High-Confidence Computing xxx (xxxx) xxx

Contents lists available at ScienceDirect

High-Confidence Computing

journal homepage: www.elsevier.com/locate/hcc

WLAN-log-based superspreader detection in the COVID-19 pandemic

Cheng Zhang[a,*], Yunze Pan[a], Yunqi Zhang[a], Adam C. Champion[a], Zhaohui Shen[c], Dong Xuan[a], Zhiqiang Lin[a], Ness B. Shroff[a,b]

[a] *Department of Computer Science and Engineering, The Ohio State University, USA*
[b] *Department of Electrical and Computer Engineering, The Ohio State University, USA*
[c] *VirtualKare LLC, USA*

ARTICLE INFO

Keywords:
Superspreader detection
WLAN logs
Contact tracing
Network analysis
COVID-19 pandemic

ABSTRACT

Identifying “superspreaders” of disease is a pressing concern for society during pandemics such as COVID-19. Superspreaders represent a group of people who have much more social contacts than others. The widespread deployment of WLAN infrastructure enables non-invasive contact tracing via people’s ubiquitous mobile devices. This technology offers promise for detecting superspreaders. In this paper, we propose a general framework for WLAN-log-based superspreader detection. In our framework, we first use WLAN logs to construct contact graphs by jointly considering human symmetric and asymmetric interactions. Next, we adopt three vertex centrality measurements over the contact graphs to generate three groups of superspreader candidates. Finally, we leverage SEIR simulation to determine groups of superspreaders among these candidates, who are the most critical individuals for the spread of disease based on the simulation results. We have implemented our framework and evaluate it over a WLAN dataset with 41 million log entries from a large-scale university. Our evaluation shows superspreaders exist on university campuses. They change over the first few weeks of a semester, but stabilize throughout the rest of the term. The data also demonstrate that both symmetric and asymmetric contact tracing can discover superspreaders, but the latter performs better with daily contact graphs. Further, the evaluation shows no consistent differences among three vertex centrality measures for long-term (i.e., weekly) contact graphs, which necessitates the inclusion of SEIR simulation in our framework. We believe our proposed framework and these results can provide timely guidance for public health administrators regarding effective testing, intervention, and vaccination policies.

1. Introduction

The COVID-19 pandemic has devastated many communities worldwide. The presence of the novel coronavirus (that causes COVID-19) in a community with high population density, such as a large public university, significantly increases the risk of contracting the disease. To fight COVID-19, contact tracing [1–5] is especially important to discover active individuals, known as *superspreaders*,[1] who lead to numerous COVID-19 transmission cases. Tracing human contacts to understand superspreader events is vital for preventing the spread of disease in communities such as university campuses, and such tracing has thus attracted a flurry of research interest [7–10].

Typically, contact tracing is conducted manually [11] (e.g., through questionnaires and interviews), initially collecting necessary information from infected patients (such as locations they visited and people with whom they had contact). Unfortunately, manual contact tracing can result in inaccurate results due to people’s unreliable memories and long delays. Hence, to fight the COVID-19 pandemic, researchers have developed numerous (partially) automated contact tracing systems. Recent efforts can be divided into two categories: *client-based* and *infrastructure-based*. Client-based approaches require pervasive deployment of apps on people’s mobile devices. Client-side apps leverage a wide variety of sources to track “encounters,” including records of credit card transactions [12], cryptographic tokens exchanged via Bluetooth Low Energy (BLE) [13–16], or acoustic channels [10]. In contrast, infrastructure-based methods exploit existing infrastructure deployed worldwide, such as CCTV footage [17], locations measured using cellular networks [18], Wi-Fi hotspots [9], and GPS [19], without requiring client-side involvement. *In this context, our paper presents an approach*

[*] Corresponding author.
E-mail address: zhang.7804@osu.edu (C. Zhang).

[1] There is no scientific definition of a “superspreader”. We use a definition similar to that in [6]: superspreaders are people with far more social connections than others, are more likely to be infected, and, if infected, will infect many more people than the median.

https://doi.org/10.1016/j.hcc.2021.100005
Received 6 February 2021; Accepted 27 February 2021
Available online xxx

Please cite this article as: C. Zhang, Y. Pan, Y. Zhang et al., WLAN-log-based superspreader detection in the COVID-19 pandemic, High-Confidence Computing, https://doi.org/10.1016/j.hcc.2021.100005

leveraging Wi-Fi local area network (WLAN) logs to identify potential superspreaders on the campus of a large public university.

However, leveraging WLAN logs for superspreader detection is nontrivial, with two major issues. First, conventional WLAN-based solutions (e.g., WiFiTrace [9]) infer whether students have contacted with each other based on their associations with specific access points (APs) during certain time intervals (e.g., > 15 min). Such *symmetric* contact detection neglects an important fact: the virus carried by people who have tested positive may infect others and replicate via pathogens in the environment. Therefore, others may be infected even if they linger in the environment over very short periods of time (e.g., < 15 min). Obviously, the current definition of human contact cannot handle this scenario. Second, existing Wi-Fi-based methods [9] quantify a superspreader by the number of associated devices from the same access point. However, the number of contacts may be unable to truly reflect how critical an individual is to spreading disease amidst the population. For example, previous work on vertex centrality measurement for social network analysis [20] demonstrates that the "importance" of a specific vertex in message-passing not only depends on the number of connected vertices, but is relevant to the vertex's location in social networks. Moreover, ground truth remains unknown in WLAN-based contact graphs, making it hard to understand how fast the disease propagates and progresses in order to determine superspreaders.

To tackle the first issue, we introduce *asymmetric contact*, a new type of human contact. Two persons in asymmetric contact are not necessarily associated with specific APs for the same period. For example, assume Persons A and B are in asymmetric contact. Person A may stay with one specific AP for a short time (e.g., 5 min) whereas Person B stays longer (50 min). Due to Person B's longer stay time, he generates a much stronger "environment" with his microbes than Person A does. If B tests positive, he may infect A even if the latter's stay time is only 5 min. On the other hand, A will not infect B due to her short stay. Hence the contact between these two persons is asymmetric. When we count the contact number, B's contact with A is counted, but A's contact with B is not. The concept of asymmetric contact partially captures the notion of *environmental infection*[2] [21].

As to the second issue, ideally, we can choose a vertex measure to determine superspreaders using either analytical solutions or prior experimental tests. Unfortunately, due to the diversity of contact graphs and the complexity of virus propagation, it is very difficult (if not impossible) to do so. In this paper, we propose an empirical approach. We include SEIR simulation, a necessary component in our solution, to ultimately determine superspreaders among the vertex-measure outputs. Specifically, we use the SEIR model to simulate the spread of the virus, followed by adaptive interventions on groups of superspreaders identified via different measures. We then finalize superspreaders who have the most crucial virus spread impacts, over the given contact graph, according to the simulations.

Incorporating the above two ideas, we propose a general framework for WLAN-log-based superspreader detection, which includes three key steps. First, we extract the individual's trajectory from wireless local area network (WLAN) logs to construct contact graphs, where vertices correspond to individual students and edges indicate physical contacts. In particular, we include both symmetric and asymmetric contact tracing to reveal potential directional interactions. Second, we adopt three vertex centrality measurements to identify three groups of potential superspreaders given a contact graph. Finally, we leverage the SEIR model to compare different vertex centrality measures and determine superspreaders based on simulation results.

The WLAN dataset [24] used in this paper contains over 5,000 students at a large public university, which represents a random sample of the overall student body. Over 8,000 APs are deployed among more than 200 buildings on campus, including lecture halls, dormitories, and restaurants. There are over 41 million device (dis)associations with WLAN APs at the university over a 139-day observation period in 2015. Although insufficient WLAN logs are available in 2020 due to school closures originating from COVID-19, the 2015 logs describe campus interactions before. Fig. 1 shows the locations of APs in multiple buildings on campus. Since the whole campus spans over 1,500 acres, our framework offers potential assisting superspreader detection efforts in large communities.

The main findings of this work include the following: (1) We find that there is a group of students that is critical in spreading the virus throughout the university's social contact networks. (2) We show the importance of symmetric and asymmetric contact tracing in superspreader detection. Specifically, we show that asymmetric contact tracing helps to discover hidden superspreaders in daily contact graphs and proper interventions with identified superspreaders greatly boosts efforts to contain the spread of disease. (3) We find that simple *betweenness* centrality better reveals the most critical individuals in daily contact graphs. We do not observe notable differences between vertex centrality measures in longer-term (*i.e.,* weekly) contact graphs with epidemic control. (4) For resource-constrained quarantine, we observe that increasing the percentage of the quarantined individuals to over 20% of the population yields limited extra benefits. (5) We find that superspreaders change heavily over the first few weeks, then remain stable during the rest of the semester. The similarity of superspreaders between the first 20 weeks and 15 weeks is around 0.8 using the rank-biased overlap metric [25], opening up opportunities to discover superspreaders as early as possible for efficient pandemic mitigation.

