.0106	0.6986 0.0054	0.5895 0.0019	0.3384 0.0019	0.7005 0.0087

Baselines. We adapted the following state-of-the-art algorithms for our task of DoD prediction based on the dynamic cell-phone call graphs constructed from the CDR dataset.

- NetWalk[53]: an anomalous node detection method that is closest to our dynamic network setting. It learns and dynamically updates the representations of non-attributed networks as they evolve in an unsupervised manner.
- LSTM-AD[33]: an algorithm using stacked LSTM networks for anomaly detection in multi-variate time-series data. Since it cannot model network data, we provide it only with dynamic node features.
- OddBall[3]: an unsupervised method to detect abnormal nodes in static networks. Since it cannot handle dynamic networks, we compute a separate model of it for every timestamp.
- OCGNN[47]: a one-class classi cation framework that combines GNN with the one-class objective for attributed network anomaly detection in a supervised manner. Since it cannot handle dynamic networks, we compute a separate model for every timestamp.

For the supervised baselines, the same $+^1_{train}$ [$+^2_{train}$ [$+$ $+^0$ is used for training, $+_{val}$ is used for hyper-parameter tuning, and $+_{test}$ is used for performance reporting. The unsupervised baselines are run on the whole $+$ and tested on $+_{test}$. When making predictions on $+_{test}$, the same top- selection mechanism is used to predict positive DoDs for each individual.

Evaluation metrics. Based on the predicted DoD labels $\hat{}$ and extended true DoD labels $\bar{}$, we compute the following metrics adopted from the standard evaluation of group classi cations.

- Micro Precision, Micro Recall, Micro AUC, and Micro F1, which represent the Precision, Recall, AUC and F1 scores averaged across all the testing individuals in $+_{test}$.
- Macro Accuracy, which is the percentage of testing individuals in $+_{test}$ who have at least one correct DoD prediction.

The suite of metrics compares prediction results with ground-truth from di erent perspectives, thus comprehensively comparing the performance of evaluated algorithms.

Parameter settings. We tune and set the hyper-parameters of GraphDNA as the following default values: we set the number of GCN layers $!_1$ to 2, LSTM layers $!_2$ to 1, and LR layers $!_3$ to 2; we set the embedding size of all layers in all models to 16; the sizes of rolling windows in GTP are set to $_1 = 100$ and $_2 = 3$. To ensure fair comparison, we use the same hyper-parameters for all of our model ablations. For the baselines, we also optimize their hyper-parameters on $+_{val}$.

Figure 4: Average disease scores (ADS) of diagnosed group and whole population vs. daily diagnosed number (DDN) in the period of 2009 H1N1v outbreak in Iceland. Thin lines denote the medians/values of the ADS/DDN, thick lines indicate the smoothed medians/values, and shading delineates the 1st 3rd quantiles of the ADS.

4.2 DoD Prediction Comparison (RQ1)

Table 1 shows that GraphDNA achieves the best performance across all metrics in the scenario of CDR-based DoD prediction. We highlight the following detailed observations.

- While not being fully consistent across the baselines, the multiple metrics we use demonstrate the same signi cant improvements of GraphDNA. Speci cally, GraphDNA achieves 16.9%-33.6% relative gains over the strongest baseline across all metrics, indicating its superiority in the task of CDR-based DoD prediction.
- Although we have included the most relevant algorithms as baselines, none of them can properly integrate all important signals in our scenario, thus leading to unsatisfactory results across all metrics.
- Compared with LSTM-AD and OddBall, NetWalk focuses on structural anomalies and make cautious predictions, thus achieving better precision but worse recall.
- OCGNN is the strongest baseline, likely due to its proper leverage of imbalanced task supervision, which indicates the importance of available DoD labels from the CDR data.

(a) Hidden Layer Sizes (b) # LSTM Layers $!_2$ (c) # LR Layers $!_3$

Figure 5: Performance of GraphDNA with varying hyper-parameters (averaged from 5 random data splits). The best baseline here is OCGNN.

