An Alternative Paradigm for Digital Exposure Notification: Detecting Super-spreading Events

Brian Thompson, Kyle Lemoi, Phil LaGambino, Molly McKnight

NIST Workshop on Challenges in Digital Proximity Detection in Pandemics
January 28, 2021
Overview

**Motivation**
- Much of COVID-19 spread is attributable to super-spreading events*, where aerosolized virus can accumulate and permeate a room due to poor ventilation
- Current GAEN/TC4TL paradigm does not capture most of these infections, e.g., because they are not close enough or index case does not have the app

**Idea**
- Send an exposure notification if the user spent a significant amount of time in the simultaneous presence of multiple others who later report a positive test
- Threshold based on an a posteriori estimate of likelihood the user got infected

**Properties**
- Works within the GAEN system but fundamentally changes configuration of the “attenuation” and “days” parameters and criteria for sending a notification
- Targets high-impact events: the more people, the greater the statistical power
- Linear dependence on adoption rate: works even if index case not using app
- Implicitly captures transmission factors GAEN cannot: biological, situational

*E.g., see Goyal et al. (2020)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Configuration under current paradigm</th>
<th>Configuration under new paradigm</th>
<th>Motivation/Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days</td>
<td># of days prior to positive test or symptom onset</td>
<td>High score if at most 2 days; based on infectiousness curve</td>
<td>High score if at least 3 days; based on incubation period</td>
<td>Current paradigm considers whether the person was infectious at the time of contact; new paradigm considers whether the person got infected then</td>
</tr>
<tr>
<td>Attenuation</td>
<td>Bluetooth signal attenuation; proxy for distance</td>
<td>High score if estimated to be within 6 feet</td>
<td>High score if estimated to be within 30 feet</td>
<td>Current paradigm considers whether an infectious person was in close contact; new paradigm considers whether people who later got infected were in the same room</td>
</tr>
<tr>
<td>Duration</td>
<td>Duration of contact</td>
<td>High score if at least 15 minutes</td>
<td>High score if at least 30 minutes</td>
<td>Current paradigm based on “plume” model of transmission during close contact; new paradigm considers “room” model where virus-laden aerosols accumulate in the air</td>
</tr>
<tr>
<td>Notification criteria</td>
<td>When to send a notification to the app user</td>
<td>“Too close for too long” contact with an infected person</td>
<td>Nearby multiple others who likely got infected around that time</td>
<td>Current paradigm considers whether the app user got infected by a specific person; new paradigm flags when the user was present at a likely super-spreading event</td>
</tr>
</tbody>
</table>
Example Scenario

- 80 people in a bar with a highly infectious person who does not have the app
Example Scenario

- 80 people in a bar with a highly infectious person who does not have the app
- 30 of the 80 susceptible people are app users (38% adoption rate)
Example Scenario

- 80 people in a bar with a highly infectious person who does not have the app
- 30 of the 80 susceptible people are app users (38% adoption rate)
- 40 people get infected, including 15 app users
Example Scenario

- 80 people in a bar with a highly infectious person who does not have the app
- 30 of the 80 susceptible people are app users (38% adoption rate)
- 40 people get infected, including 15 app users
- 25 of those infected eventually become symptomatic, including 10 app users
Example Scenario — Outcome

- **Current paradigm:** Nobody gets notified because the infectious person does not have the app and others were not infectious at the time.
Example Scenario — Outcome

- **Current paradigm**: Nobody gets notified because the infectious person does not have the app and others were not infectious at the time.
- **New paradigm**: After the first 3 symptomatic app users report a positive test via the app, the other **27 app users get notified, 12 of whom were infected**
## Example Scenario — Timeline

<table>
<thead>
<tr>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Day 8</td>
</tr>
<tr>
<td>Day 2</td>
<td>Day 3</td>
<td>Day 4</td>
<td>Day 5</td>
<td>Day 6</td>
<td>Day 7</td>
<td>Day 8</td>
</tr>
</tbody>
</table>

- **Day 0**: Super-spreading event
- **Day 1**: Day 1
- **Day 2**: Day 2
- **Day 3**: Day 3
- **Day 4**: Day 4
- **Day 5**: Day 5
- **Day 6**: Day 6
- **Day 7**: Day 7
- **Day 8**: Day 8

<table>
<thead>
<tr>
<th>Day 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated 26%* of symptomatic secondary cases will still be pre-symptomatic 7 days after getting infected — those are the people most likely to spread the virus the following weekend</td>
</tr>
</tbody>
</table>

*Based on maximum likelihood estimate of incubation period distribution by [Ferretti et al. (2020)](https://www.tandfonline.com/doi/full/10.1089/cvd.2020.0006)
Vision

- GAEN could provide users with several different notification types, each with appropriate messaging/recommendations:
  - Close contact with infected person, low likelihood of transmission
    - Inform user of exposure to raise awareness and motivate behavior change
  - Close contact with infected person, high likelihood of transmission
    - Inform user of exposure, recommend testing and/or self-quarantine
  - Attendance at likely super-spreading event
    - Inform user of exposure, recommend testing and/or self-quarantine
Next Steps

▪ **Assess the feasibility and efficacy of the new paradigm**
  – Can it notify enough people early enough to significantly reduce spread?
  – How does its precision-recall tradeoff compare to the current paradigm?