Practical significance for university/city administrators: We believe our proposed contact tracing method will enable both proactive and reactive interventions. For the former, our method can help administrators rapidly identify superspreaders for health warnings and frequent testing, using data from just the first few weeks of the semester. For the latter, our method can assist efforts in contact tracing, quarantine, medical support, and prioritized patient care.

In summary, our main contributions are threefold:

- We propose a general framework for WLAN-log-based contact analysis and superspreader detection. The framework applies to a wide range of working scenarios based on users' preferences, environmental dynamics, and resource availability.
- We present a set of initial work using the WLAN-log-based superspreader detection framework, including asymmetric contact tracing, vertex centrality measurement, and simulation-based superspreader determination.
- We implement the framework and evaluate it on a large-scale real-world WLAN log dataset. Our empirical results show the efficacy of the proposed contact tracing approaches and uncover insightful findings for public health administrators.

The rest of this paper is organized as follows. Section 2 provides background on epidemic models. Section 3 presents our framework on WLAN-log-based superspreader detection. Section 4 illustrates our evaluation results and analyses. Section 5 reviews related work. Finally, Section 6 concludes the paper.

2. Background: compartmental epidemic models

In this section, we discuss the background of compartmental epidemic models, which are simplified mathematical models of infectious diseases [26–30]. Recently, the SEIR (**S**usceptible, **E**xposed, **I**nfectious, **R**ecovered) model has shown promise combating COVID-19 in disease modeling [6,31], forecasting [32,33], and intervention [34]. In the SEIR model, the population is assigned to labeled compartments between which people move based on their health status.

Following the equivalent compartmental diagram shown in Fig. 2, we can use the following differential equations to describe the SEIR

[2] The physical environment represents an important source of pathogens that can cause infections or carry antibiotic resistance.

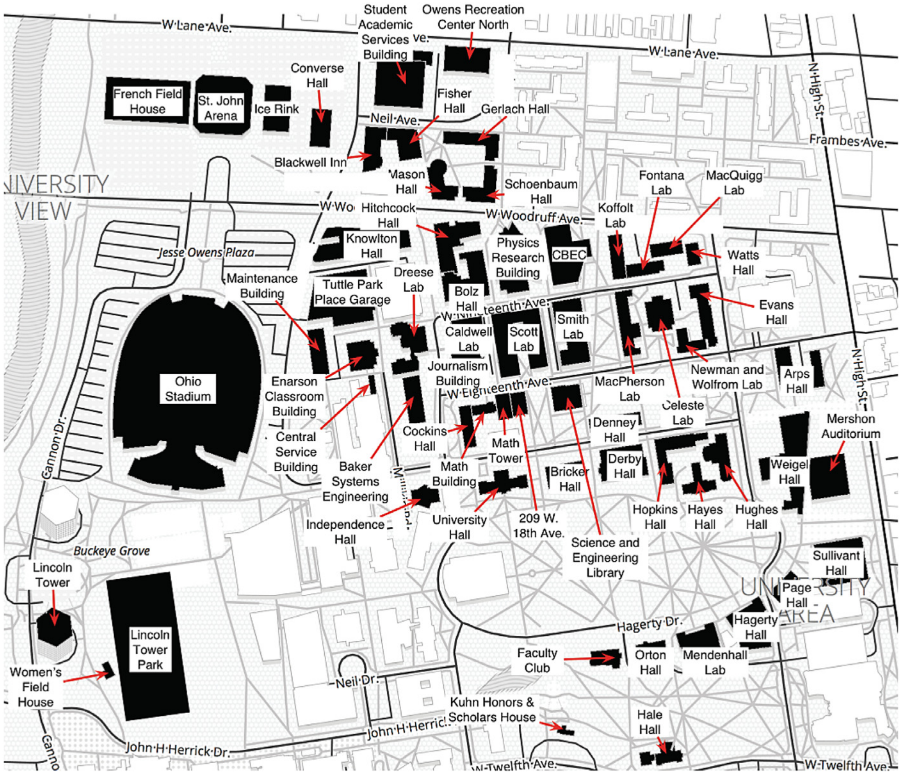


Fig. 1. Campus buildings with AP deployment information (shaded). Other buildings include: 22 E. 16th Avenue, 53 W. 11th Avenue, Knight House, North Commons, Northwood-High Building, Raney Commons, Riverwatch Tower, and the Wexner Center for the Arts (not shown). We generate the map using Mapzen [22] with OpenStreetMap data [23].

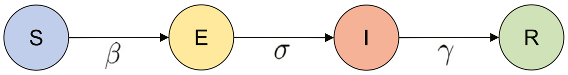


Fig. 2. Illustration of the popular SEIR compartmental model in epidemiology. The population is assigned to one of several labeled compartments: Susceptible, Exposed, Infectious, or Recovered. The order of the labels usually shows flow patterns between the compartments with epidemiological parameters β, σ, and γ. Details are explained in the text.

model involving variables S, E, I, and R and their rates of change with respect to time t:

$$\begin{aligned} \frac{dS}{dt} &= -\beta\frac{IS}{N}, \\ \frac{dE}{dt} &= \beta\frac{IS}{N} - \sigma E, \\ \frac{dI}{dt} &= \sigma E - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{1}$$

where β is the probability of transmitting disease between a susceptible and an infectious individual, σ is the inverse of the average incubation time (the rate of latent individuals becoming infectious), and γ is the recovery rate. In this model, recovered individuals are permanently immune to disease. In practice, all parameters are constant values that can be obtained via maximum likelihood estimation with real pandemic data. In this work, we leverage SEIR simulations [6] to model quarantine (self-isolation) of identified superspreaders (cf. Section 4.1).

3. Methodology

Fig. 3 shows an overall framework of our WLAN-log-based superspreader detection. We describe each component as follows.

3.1. WLAN data collection

The WLAN logs often include the (dis)association of mobile devices with respect to APs. In this paper, we use the same dataset in [24]. A sample log entry has the following format:

`timestamp,process,ap-name,student-id,role,MAC, SSID,result`

The fields in the log represent the event's UNIX timestamp, the process that generated the log entry, the AP name, the encrypted student ID, the role assigned to the device, the anonymized MAC address (preserving the OUI), the SSID name, and the authentication result (success or failure), respectively.

Our WLAN dataset collection consists of three steps: (1) We first filter out students who use the university's unsecured WLAN from the dataset. Some information is missing regarding student ID and AP's name. We consider these log entries invalid in this work. After removing invalid entries from the dataset, 39 million log entries remain. (2) Since WLAN logs only provide the association (arrival) time of the person at the corresponding AP, we need to estimate the disassociation (leave) time. We first sort the log entries of each student in ascending order (based on timestamps) to ensure sequential order. For APs within the same building, the stay time of each AP is the duration between the arrival time of the next AP and the current one. Following [24], we also calculate the estimated walking time between two buildings using the Google Maps API [35]. (3) In [24], the location granularity is building-level as that work focuses on human mobility measurement [36]. In contrast, we treat the AP as the base unit in the trajectory in order to study human proximity tracing. Therefore, after data processing, each user/MAC's trajectory becomes a time series of APs and their corresponding stay times. A person's trajectory T can be defined as:

$$T = (AP_1, t_1, ST_1) \rightarrow (AP_2, t_2, ST_2) \rightarrow \cdots \rightarrow (AP_M, t_M, ST_M), t_1 < t_2 < \ldots < t_M,$$

where AP_i is the ith AP in trajectory T, t_i is the arrival time of the person at AP_i, and ST_i is the stay time of the person at AP_i. We refer to

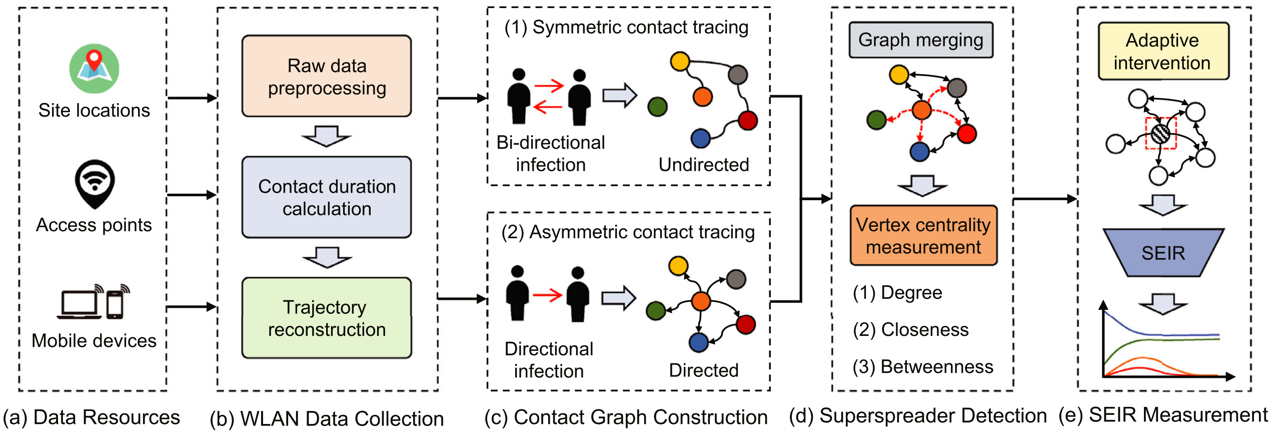


Fig. 3. Overview of WLAN-log-based superspreader detection. First, we extract contact graphs from WLAN logs via symmetric and asymmetric contact tracing. Second, we perform vertex centrality measurement to discover potential superspreaders. Finally, we simulate adaptive interventions using the SEIR model.