4.3 Anomaly Curve during Epidemic (RQ2)

Beyond predicting the DoD of diagnosed people, we examine the potential of GraphDNA to estimate wider disease infection among the entire population. In Figure 4, we visualize the average disease score (ADS) in the diagnosed group (c_{test}) and the whole population (+) predicted by GraphDNA, versus the diagnosed number (DDN) in the ground-truth of $+^0$. The main peak of the GraphDNA ADS estimate among the diagnosed group coincides with the ground-truth peak of H1N1v outbreak in Iceland in October 2009, suggesting that the ADS model captures behavioral anomalies associated with illness. The model also picks up anomalies during the winter holidays in December 2009. Interestingly, a small but signi cant anomaly signal also arises in the whole population during the epidemic (green curve). Notably, the model was not picking up time-of-year related artefacts, as evidenced by the baseline (orange curve) showing ADS inference by the same model trained on a control group in which we matched an undiagnosed person with each diagnosed person 1:1 at random and assigned them the latter's DoD. This supports the conclusion that the model is identifying illness-speci c anomalies in the whole population a promising information source. We caution, however, that further research is warranted for predictive epidemic estimation since the ADS scores in our model are based on training data from the entire 1-year period.

4.4 In-depth Model Analysis (RQ3-5)

Ablation analysis (RQ3). Table 1 also shows that each constituent part of GraphDNA contributes signi cantly to its overall performance. We further summarize several key observations as follows.

- Removing the GCN model causes the most signi cant performance drop, demonstrating the importance of modeling the neighborhood behaviors for DoD prediction the key di erence from our work to previous studies on the same CDR datasets [45].
- Surprisingly, removing the LSTM model actually does not signi cantly degrade performance consistent with the reasonable performance of OCGNN. Perhaps evolutionary patterns are not be a key factor for DoD prediction; perhaps LSTM is not the ideal model to capture such network evolution.
- Both the LR and GTP models are indispensable to GraphDNA, supporting our design principle of integrating the e ective data-driven learning ability of LR with the anomaly-based feature engineering of GTP.

In summary, the ablation test justi es the e ciency of our model design. Each of the main components contributes to the accuracy and robustness of GraphDNA.

Hyper-parameter analysis (RQ4). Comprehensive experiments are done for hyper-parameter tuning, and the results are presented in Figure 5. We run GraphDNA with various combinations of hyper-parameters and plot the performances holding each hyper-parameter to be xed. We highlight three important observations:

- The hyper-parameters we tested have minimal impact on the performance of GraphDNA, maintaining signi cant margins from the best baseline across a vast range of values.
- Larger embedding sizes, fewer LSTM layers, and fewer LR layers generally improve results due to di erent trade-o s between model capacity and over tting.
- The standard deviations remain acceptable across di erent settings, indicating that GraphDNA's hyper-parameter are robust.

Due to the di culty in implementing and running deep GCNs, we have not studied the performance of GraphDNA with the number of GCN layers $!_1$ greater than 2. While having signi cantly larger training and testing times, we have observed the performance of GCN with $!_1 = 2$ to be only slightly better than that with $!_1 = 1$, and thus lack compelling need to grow $!_1$ beyond 2 at the moment.

E ciency analysis (RQ5). We observe the computational cost of GraphDNA to be similar to those of OCGNN, which is slightly larger than those of LSTM-AD and NetWalk, yet within the same order of magnitude (detailed results and analysis in Appendix B).

5 CONCLUSION

Disease outbreak detection is di cult: population surveys are slow and skewed, and traditional syndromic surveillance requires the integration of a health-care data collection system with a responsive public health body to function adequately. Detecting behavioral anomalies through cell-phone metadata, as discussed here, o ers a passive and universal alternative to infectious disease surveillance. Using real-world linked cell-phone and health data from the H1N1v pandemic in Iceland in 2009, we showed how GraphDNA identi ed individual behavior change indicative of disease symptoms and found evidence of illness-related anomalies in the entire population that could be used to track the prevalence of symptoms. These estimates could inform transmission models, policy choices (e.g., targeted lockdowns, quarantines, vaccination campaigns) and provide direct observation of societal costs.