▪ **Estimate impact on population-level health outcomes and social burden**
  – Building on Oxford’s open-source agent-based simulation model to include super-spreading events, both observed and unobserved factors in viral transmission, and explicit representation of contact tracing mechanisms
  – Consider new paradigm as alternative and complement to current paradigm

▪ **Collaborate with others to reduce the spread of COVID-19**
  – Explore the technical and operational challenges involved in adapting GAEN for super-spreading event detection
  – Integrate our changes into Oxford’s open-source repository for public use
  – Help inform PHAs in determining messaging around exposure notifications
BACKUP
Efficacy of GAEN — Study Design

- **Simulate interactions** between infected and susceptible users of GAEN-based apps, then use the GAEN risk formula to infer whether transmission occurred
  - Transmission model integrates components from established and recent literature: infectiousness, emission, transport, and dose-response models
  - Consider both observed and unobserved biological and situational factors

- **Measure efficacy** as GAEN’s ability to achieve both high recall (detection rate of true transmission events) and high precision (low false alarm/notification rate)
  - Evaluate under both ideal and noisy conditions (e.g., attenuation -> distance)
Efficacy of GAEN — Study Conclusions

- **Fundamental limit on GAEN’s ability to accurately predict transmission**
  Detecting 50% of transmissions means that more than 90% of exposure notifications will be false alarms — even if GAEN perfectly infers distance, duration, and days — which may drive down app usage and compliance.

- **New solutions must break out of the current paradigm** if a more favorable precision-recall tradeoff is to be achieved.

---

**Precise formula using observable factors captured by GAEN**

**GAEN under noisy conditions, using recommended configuration**

**Manual contact tracing, assumes perfect memory and doesn’t work for strangers**
Efficacy of GAEN — Transmission Factors Considered

GAEN
- distance
- duration
- days after symptom onset

MITRE model
- incubation period
- peak viral load
- emission rate of aerosols during normal speech
- loudness of speech
- exercise level
## Efficacy of GAEN — Impact of Transmission Factors

<table>
<thead>
<tr>
<th>Transmission Factor</th>
<th>Used by GAEN?</th>
<th>How to infer</th>
<th>Accuracy</th>
<th>Impact on effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Who</strong> (person)</td>
<td>No</td>
<td>Only evident in retrospect if many infections result</td>
<td>Moderate with data on forward-traced infections, low otherwise</td>
<td>High — maximum amount of virus transmitted can vary by multiple orders of magnitude across individuals [Jacot et al. (2020), Asadi et al. (2020)]</td>
</tr>
<tr>
<td><strong>What</strong> (activity)</td>
<td>No</td>
<td>Assumption based on environment</td>
<td>Moderate with location data (bar vs. restaurant vs. train), low otherwise</td>
<td>High — amount of virus transmitted can vary by over two orders of magnitude from breathing to singing or shouting [Morawska et al. (2009), Asadi et al. (2020)]</td>
</tr>
<tr>
<td><strong>Where</strong> (environment)</td>
<td>No</td>
<td>Location data, RSSI/sensors for indoor/outdoor</td>
<td>High with location data, moderate with sensor data, low otherwise</td>
<td>Moderate — size of space is helpful for identifying super-spreading events, less important for tracing individual contact events except for indoor/outdoor</td>
</tr>
<tr>
<td><strong>When</strong> (time)</td>
<td>Yes</td>
<td>date of self-report or + test, proxy for symptom onset</td>
<td>Moderate if infected person is symptomatic, low otherwise</td>
<td>High — amount of viable virus transmitted decreases by multiple orders of magnitude if more than 1-3 days before or after the time of peak infectiousness</td>
</tr>
<tr>
<td><strong>How close/crowded</strong> (proximity)</td>
<td>Yes</td>
<td>Bluetooth RSSI and/or other sensor data</td>
<td>High with sensor data, moderate with RSSI only</td>
<td>Moderate — high accuracy is helpful for tracing individual contacts; moderate accuracy is probably sufficient for identifying super-spreading events</td>
</tr>
<tr>
<td><strong>How long</strong> (duration)</td>
<td>Yes</td>
<td>Timestamps of scans when ID appeared</td>
<td>High (within 5 minutes)</td>
<td>Moderate — amount of virus transmitted grows between linearly and quadratically with duration, depending on ventilation rate and other factors</td>
</tr>
</tbody>
</table>
Measuring Impact — Integration with Oxford model *

- **What is the Oxford model?**
  - Population-level agent-based simulation model capturing demographics, person-to-person interactions, virus spread, NPIs, and health outcomes
  - Developed at Oxford Big Data Institute, published in Science, open-source

- **Proposed augmentations to the Oxford model:**
  - **Group interactions** — many people in an enclosed space for a prolonged period of time (bars, public transit, etc.); “room” model vs. “plume” model; enables modeling of super-spreader events
  - **Contact tracing apps** — higher-fidelity model enables comparison of effectiveness of different exposure risk inference algorithms

Measuring Impact — Integration with Oxford model

- Bluetooth signal + sensor data
- contact inference model
- proximity
- duration
- environment
- virus transmission model
- exposure risk

Contact tracing app

Holistic model

- population-level demographics and intervention data
- agent-based simulation model
- impact metrics

© 2021 THE MITRE CORPORATION. ALL RIGHTS RESERVED.