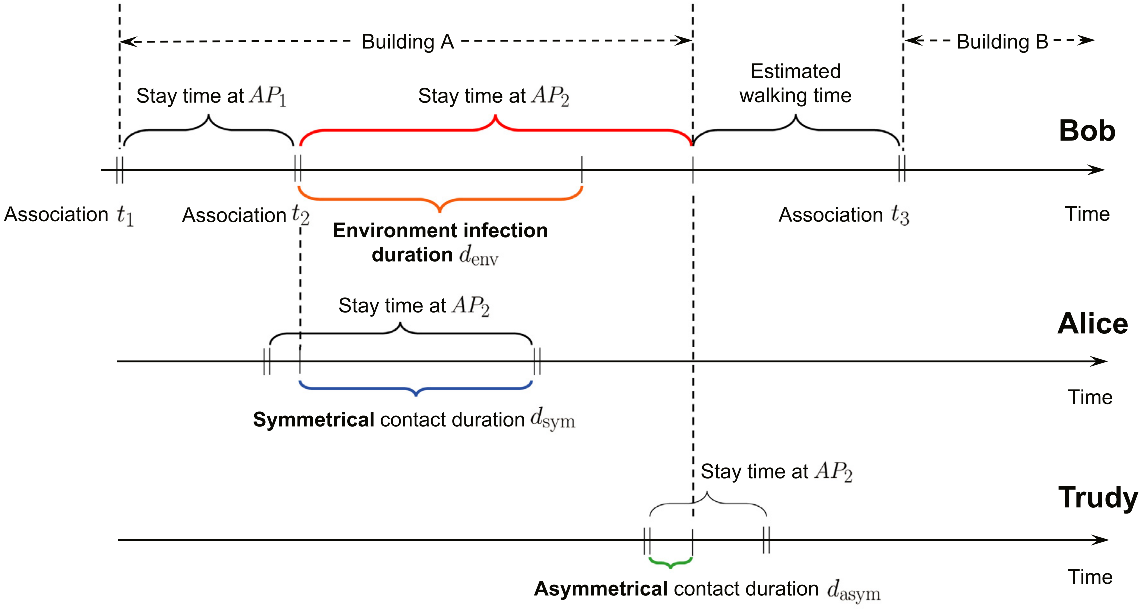


Fig. 4. Contact tracing using persons' trajectories. We show trajectories of three persons, i.e., Bob, Alice, and Trudy. At AP_2, Bob's stay time (red) is longer than the environmental infection duration (orange). There is a symmetric contact (blue) between Bob and Alice and an asymmetric contact (green) between Bob and Trudy.

(AP_i, t_i, ST_i) as a *tracklet*. Fig. 4 (top) shows how we estimate stay time for intra- and inter-building AP connections for a person's trajectory and illustrate Bob's trajectory between two buildings.

3.2. Contact graph construction

Next, we describe the contact tracing method using persons' trajectories. Given a student's trajectory T with sequential tracklets, we take each tracklet as a query and apply beam search on all other students' tracklets to determine if there is an overlapping duration for physical interaction between two persons. Fig. 4 shows an example where we consider two contact tracing methods—symmetric and asymmetric—to compute the overlapping duration between Bob and two other persons, Alice and Trudy.

Symmetric contact tracing Intuitively, if Bob and Alice connect to the same AP with a certain overlapping period, we assume there may be a potential physical interaction between them. Thus, given a tracklet (AP_q, t_q, ST_q) from student q (e.g., Bob) and a tracklet (AP_p, t_p, ST_p) from student p (e.g., Alice), we assign a bidirectional[3] contact edge between q and p if $AP_q = AP_p$ and the following criterion is satisfied:

$$ST_q + ST_p - \max\{t_q + ST_q, t_p + ST_p\} + \min\{t_q, t_p\} \geq d_{\text{sym}}, \tag{2}$$

where d_{sym} is a constant value of symmetric contact duration. Empirically, we set d_{sym} to 15 min in the experiments.

Asymmetric contact tracing However, the above symmetric tracing method omits environmental infection (cf. Section 1). In that situation, Bob may stay at AP_2 for a long enough period, making the environment infected. Thus, the virus will spread to another person, Trudy, even though the overlapping contact duration is short. To resolve this problem, we propose a new asymmetric contact tracing method that can discover such directional interactions. Concretely, we take Bob's tracklet whose stay time ST_q exceeds a certain duration d_{env} and assign a directional[4] contact edge between q (e.g., Bob) and p (e.g., Trudy) if

[3] The virus can spread from person q to person p, and vice versa.

[4] The virus may only spread from one person to another.

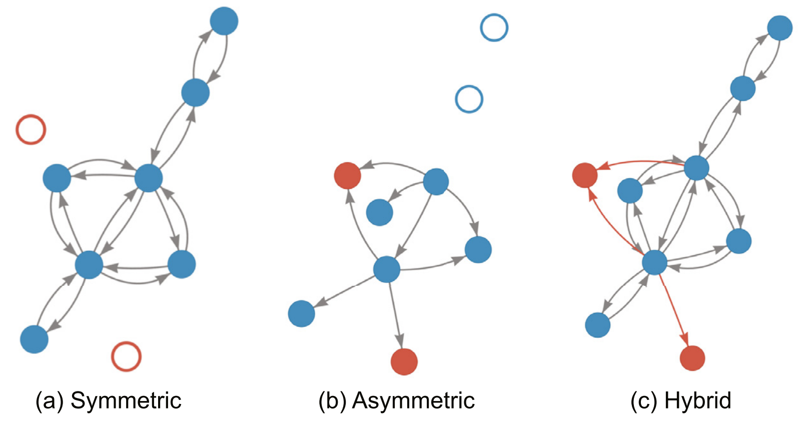


Fig. 5. Symmetric, asymmetric, and hybrid contact graphs. We show different contact tracing results of a real case from a group of students in our WLAN dataset. (a) Contact graph only with symmetric tracing: the *unfilled* red nodes are overlooked due to short overlapped stay time with other blue nodes. (b) Contact graph with asymmetric tracing: we observe that *filled* red nodes are included if directional contact is considered. (c) Merging symmetric and asymmetric graphs to construct a hybrid graph: red nodes and edges indicate newly discovered information compared to the symmetric contact graph.

$AP_q = AP_p$ and the following criterion is satisfied:

$$(ST_q - d_{\text{env}}) + ST_p - \max\{t_q + ST_q, t_p + ST_p\} + \min\{t_q + d_{\text{env}}, t_p\} \geq d_{\text{asym}}, \tag{3}$$

where d_{env} and d_{asym} are constant values of environmental infection time and asymmetric contact duration, respectively. Empirically, we set d_{env} to 50 min and d_{asym} to 5 min in the experiments.

Graph merging Once both symmetric and asymmetric contact graphs are obtained, we merge two graphs into one hybrid graph by aligning nodes and edges. The hybrid graph can reveal realistic contacts in our social interactions evidenced by WLAN logs. Fig. 5 gives an example for each graph.

3.3. Superspreader detection via vertex centrality measurement

The reader may ask a key question about a vertex in the hybrid graph: *How "important" is a specific person in the spread of disease?* Centrality measurements [20] are designed to quantify a person's importance, helping answer this question. Accordingly, the purpose of this subsection is *not* to propose a new metric for vertex measurement. Rather, we investigate the efficacy of three metrics in representing superspreaders in the Wi-Fi-based contact graphs. Fig. 6 shows their differences.

Degree centrality Degree centrality is defined as the number of edges incident upon a vertex (i.e., the vertex's number of social ties). If the network is directed, then two separate measures of degree centrality are defined: in-degree and out-degree. In this paper, we define each vertex's out-degree as follows:

$$deg(u) = \frac{|E_u^o|}{N-1}, \tag{4}$$

where $|E_u^o|$ is the total number of edges directed out of a vertex u in a directed hybrid contact graph, and N is the number of vertices in the graph.