6 ACKNOWLEDGEMENTS

We thank the anonymous reviewers of our manuscript for constructive feedback, and past members and a liates of Emory SimBioSys lab for early work on the problem. This research is partly supported by equipment and internal grants from Emory University, a hardware donation from NVIDIA Corporation, Icelandic Centre for Research Award 152620-051, UKRI (MRC and EPSRC) through grants MC/PC/19067, MR/V038613/1, EP/V051555/1, EP/N510129/1 and investigator-led grants from Pfizer.

REFERENCES

[1] Subutai Ahmad and Scott Purdy. 2016. Real-time anomaly detection for streaming analytics. arXiv preprint arXiv:1607.02480 (2016).
[2] Mohiuddin Ahmed, Abdun Naser Mahmood, and Jiankun Hu. 2016. A survey of network anomaly detection techniques. Journal of Network and Computer Applications 60 (2016), 19 31.
[3] Leman Akoglu, Mary McGlohon, and Christos Faloutsos. 2010. Oddball: Spotting anomalies in weighted graphs. In PAKDD.
[4] Samuel Altmann, Luke Milsom, Hannah Zillessen, Raffaele Blasone, Frederic Gerdon, Ruben Bach, Frauke Kreuter, Daniele Nosenzo, Séverine Toussaert, Johannes Abeler, et al. 2020. Acceptability of app-based contact tracing for COVID-19: Cross-country survey study. JMIR mHealth and uHealth 8, 8 (2020), e19857.
[5] Khaled Ammar. 2016. Techniques and systems for large dynamic graphs. In SIGMOD.
[6] Linus Bengtsson, Jean Gaudart, Xin Lu, Sandra Moore, Erik Wetter, Kankoe Sallah, Stanislas Rebaudet, and Renaud Piarroux. 2015. Using mobile phone data to predict the spatial spread of cholera. Sci. Rep. 5 (2015), 8923.
[7] Nita Bharti. 2021. Linking human behaviors and infectious diseases. PNAS 118, 11 (2021).
[8] Caroline O Buckee, Amy Wesolowski, Nathan N Eagle, Elsa Hansen, and Robert W Snow. 2013. Mobile phones and malaria: modeling human and parasite travel. Travel Med Infect Dis 11, 1 (2013), 15 22.
[9] Edenilson E Calore, David E Uip, and Nilda M Perez. 2011. Pathology of the swine-origin influenza A (H1N1) flu. JPRP 207, 2 (2011), 86 90.
[10] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. ACM computing surveys (CSUR) 41, 3 (2009), 1 58.
[11] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2010. Anomaly detection for discrete sequences: A survey. TKDE 24, 5 (2010), 823 839.
[12] Serina Chang, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2021. Mobility network models of COVID-19 explain inequities and inform reopening. Nature 589, 7840 (2021), 82 87.
[13] Jinyin Chen, Xuanheng Xu, Yangyang Wu, and Haibin Zheng. 2018. Gc-lstm: Graph convolution embedded lstm for dynamic link prediction. arXiv preprint arXiv:1812.04206 (2018).
[14] Forrest W Crawford, Sydney A Jones, Matthew Cartter, et al. 2022. Impact of close interpersonal contact on COVID-19 incidence: Evidence from 1 year of mobile device data. Science advances 8, 1 (2022).
[15] Laxman Dhulipala, Guy E Blelloch, and Julian Shun. 2019. Low-latency graph streaming using compressed purely-functional trees. In PLDI.
[16] Tolga Ergen and Suleyman Serdar Kozat. 2019. Unsupervised anomaly detection with LSTM neural networks. TNNLS 31, 8 (2019), 3127 3141.
[17] Susan L. Erikson. 2018. Cell Phones ≠ Self and Other Problems with Big Data Detection and Containment during Epidemics. MAQ 32, 3 (2018), 315 339.