Closeness centrality One common notion of centrality is a vertex's "nearness" to many other vertices, which closeness centrality metrics aim to capture. For a given vertex, closeness centrality varies inversely with the vertex's distance of a vertex from all others. Formally, for a connected graph, this measure is defined as:

$$cl(u) = \frac{1}{\sum_v dist(u,v)}, \tag{5}$$

where $dist(u,v)$ denotes the geodesic (shortest-path) distance between vertices u and v. Intuitively, this measure looks at how fast information can spread from one vertex to all others. For example, a vertex that is close to many other vertices may easily transmit the disease to them.

Betweenness centrality Another popular class of centralities is based upon the perspective that "importance" relates to a vertex's position regarding paths in the graph. If we picture those paths as the routes by which communication takes place, vertices situated on many paths tend to be more critical to the communication process. Betweenness centrality metrics are aimed at summarizing the extent to which a vertex is located "between" other pairs of vertices:

$$bw(u) = \sum_{s \neq t \neq v} \frac{\sigma(s,t|v)}{\sigma(s,t)}, \tag{6}$$

where $\sigma(s,t|v)$ is the total number of shortest paths between s and t that pass through v and $\sigma(s,t) = \sum_v \sigma(s,t|v)$. Vertices with high betweenness centrality are critical for maintaining graph connectivity.

SEIR measurement Based on these centrality measures, we are able to identify potential superspreaders. Next, we perform adaptive interventions on those active nodes using SEIR simulations to measure who are the most critical individuals for the spread of disease based on the simulation results.

4. Evaluation of the proposed framework

In this section, we first describe our methodology. Next, we present our experimental results.

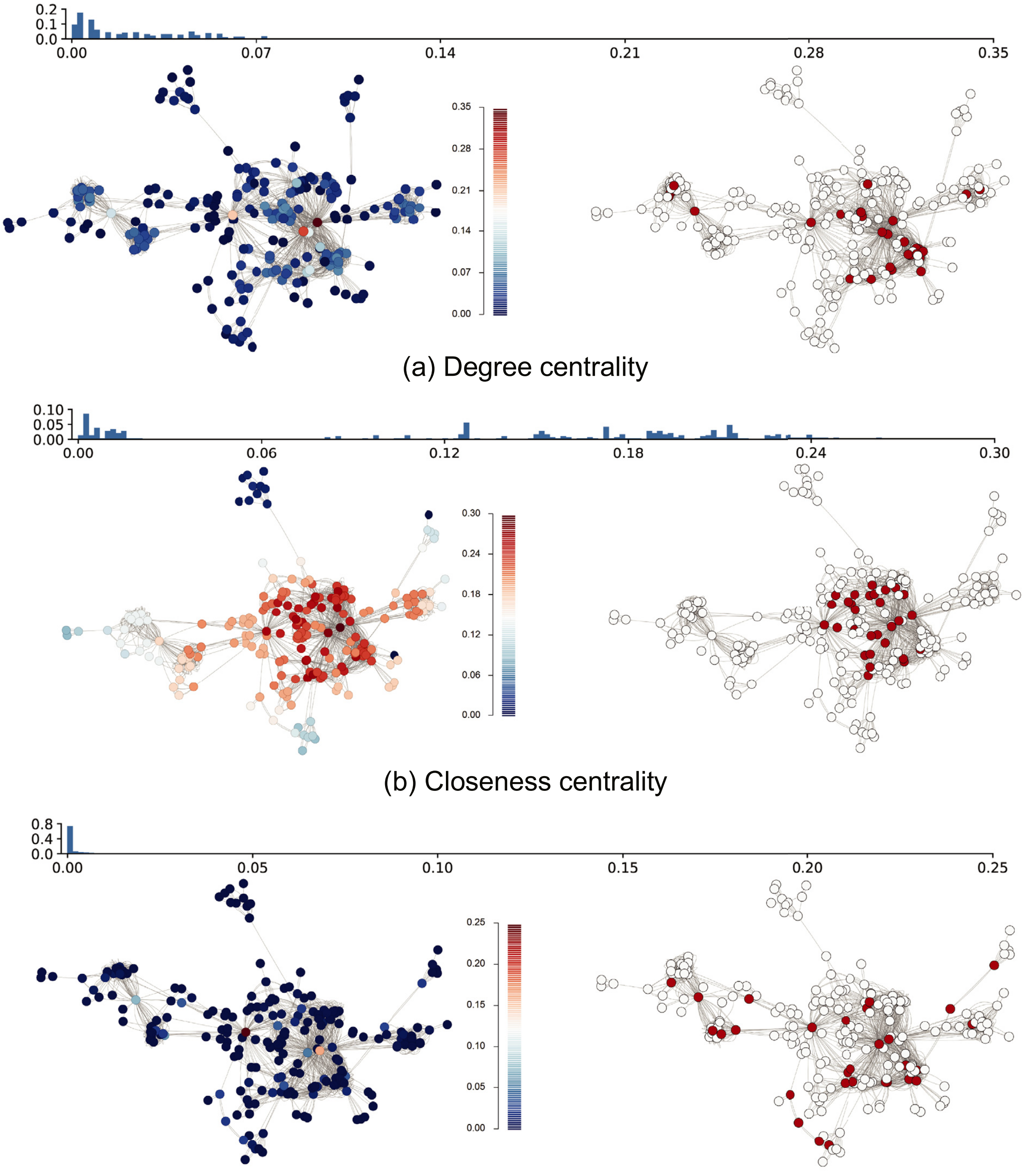
4.1. Evaluation setup

WLAN dataset We use the WLAN dataset from Cao et al. [24], which contains WLAN log data with demographic information at a large public university spanning 139 days in 2015. Cao et al. [24] found that university students' mobility patterns change periodically on a weekly basis. In our study, we focus on analyzing the contact graph from a specific day of the week in the dataset. Specifically, we use the WLAN log information to compute the contact graph for each weekday from a randomly selected week in the dataset. We also report results on the contact graphs computed from a weekly period. We construct three types of contact graphs: symmetric, asymmetric, and hybrid.

Evaluation metrics Based on the SEIR model, we use the following realistic epidemiological measures to estimate the effect of different approaches:

- **Doubling Time (day):** the time it takes for the number of cumulative infections to double.
- **Total Infected Fraction (%)**: the fraction of the total accumulated infected population during the entire epidemic.
- **Peak Infected Time (day)**: the time required to infect the largest possible population.
- **Peak Infected Fraction (%)**: the fraction of infected persons when peak infection is reached.

Experimental comparison We quarantine persons with higher centrality based on the hybrid contact graph and simulate the epidemic on the



(a) Degree centrality

(b) Closeness centrality

(c) Betweenness centrality

Fig. 6. Visualization of vertex centrality measurement. We show a one-day contact graph of a building on campus with (a) degree centrality, (b) closeness centrality, and (c) betweenness centrality measurements. The top, left, and right part of each indicates the relative frequency histogram, centrality graphs, and the top 10% of highlighted nodes (red), respectively. Warmer colors indicate larger values. Discrepancies among the three measurements are visible.

hybrid graph. We test three vertex centrality measurement methods and compare our results to the following baselines:

- **No quarantine**: we let the virus spread naturally on the hybrid graph without intervention.
- **Random**: we randomly quarantine a certain number of persons and simulate the epidemic on the hybrid graph.
- **Symmetric contact tracing (SymC)**: we quarantine persons with higher centrality based on the symmetric contact graph and simulate the epidemic on the hybrid graph.
- **Symmetric and asymmetric contact tracing (Hybrid)**: we quarantine persons with higher centrality based on the hybrid contact graph and simulate the epidemic on the hybrid graph.

Implementation details We follow [6] in order to simulate an epidemic using the SEIR model. We use the default SEIR parameters, as they are calculated from a real-world infectious dataset.[5] In particular, the total

[5] Other toolkits could be used to simulate the spread of disease elsewhere.

Table 1
Main results on single-day contact graph. We compare different methods with various centrality metrics. Next, we perform SEIR simulation by quarantining 100 persons based on these metrics. We observe that our hybrid graph, which jointly considers symmetric and asymmetric contact tracing, achieves better performance than the baseline model and the symmetric contact tracing method alone. **DB-Time:** Doubling Time (day); **T-Inf:** Total Infected Fraction (%); **PK-Time:** Peak Infection Time (day); **PK-Inf:** Peak Infection Fraction (%). Results in blue show where the hybrid graph outperforms SymC. The top result in each column is in **bold**.

Method	Measure	DB-Time (↑)	T-Inf (↓)	PK-Time (↑)	PK-Inf (↓)
No quarantine	-	3.24	48.45	29.00	4.17
Random	-	3.29	44.90	30.40	3.91
SymC	Degree	5.61	40.69	40.80	2.53
Hybrid		(−0.70) 4.91	(−1.92) 38.77	(−0.60) 40.20	(−0.27) **2.26**
SymC	Closeness	6.08	40.21	39.80	2.37
Hybrid		(+0.51) **6.59**	(−1.74) 38.47	(+1.20) 41.00	(−0.09) 2.28
SymC	Betweenness	5.44	42.61	39.80	2.51
Hybrid		(+0.21) 5.65	(−5.20) **37.41**	(+1.40) **41.20**	(−0.06) 2.45

population size in the our experiments is 3748. We set the initial number of infected persons to 50, which we fix across all experiments. In order to achieve stable observations, We run our simulation 50 times in each group of experiments until convergence is reached.

4.2. Results and analyses

Main results on single-day contact graph We report the main results of a single-day contact graph in Table 1. Identifying superspreaders using a hybrid graph with asymmetric and symmetric contact tracing outperforms baseline methods substantially in terms of all centralities, justifying our motivation: *symmetric and asymmetric contact tracing, which naturally reflects environmental infection, can be a valuable factor to contain the spread of disease.* In addition, we find similar observations from other days of the week in the WLAN dataset. Next, we detail our analyses.

Superspreaders exist on the university campus. We notice that both SymC and Hybrid significantly outperform baseline and “random quarantine,” suggesting the existence of superspreaders and the importance of contact tracing to limit the spread of disease. To analyze these superspreaders’ extent of spread, we conduct a simulated comparison by initializing different groups of individuals. As shown in Fig. 7, we observe that the virus carried by students with higher centrality causes a much faster spread than with randomly selected students. Further, students with higher betweenness centrality are critical to the spread.

Asymmetric contact tracing is efficient. We found that asymmetric contact tracing with a simple vertex measure leads to a notable gain for all metrics. Especially for the total infected fraction (T-Inf), Hybrid is ∼1% better than symmetric contact tracing (SymC), which represents around 40 persons in our WLAN dataset. We also show the SEIR simulation curves in Fig. 8: both symmetric and asymmetric contact tracing methods significantly outperform random quarantine methods, demonstrating the effectiveness of our superspreader detection framework.

Betweenness centrality strongly limits the total infected population on daily contact graphs. By comparing different centrality measurements for the selection of quarantine populations, we found that betweenness centrality leads to the strongest reduction in the total infected fraction (from 42.61% to 37.41%) in the daily contact graph (cf. Table 1). One reason is that betweenness metrics can effectively discover vertices that sit on many paths are likely more critical to the spread process in social graphs. This verifies our observation in Fig. 6 that betweenness centrality identifies a very different group of persons compared to degree centrality and closeness centrality (cf. Section 3.3).

Further, we extend the simulation on contact graphs computed over longer weekly periods. Compared to daily contact graphs, weekly graphs generated from the WLAN logs are more densely connected. We focus on betweenness centrality and Table 2 shows the results. We find that the difference between the symmetric and the hybrid graphs is marginal. This is because, in long-term contact tracing, the top superspreaders between the symmetric and hybrid graphs overlap highly, suggesting that early-stage pandemic control helps identify superspreaders who may be missed otherwise. We observe similar patterns for other weeks throughout the study period.

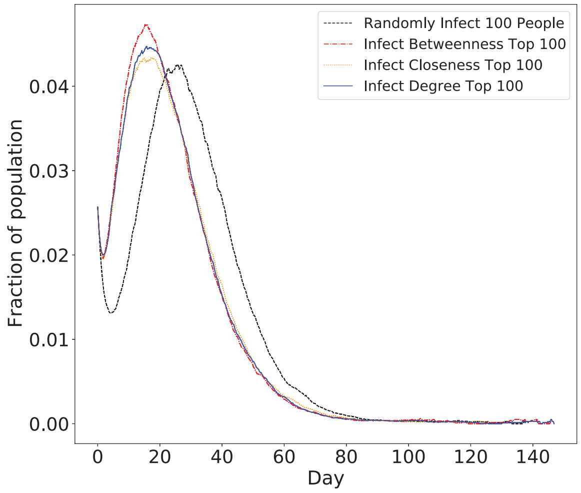


Fig. 7. **Effect of the infected population on the spread of the pandemic.** We select 100 students and set their initial conditions as infectious based on different criteria. We run SEIR simulation and show the fraction of the infected population on different days.

Table 2
Results on one-week contact graph. We compare different methods using a one-week contact graph. We perform SEIR simulation by quarantining 100 persons based on centrality metrics. We use betweenness centrality to discover superspreaders. **CM:** Centrality Measure; **DB-Time:** Doubling Time (day); **T-Inf:** Total Infected Fraction (%); **PK-Time:** Peak Infection Time (day); **PK-Inf:** Peak Infection Fraction (%).

Method	DB-Time (↑)	T-Inf (↓)	PK-Time (↑)	PK-Inf (↓)
No quarantine	0.98	86.15	13.04	17.17
Random	0.98	83.93	13.04	16.61
SymC	1.09	81.88	13.96	16.18
Hybrid	1.11	82.57	13.84	16.17

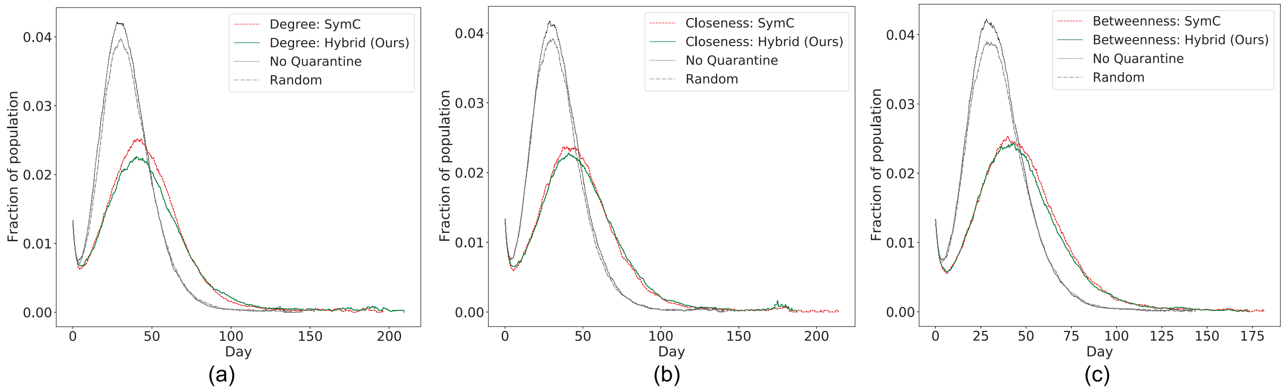


Fig. 8. Spread of the pandemic during the period. We show comparison results on (a) degree centrality, (b) closeness centrality, and (c) betweenness centrality measurements. Our asymmetric contact tracing and symmetric contact tracing (green and red) outperforms the baseline approaches with random quarantine (gray).

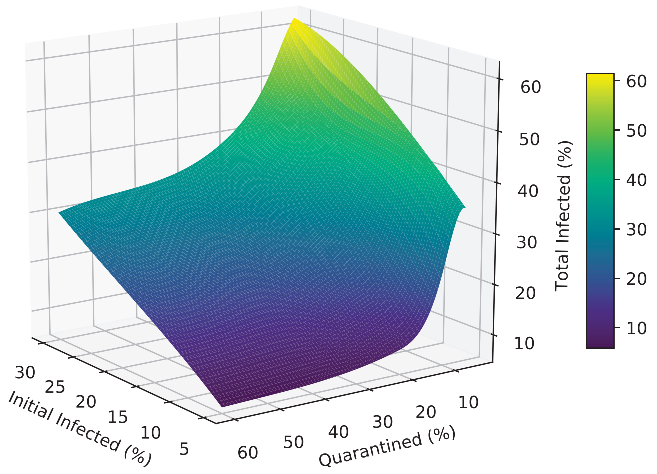


Fig. 9. Effect of infected population in the spread of the pandemic. We show the fraction of total infected in terms of fractions of initial infected and quarantined people.

How to perform quarantine with constrained resources? Next, we study suitable proportions of the population for intervention. We show the total infected fraction with respect to different amounts of infectious and quarantined populations. Fig. 9 shows the results based on the betweenness centrality measure. We observe a clear turning point where quarantining 20% of the whole population reduces the spread of disease among all infected ratios. This suggests that increasing the quarantine percentage over 20% provides only marginal benefits.