[18] Guoyao Feng, Xiao Meng, and Khaled Ammar. 2015. Distinger: A distributed graph data structure for massive dynamic graph processing. In IEEE BigData.
[19] Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant. 2009. Detecting influenza epidemics using search engine query data. Nature 457, 7232 (2009), 1012 1014.
[20] David M. Halbfinger, Isabel Kershner, and Ronen Bergman. 2020. To Track Coronavirus, Israel Moves to Tap Secret Trove of Cellphone Data. The New York Times (16 March 2020). Issue 2020-03-16.
[21] William L Hamilton, Rex Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In NeurIPS.
[22] Kelly J Henning. 2004. What is syndromic surveillance? MMWR (2004), 7 11.
[23] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. Neural computation 9, 8 (1997), 1735 1780.
[24] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. 2013. Applied logistic regression. Vol. 398.
[25] Wenjie Hu, Yang Yang, Ziqiang Cheng, Carl Yang, and Xiang Ren. 2021. Time-Series Event Prediction with Evolutionary State Graph. In WSDM.
[26] International Telecommunication Union. 2020. Measuring digital development: Facts and figures 2020.
[27] George K Karagiannidis and Athanasios S Lioumpas. 2007. An improved approximation for the Gaussian Q-function. IEEE COMML 11, 8 (2007), 644 646.
[28] Thomas N Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In ICLR.
[29] Nishant Kishore, Mathew V Kiang, Kenth Engø-Monsen, Navin Vembar, Andrew Schroeder, Satchit Balsari, and Caroline O Buckee. 2020. Measuring mobility to monitor travel and physical distancing interventions: a common framework for mobile phone data analysis. The Lancet Digital Health (2020).
[30] Andrew M Kramer, J Tomlin Pulliam, Laura W Alexander, Andrew W Park, Pejman Rohani, and John M Drake. 2016. Spatial spread of the West Africa Ebola epidemic. Royal Society Open Science 3, 8 (2016), 160294.
[31] David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. The parable of Google Flu: traps in big data analysis. Science 343, 6176 (2014), 1203 1205.
[32] Xiaoxiao Ma, Jia Wu, Shan Xue, Jian Yang, Quan Z Sheng, and Hui Xiong. 2021. A Comprehensive Survey on Graph Anomaly Detection with Deep Learning. arXiv preprint arXiv:2106.07178 (2021).
[33] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, and Puneet Agarwal. 2015. Long short term memory networks for anomaly detection in time series. In ESANN.
[34] Emmanuel Müller, Patricia Iglesias Sánchez, Yvonne Mülle, and Klemens Böhm. 2013. Ranking outlier nodes in subspaces of attributed graphs. In ICDEW.
[35] Aldo Pareja, Giacomo Domeniconi, Jie Chen, Tengfei Ma, Toyotaro Suzumura, Hiroki Kanezashi, Tim Kaler, Tao Schardl, and Charles Leiserson. 2020. Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In AAAI.
[36] Adam Sadilek, Stephanie Caty, Lauren DiPrete, Raed Mansour, Tom Schenk, Mark Bergtholdt, Ashish Jha, Prem Ramaswami, and Evgeniy Gabrilovich. 2018. Machine-learned epidemiology: real-time detection of foodborne illness at scale. NPJ digital medicine 1, 1 (2018), 1 7.
[37] Haşim Sak, Andrew Senior, and Françoise Beaufays. 2014. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling. In INTERSPEECH.
[38] Loukas Samaras, Elena García-Barriocanal, and Miguel-Angel Sicilia. 2020. Syndromic surveillance using web data: a systematic review. Innov Health Inform (2020), 39 77.