Will superspreaders change during the whole semester? To further analyze the stability of superspreaders among different periods, we compute the similarities of the identified superspreaders from any two accumulated weeks, whose results are shown in Fig. 10. In this study, we first generate the contact graphs based on the first N weeks in the WLAN dataset, where N ranges from 1 to 20. Next, we select the top 100 students based on our centrality measurements. We adopt rank-biased overlap (RBO) [25] to compute the similarity of two ranked student lists from any two accumulated weeks. Our results show that the superspreaders change during the first few weeks, but remain stable throughout the rest of the semester. For example, the similarity between the first 20 weeks and 15 weeks is around 0.8, opening up opportunities to discover the superspreaders as early as possible for efficient pandemic mitigation.

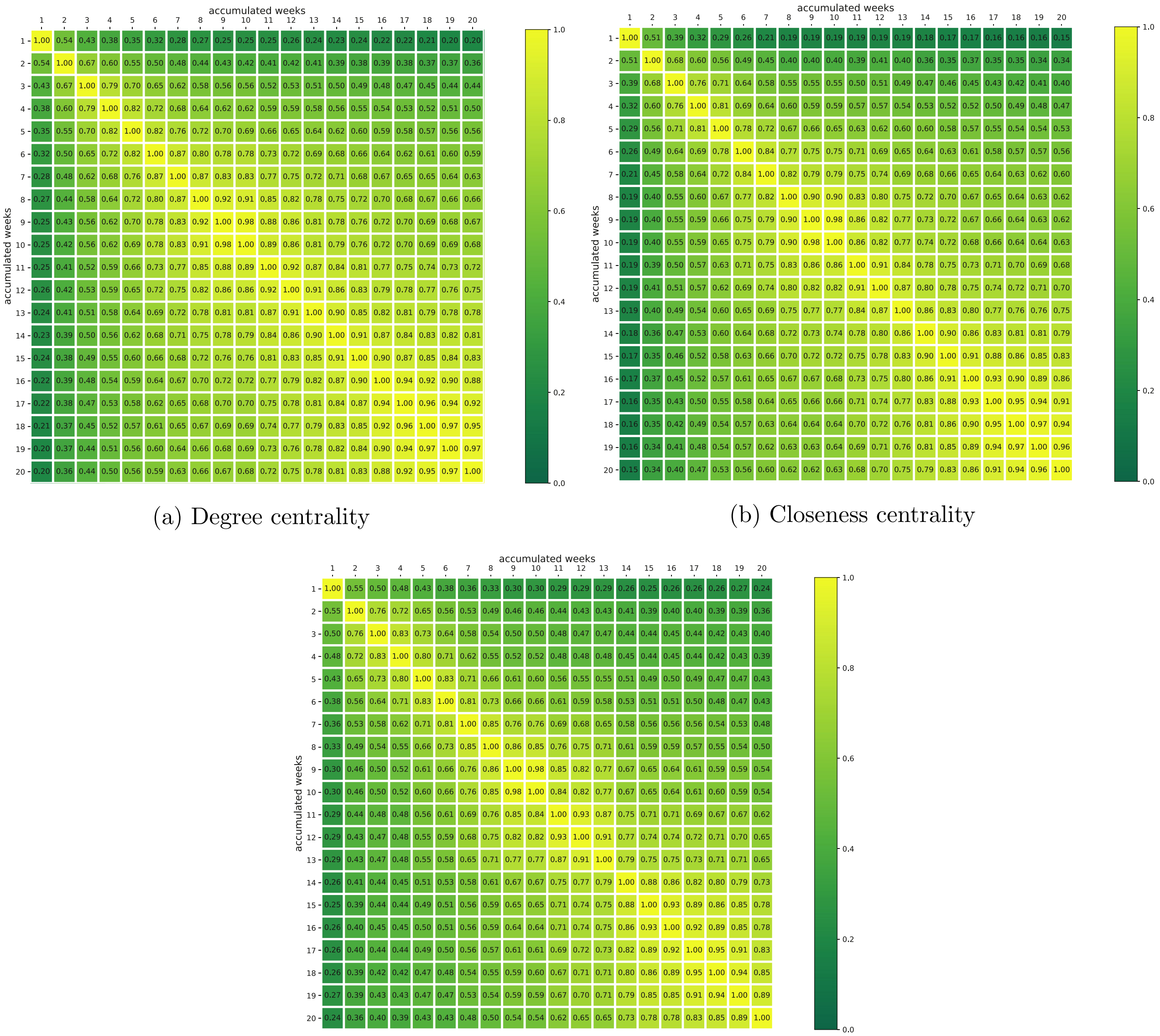
5. Related work

5.1. Client-based contact tracing

Researchers have devoted considerable attention to mobile application (app) technology for COVID-19 contact tracing. For example, Covid Watch [13] uses Bluetooth signals to detect when users are near each other and alerts them anonymously if they were in contact with someone who is later diagnosed with COVID-19. Similarly, PACT [14] uses inter-phone Bluetooth communications (including energy measurements) as a proxy for inter-person distance measurement. Through applied cryptography, this system can collect and maintain weeks of contact events. Later, PACT augments these events with infection notifications leading to exposure notifications to all mobile phone owners who have had medically significant contact (in terms of distance and time) with infected people in the past medically significant period (e.g., two weeks). In addition, Singapore launched the TraceTogether [16] app to boost COVID-19 contact tracing efforts. By downloading the app and consenting to participate in it, TraceTogether lets users “proactively help” in the contact tracing process [16]. The app works by exchanging short-range Bluetooth signals between phones to detect other app users who are nearby. Apple and Google [15] are working together for the first time on a protocol that will alert users if they have been exposed to the coronavirus. Luo et al. propose A-Turf [10], an acoustic encounter detection method for COVID-19 contact tracing. Compared with Bluetooth technology, the system more precisely detects encounters within 6-foot ranges (social distancing). Unlike the WLAN-log-based contact tracing presented in this paper, client-based contact tracing requires users’ widespread adoption and active participation.

5.2. Infrastructure-based contact tracing

Infrastructure-based methods take advantage of existing infrastructure deployed worldwide such as CCTV footage [17], locations measured using cellular networks [18], Wi-Fi hotspots [9,37], and GPS [19], without requiring client-side involvement. Similar to our approach, recent efforts [9,37] use passive Wi-Fi sensing for network-based contact tracing for infectious diseases, particularly focused on the COVID-19 pandemic. Those works mainly use location occupancy or number of contact as the measure to identify the superspreaders while we consider different types of centrality for measuring the “importance” of the vertex in the social networks. Moreover, we adopt SEIR simulation to justify which measure is better in discovering the superspearders.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.00	0.54	0.43	0.38	0.35	0.32	0.28	0.27	0.25	0.25	0.25	0.26	0.24	0.23	0.24	0.22	0.22	0.21	0.20	0.20
2	0.54	1.00	0.67	0.60	0.55	0.50	0.48	0.44	0.43	0.42	0.41	0.42	0.41	0.39	0.38	0.39	0.38	0.37	0.37	0.36
3	0.43	0.67	1.00	0.79	0.70	0.65	0.62	0.58	0.56	0.56	0.52	0.53	0.51	0.50	0.49	0.48	0.47	0.45	0.44	0.44
4	0.38	0.60	0.79	1.00	0.82	0.72	0.68	0.64	0.62	0.62	0.59	0.59	0.58	0.56	0.55	0.54	0.53	0.52	0.51	0.50
5	0.35	0.55	0.70	0.82	1.00	0.82	0.76	0.72	0.70	0.69	0.66	0.65	0.64	0.62	0.60	0.59	0.58	0.57	0.56	0.56
6	0.32	0.50	0.65	0.72	0.82	1.00	0.87	0.80	0.78	0.78	0.73	0.72	0.69	0.68	0.66	0.64	0.62	0.61	0.60	0.59
7	0.28	0.48	0.62	0.68	0.76	0.87	1.00	0.87	0.83	0.83	0.77	0.75	0.72	0.71	0.68	0.67	0.65	0.65	0.64	0.63
8	0.27	0.44	0.58	0.64	0.72	0.80	0.87	1.00	0.92	0.91	0.85	0.82	0.78	0.75	0.72	0.70	0.68	0.67	0.66	0.66
9	0.25	0.43	0.56	0.62	0.70	0.78	0.83	0.92	1.00	0.98	0.88	0.86	0.81	0.78	0.76	0.72	0.70	0.69	0.68	0.67
10	0.25	0.42	0.56	0.62	0.69	0.78	0.83	0.91	0.98	1.00	0.89	0.86	0.81	0.79	0.76	0.72	0.70	0.69	0.69	0.68
11	0.25	0.41	0.52	0.59	0.66	0.73	0.77	0.85	0.88	0.89	1.00	0.92	0.87	0.84	0.81	0.77	0.75	0.74	0.73	0.72
12	0.26	0.42	0.53	0.59	0.65	0.72	0.75	0.82	0.86	0.86	0.92	1.00	0.91	0.86	0.83	0.79	0.78	0.77	0.76	0.75
13	0.24	0.41	0.51	0.58	0.64	0.69	0.72	0.78	0.81	0.81	0.87	0.91	1.00	0.90	0.85	0.82	0.81	0.79	0.78	0.78
14	0.23	0.39	0.50	0.56	0.62	0.68	0.71	0.75	0.78	0.79	0.84	0.86	0.90	1.00	0.91	0.87	0.84	0.83	0.82	0.81
15	0.24	0.38	0.49	0.55	0.60	0.66	0.68	0.72	0.76	0.76	0.81	0.83	0.85	0.91	1.00	0.90	0.87	0.85	0.84	0.83
16	0.22	0.39	0.48	0.54	0.59	0.64	0.67	0.70	0.72	0.72	0.77	0.79	0.82	0.87	0.90	1.00	0.94	0.92	0.90	0.88
17	0.22	0.38	0.47	0.53	0.58	0.62	0.65	0.68	0.70	0.70	0.75	0.78	0.81	0.84	0.87	0.94	1.00	0.96	0.94	0.92
18	0.21	0.37	0.45	0.52	0.57	0.61	0.65	0.67	0.69	0.69	0.74	0.77	0.79	0.83	0.85	0.92	0.96	1.00	0.97	0.95
19	0.20	0.37	0.44	0.51	0.56	0.60	0.64	0.66	0.68	0.69	0.73	0.76	0.78	0.82	0.84	0.90	0.94	0.97	1.00	0.97
20	0.20	0.36	0.44	0.50	0.56	0.59	0.63	0.66	0.67	0.68	0.72	0.75	0.78	0.81	0.83	0.88	0.92	0.95	0.97	1.00

(a) Degree centrality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.00	0.51	0.39	0.32	0.29	0.26	0.21	0.19	0.19	0.19	0.19	0.19	0.19	0.18	0.17	0.17	0.16	0.16	0.16	0.15
2	0.51	1.00	0.68	0.60	0.56	0.49	0.45	0.40	0.40	0.40	0.39	0.41	0.40	0.36	0.35	0.37	0.35	0.35	0.34	0.34
3	0.39	0.68	1.00	0.76	0.71	0.64	0.58	0.55	0.55	0.55	0.50	0.51	0.49	0.47	0.46	0.45	0.43	0.42	0.41	0.40
4	0.32	0.60	0.76	1.00	0.81	0.69	0.64	0.60	0.59	0.59	0.57	0.57	0.54	0.53	0.52	0.52	0.50	0.49	0.48	0.47
5	0.29	0.56	0.71	0.81	1.00	0.78	0.72	0.67	0.66	0.65	0.63	0.62	0.60	0.60	0.58	0.57	0.55	0.54	0.54	0.53
6	0.26	0.49	0.64	0.69	0.78	1.00	0.84	0.77	0.75	0.75	0.71	0.69	0.65	0.64	0.63	0.61	0.58	0.57	0.57	0.56
7	0.21	0.45	0.58	0.64	0.72	0.84	1.00	0.82	0.79	0.79	0.75	0.74	0.69	0.68	0.66	0.65	0.64	0.63	0.62	0.60
8	0.19	0.40	0.55	0.60	0.67	0.77	0.82	1.00	0.90	0.90	0.83	0.80	0.75	0.72	0.70	0.67	0.65	0.64	0.63	0.62
9	0.19	0.40	0.55	0.59	0.66	0.75	0.79	0.90	1.00	0.98	0.86	0.82	0.77	0.73	0.72	0.67	0.66	0.64	0.63	0.62
10	0.19	0.40	0.55	0.59	0.65	0.75	0.79	0.90	0.98	1.00	0.86	0.82	0.77	0.74	0.72	0.68	0.66	0.64	0.64	0.63
11	0.19	0.39	0.50	0.57	0.63	0.71	0.75	0.83	0.86	0.86	1.00	0.91	0.84	0.78	0.75	0.73	0.71	0.70	0.69	0.68
12	0.19	0.41	0.51	0.57	0.62	0.69	0.74	0.80	0.82	0.82	0.91	1.00	0.87	0.80	0.78	0.75	0.74	0.72	0.71	0.70
13	0.19	0.40	0.49	0.54	0.60	0.65	0.69	0.75	0.77	0.77	0.84	0.87	1.00	0.86	0.83	0.80	0.77	0.76	0.76	0.75
14	0.18	0.36	0.47	0.53	0.60	0.64	0.68	0.72	0.73	0.74	0.78	0.80	0.86	1.00	0.90	0.86	0.83	0.81	0.81	0.79
15	0.17	0.35	0.46	0.52	0.58	0.63	0.66	0.70	0.72	0.72	0.75	0.78	0.83	0.90	1.00	0.91	0.88	0.86	0.85	0.83
16	0.17	0.37	0.45	0.52	0.57	0.61	0.65	0.67	0.67	0.68	0.73	0.75	0.80	0.86	0.91	1.00	0.93	0.90	0.89	0.86
17	0.16	0.35	0.43	0.50	0.55	0.58	0.64	0.65	0.66	0.66	0.71	0.74	0.77	0.83	0.88	0.93	1.00	0.95	0.94	0.91
18	0.16	0.35	0.42	0.49	0.54	0.57	0.63	0.64	0.64	0.64	0.70	0.72	0.76	0.81	0.86	0.90	0.95	1.00	0.97	0.94
19	0.16	0.34	0.41	0.48	0.54	0.57	0.62	0.63	0.63	0.64	0.69	0.71	0.76	0.81	0.85	0.89	0.94	0.97	1.00	0.96
20	0.15	0.34	0.40	0.47	0.53	0.56	0.60	0.62	0.62	0.63	0.68	0.70	0.75	0.79	0.83	0.86	0.91	0.94	0.96	1.00

(b) Closeness centrality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.00	0.55	0.50	0.48	0.43	0.38	0.36	0.33	0.30	0.30	0.29	0.29	0.29	0.26	0.25	0.26	0.26	0.26	0.27	0.24
2	0.55	1.00	0.76	0.72	0.65	0.56	0.53	0.49	0.46	0.46	0.44	0.43	0.43	0.41	0.39	0.40	0.40	0.39	0.39	0.36
3	0.50	0.76	1.00	0.83	0.73	0.64	0.58	0.54	0.50	0.50	0.48	0.47	0.47	0.44	0.44	0.45	0.44	0.42	0.43	0.40
4	0.48	0.72	0.83	1.00	0.80	0.71	0.62	0.55	0.52	0.52	0.48	0.48	0.48	0.45	0.44	0.45	0.44	0.42	0.43	0.39
5	0.43	0.65	0.73	0.80	1.00	0.83	0.71	0.66	0.61	0.60	0.56	0.55	0.55	0.51	0.49	0.50	0.49	0.47	0.47	0.43
6	0.38	0.56	0.64	0.71	0.83	1.00	0.81	0.73	0.66	0.66	0.61	0.59	0.58	0.53	0.51	0.51	0.50	0.48	0.47	0.43
7	0.36	0.53	0.58	0.62	0.71	0.81	1.00	0.85	0.76	0.76	0.69	0.68	0.65	0.58	0.56	0.56	0.56	0.54	0.53	0.48
8	0.33	0.49	0.54	0.55	0.66	0.73	0.85	1.00	0.86	0.85	0.76	0.75	0.71	0.61	0.59	0.59	0.57	0.55	0.54	0.50
9	0.30	0.46	0.50	0.52	0.61	0.66	0.76	0.86	1.00	0.98	0.85	0.82	0.77	0.67	0.65	0.64	0.61	0.59	0.59	0.54
10	0.30	0.46	0.50	0.52	0.60	0.66	0.76	0.85	0.98	1.00	0.84	0.82	0.77	0.67	0.65	0.64	0.61	0.60	0.59	0.54
11	0.29	0.44	0.48	0.48	0.56	0.61	0.69	0.76	0.85	0.84	1.00	0.93	0.87	0.75	0.71	0.71	0.69	0.67	0.67	0.62
12	0.29	0.43	0.47	0.48	0.55	0.59	0.68	0.75	0.82	0.82	0.93	1.00	0.91	0.77	0.74	0.74	0.72	0.71	0.70	0.65
13	0.29	0.43	0.47	0.48	0.55	0.58	0.65	0.71	0.77	0.77	0.87	0.91	1.00	0.79	0.75	0.75	0.73	0.71	0.71	0.65
14	0.26	0.41	0.44	0.45	0.51	0.53	0.58	0.61	0.67	0.67	0.75	0.77	0.79	1.00	0.88	0.86	0.82	0.80	0.79	0.73
15	0.25	0.39	0.44	0.44	0.49	0.51	0.56	0.59	0.65	0.65	0.71	0.74	0.75	0.88	1.00	0.93	0.89	0.86	0.85	0.78
16	0.26	0.40	0.45	0.45	0.50	0.51	0.56	0.59	0.64	0.64	0.71	0.74	0.75	0.86	0.93	1.00	0.92	0.89	0.85	0.78
17	0.26	0.40	0.44	0.44	0.49	0.50	0.56	0.57	0.61	0.61	0.69	0.72	0.73	0.82	0.89	0.92	1.00	0.95	0.91	0.83
18	0.26	0.39	0.42	0.42	0.47	0.48	0.54	0.55	0.59	0.60	0.67	0.71	0.71	0.80	0.86	0.89	0.95	1.00	0.94	0.85
19	0.27	0.39	0.43	0.43	0.47	0.47	0.53	0.54	0.59	0.59	0.67	0.70	0.71	0.79	0.85	0.85	0.91	0.94	1.00	0.89
20	0.24	0.36	0.40	0.39	0.43	0.43	0.48	0.50	0.54	0.54	0.62	0.65	0.65	0.73	0.78	0.78	0.83	0.85	0.89	1.00

(c) Betweenness centrality

Fig. 10. Similarity matrix of superspreaders between accumulated weeks.