[39] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2008. The graph neural network model. TNNLS 20, 1 (2008), 61 80.
[40] Tamar Sharon. 2020. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. Ethics and Information Technology (2020), 1 13.
[41] G Sigmundsdottir, T Gudnason, Ö Ólafsson, GE Baldvinsdottir, A Atladottir, A Löve, L Danon, and H Briem. 2010. Surveillance of influenza in Iceland during the 2009 pandemic. Euro Surveill 15, 49 (2010), 19742.
[42] Charles Spearman. 1987. The proof and measurement of association between two things. Am. J. Psychol. 100, 3/4 (1987), 441 471.
[43] Xian Teng, Yu-Ru Lin, and Xidao Wen. 2017. Anomaly detection in dynamic networks using multi-view time-series hypersphere learning. In CIKM.
[44] Michele Tizzoni, Paolo Bajardi, Adeline Decuyper, Guillaume Kon Kam King, Christian M Schneider, Vincent Blondel, Zbigniew Smoreda, Marta C González, and Vittoria Colizza. 2014. On the use of human mobility proxies for modeling epidemics. PLoS Comput Biol 10, 7 (2014), e1003716.
[45] Ymir Vigfusson, Thorgeir A Karlsson, Derek Onken, Congzheng Song, Atli F Einarsson, Nishant Kishore, Rebecca M Mitchell, Ellen Brooks-Pollock, Gudrun Sigmundsdottir, et al. 2021. Cell-phone traces reveal infection-associated behavioral change. PNAS 118, 6 (2021).
[46] Dashun Wang, Dino Pedreschi, Chaoming Song, Fosca Giannotti, and Albert-Laszlo Barabasi. 2011. Human mobility, social ties, and link prediction. In KDD.
[47] Xuhong Wang, Baihong Jin, Ying Du, Ping Cui, Yingshui Tan, and Yupu Yang. 2021. One-class graph neural networks for anomaly detection in attributed networks. Neural. Comput. Appl. (2021), 1 13.
[48] Amy Wesolowski, Nathan Eagle, Andrew J Tatem, David L Smith, Abdisalan M Noor, Robert W Snow, and Caroline O Buckee. 2012. Quantifying the impact of human mobility on malaria. Science 338, 6104 (2012), 267 270.
[49] Amy Wesolowski, Taimur Qureshi, Maciej F Boni, Pål Roe Sundsøy, Michael A Johansson, Syed Basit Rasheed, Kenth Engø-Monsen, and Caroline O Buckee. 2015. Impact of human mobility on the emergence of dengue epidemics in Pakistan. PNAS 112, 38 (2015), 11887 11892.
[50] Yuanzhen Xie, Zijing Ou, Liang Chen, Yang Liu, Kun Xu, Carl Yang, and Zibin Zheng. 2021. Learning and Updating Node Embedding on Dynamic Heterogeneous Information Network. In WSDM.
[51] Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, and Kannan Achan. 2020. Inductive representation learning on temporal graphs. arXiv preprint arXiv:2002.07962 (2020).
[52] Takahiro Yabe, Nicholas KW Jones, P Suresh C Rao, Marta C Gonzalez, and Satish V Ukkusuri. 2022. Mobile phone location data for disasters: A review from natural hazards and epidemics. Computers, Environment and Urban Systems 94 (2022), 101777.
[53] Wenchao Yu, Wei Cheng, Charu C Aggarwal, Kai Zhang, Haifeng Chen, and Wei Wang. 2018. Netwalk: A flexible deep embedding approach for anomaly detection in dynamic networks. In KDD.
[54] Li Zheng, Zhenpeng Li, Jian Li, Zhao Li, and Jun Gao. 2019. AddGraph: Anomaly Detection in Dynamic Graph Using Attention-based Temporal GCN. In IJCAI.
[55] Panpan Zheng, Shuhan Yuan, Xintao Wu, Jun Li, and Aidong Lu. 2019. One-class adversarial nets for fraud detection. In AAAI, Vol. 33. 1286 1293.