6. Conclusion

In this paper, we focused on WLAN-log-based superspreader detection in the COVID-19 pandemic. We proposed a general framework with applications to a wide range of working scenarios based on users' preferences, environmental dynamics, and resource availability. Moreover, we presented asymmetric contact, a new type of human contact. The concept of asymmetric contact partially captured the notion of environmental infection. We required that persons in asymmetric contact must have had a certain overlap time between their association times with a specific AP. In fact, we can generalize by eliminating this constraint. We can treat the overlap time as a control knob to adjust the degree of "asymmetry". Due to space limitations, this remains part of our future work. We have implemented our framework, conducted an extensive evaluation, and obtained a set of important findings. Our proposed contact tracing framework and our findings provided a tool as well as guidelines for public health administrators regarding both proactive and reactive interventions against the pandemic.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The work was supported in part by the National Science Foundation (NSF) under Grant No. CNS 2028547. Any opinions, findings, conclusions, and recommendations in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

References

[1] K.T. Eames, M.J. Keeling, Contact tracing and disease control, Proc. R. Soc. Lond. Ser. B: Biol. Sci. 270 (1533) (2003) 2565–2571.

[2] J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday, A.J. Kucharski, W.J. Edmunds, F. Sun, et al., Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts, Lancet Global Health (2020).

[3] D. Klinkenberg, C. Fraser, H. Heesterbeek, The effectiveness of contact tracing in emerging epidemics, PloS One 1 (1) (2006) e12.

[4] M. Salathé, C.L. Althaus, R. Neher, S. Stringhini, E. Hodcroft, J. Fellay, M. Zwahlen, G. Senti, M. Battegay, A. Wilder-Smith, et al., Covid-19 epidemic in switzerland: on the importance of testing, contact tracing and isolation., Swiss Med. Wkly. 150 (11-12) (2020) w20225.

[5] R. Singh, W. Ren, F. Liu, D. Xuan, Z. Lin, N.B. Shroff, Aa blueprint for effective pandemic mitigation, ITU J. Fut. Evolv. Technol. (2020).

[6] O. Reich, G. Shalev, T. Kalvari, Modeling covid-19 on a network: super-spreaders, testing and containment, medRxiv (2020).

[7] H. Wen, Q. Zhao, Z. Lin, D. Xuan, N. Shroff, A study of the privacy of covid-19 contact tracing apps, in: International Conference on Security and Privacy in Communication Systems, Springer, 2020, pp. 297–317.

[8] Q. Zhao, H. Wen, Z. Lin, D. Xuan, N. Shroff, On the accuracy of measured proximity of bluetooth-based contact tracing apps, in: International Conference on Security and Privacy in Communication Systems, Springer, 2020, pp. 49–60.

[9] A. Trivedi, C. Zakaria, R. Balan, P. Shenoy, Wifitrace: network-based contact tracing for infectious diseasesusing passive wifi sensing, arXiv preprint arXiv:2005.12045(2020).

[10] Y. Luo, C. Zhang, Y. Zhang, C. Zuo, D. Xuan, Z. Lin, A.C. Champion, N. Shroff, Acoustic-turf: acoustic-based privacy-preserving covid-19 contact tracing, arXiv preprint arXiv:2006.13362(2020).

[11] Contact Tracing for COVID-19, 2020, (https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html), (Accessed on 02/15/2021).

[12] Pingbo information technology - trustkernel, 2020, (https://www.trustkernel.com/en/), (Accessed on 02/15/2021).

[13] C. Watch, 2020, (https://www.covid-watch.org/), (Accessed on 02/15/2021).

[14] P.A.C. Tracing, 2020, (https://pact.mit.edu/), (Accessed on 02/15/2021).

[15] Apple, G. partner on COVID-19 contact tracing technology, 2020, (https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/), (Accessed on 02/15/2021).

[16] G. of Singapore, Trace together, safer together, 2020, (https://www.tracetogether.gov.sg), (Accessed on 02/15/2021).

[17] D. Skoll, J. Miller, L. Saxon, Covid-19 testing and infection surveillance: is a combined digital contact tracing and mass testing solution feasible in the united states? Cardiovasc. Digit. Health J. (2020).

[18] N.A. Ayan, N.L. Damasceno, S. Chaskar, P.R. de Sousa, A. Ramesh, A. Seetharam, A.A.d. A. Rocha, Characterizing human mobility patterns during covid-19 using cellular network data, arXiv preprint arXiv:2010.14558(2020).

[19] J. Bay, J. Kek, A. Tan, C.S. Hau, L. Yongquan, J. Tan, T.A. Quy, Bluetrace: A Privacy-Preserving Protocol for Community-Driven Contact Tracing Across Borders, Government Technology Agency-Singapore, Technical Report(2020).

[20] E.D. Kolaczyk, G. Csárdi, Statistical Analysis of Network Data with R, 65, Springer, 2014.

[21] Guidelines for Environmental Infection Control in Health-Care Facilities (2003), 2020, (https://www.cdc.gov/infectioncontrol/guidelines/environmental/background/air.html), (Accessed on 02/15/2021).

[22] Mapzen, 2020, Https://www.mapzen.com.

[23] OpenStreetMap, 2020, https://www.openstreetmap.com.

[24] P.Y. Cao, G. Li, A.C. Champion, D. Xuan, S. Romig, W. Zhao, On human mobility predictability via wlan logs, in: IEEE INFOCOM, 2017, pp. 1–9.

[25] W. Webber, A. Moffat, J. Zobel, A similarity measure for indefinite rankings, ACM Trans. Inf. Syst. (TOIS) 28 (4) (2010) 1–38.

[26] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. Ser. A, Contain. Pap. Math. Phys. Charact. 115 (772) (1927) 700–721.

[27] H.W. Hethcote, Three basic epidemiological models, in: Applied Mathematical Ecology, Springer, 1989, pp. 119–144.

[28] R.M. Anderson, R.M. May, Infectious Diseases of Humans: Dynamics and Control, Oxford university press, 1992.

[29] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (4) (2000) 599–653, doi:10.1137/S0036144500371907.

[30] E. Vynnycky, R. White, An Introduction to Infectious Disease Modelling, Oxford University Press, 2010.

[31] N.M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, et al., Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. 2020, DOI 10 (2020) 77482.

[32] A.L. Bertozzi, E. Franco, G. Mohler, M.B. Short, D. Sledge, The challenges of modeling and forecasting the spread of covid-19, arXiv preprint arXiv:2004.04741(2020).

[33] J.-D. Van Wees, S. Osinga, M. van der Kuip, M. Tanck, M. Hanegraaf, M. Pluymaekers, O. Leeuwenburgh, L. Van Bijsterveldt, J. Zindler, M. Van Furth, Forecasting hospitalization and icu rates of the covid-19 outbreak: An efficient seir model, Bull. World Health Organ. (2020).

[34] A. Leung, X. Ding, S. Huang, R. Rabbany, Contact graph epidemic modelling of covid-19 for transmission and intervention strategies, arXiv preprint arXiv:2010.03081(2020).

[35] Google Maps API, 2020, (https://www.google.com/maps), (Accessed on 02/15/2021).

[36] M. Kim, D. Kotz, S. Kim, Extracting a mobility model from real user traces, IEEE INFOCOM, 2006.

[37] C. Zakaria, A. Trivedi, M. Chee, P. Shenoy, R. Balan, Analyzing the impact of covid-19 control policies on campus occupancy and mobility via passive wifi sensing, arXiv preprint arXiv:2005.12050(2020).