Multiagent Epidemiologic Inference through Realtime Contact Tracing

Guni Sharon
Computer Science & Engineering
Texas A&M University
College Station, Texas
guni@tamu.edu

James Ault
Computer Science & Engineering
Texas A&M University
College Station, Texas
jault@tamu.edu

Peter Stone
Computer Science
The University of Texas at Austin,
Sony AI
Austin, Texas
pstone@cs.utexas.edu

Varun Kompella
Sony AI
varun.kompella@sony.com

Roberto Capobianco
Sony AI,
Sapienza University of Rome
Rome, Italy
roberto.capobianco@sony.com

ABSTRACT

This paper addresses an epidemiologic inference problem where, given realtime observation of test results, presence of symptoms, and physical contacts, the most likely infected individuals need to be inferred. The inference problem is modeled as a hidden Markov model where infection probabilities are updated at every time step and evolve between time steps. We suggest a unique inference approach that avoids storing the given observations explicitly. Theoretical justification for the proposed model is provided under specific simplifying assumptions. To complement these theoretical results, a comprehensive experimental study is performed using a custom-built agent-based simulator that models inter-agent contacts. The reported results show the effectiveness of the proposed inference model when considering more realistic scenarios – where the simplifying assumptions do not hold. When pairing the proposed inference model with a simple testing and quarantine policy, promising trends are obtained where the epidemic progression is significantly slowed down while quarantining a bounded number of individuals.

KEYWORDS

Statistical Inference, Hidden Markov Model, Computational Epidemiology, COVID-19

ACM Reference Format:
Guni Sharon, James Ault, Peter Stone, Varun Kompella, and Roberto Capobianco. 2021. Multiagent Epidemiologic Inference through Realtime Contact Tracing. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), London, UK, May 3–7, 2021*, IFAAMAS, 9 pages.

1 INTRODUCTION

When combating an ongoing epidemic, many authorities attempt to hinder or even halt the disease spread across the community. Several tools are utilized towards achieving this objective. In general, these can be divided into two classes, disease surveillance [12] and containment strategies [23]. Disease surveillance tools commonly use observations such as test results, presence of symptoms, and physical contact patterns along with epidemic models for anticipating the disease progression. Based on the surveillance projection, containment strategies can appropriately be applied. Containment strategies may include individual or collective quarantine orders, enforcing social distancing, closing certain public facilities such as schools or shops, etc. See Walensky and Del-Rio [23] for a survey on such strategies.

Effective combinations between surveillance tools and containment strategies were shown to lead to desirable outcomes w.r.t. epidemic progression [8, 14, 20]. The approaches proposed in these publications, however, rely on simple inference rules regarding individual infection likelihood. For instance, by ranking the infection likelihood based on a weighted combination of observed symptoms [10].

We present a novel approach for infection inference that reports individual infection probabilities per day over the entire population. We show how such an approach, when paired with a simple containment strategy, is highly successful in hindering the epidemic progression in a simulated community.

The proposed inference approach models the problem as a hidden Markov model [9], where every time step is defined by individual infection probabilities (the belief state) and a compressed representation of the physical contact history. At every state a set of observations is provided, namely, test results for a subset of the community, symptom presence for each individual, and physical contacts within the community. Based on these observations and given conditional probabilities (false-positive/negative test and symptoms rates), the individual infection probabilities are updated. Finally, the transition probabilities are defined by the probability of recovering (transitioning from being infected to recovered) per time step.

The inference procedure is justified using a set of simplifying assumptions. In order to compliment these theoretical results, we present a comprehensive experimental study using a custom-built

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, London, UK.

agent-based simulator. Our experimental study shows the effectiveness of the proposed approach in more realistic settings where the simplifying assumptions might be violated.

2 PROBLEM DEFINITION

We consider a scenario where a population, S, is exposed to an infectious disease. At every time step, t, a subset of the population, $I_t \subset S$, is sick. The median individual sickness period is known and denoted, d. We define an infection event for individual s_i at time t by $s_i \in I_{t+1}$ and $s_i \notin I_t$. Note that we use the term 'sick' in lieu of the term infected which is more common in the epidemiological literature. We do so in order to avoid confusion between an infection event and a state of being infected which is denoted 'sick'. Once an individual is infected, they will remain sick (and infectious to others) until they recover. Each individual can be tested at any time step to determine if they are sick. The test false positive and false negative rates are known and denoted F^+_{test} and F^-_{test} respectively. Every individual can also be showing symptoms (or not) at every time step. The probability of showing symptoms and not being sick is known and denoted F^+_{symp}. Similarly, the probability of showing **no** symptoms and being sick is known and denoted F^-_{symp}. We define the *Sickness Likelihood Inference* problem as follows.

Goal: compute a sickness belief state over the entire population. That is, compute a vector of probabilities, $B = \mathbb{R}^{|S|}$, where $B[i]$ represents the probability that individual i is currently sick (and infectious).

Input: The following observations are provided at every time step, t:

- Test results for a subset, of the population.
- The existence of symptoms for each individual.
- Contact graph as a symmetric matrix, $C_t = \mathbb{R}^{|S|\times|S|}$. $C_t[i, j]$, represents the transmission probability between individuals i and j during time step t.

In a real-world scenario some of these inputs might be unknown. For instance, some individuals might refuse to report existing symptoms or who they were in contact with. We consider such partial observability in the experimental study.

Desiderata: The proposed solution should **(a)** maximize prediction accuracy regarding sick individuals, and **(b)** avoid storing the full contact, testing, and symptoms history due to computational and memory limitations as well as privacy concerns.

3 RELATED WORK

Previous work [11, 21] attempted to estimate the individual sickness (for COVID-19) probabilities based on different features, namely, age, gender, presence of prior medical conditions, general feeling, and the symptoms fever, cough, shortness of breath, sore throat and loss of taste or smell. The observed correlations between the mentioned attributes and positively testing for COVID-19 were reported. These studies, however, did not consider physical contact tracing or inference over successive days.

Grushka et al. [10] considered similar symptom attributes along with a binary attribute designating contact with a confirmed case. Full contact history was not considered based on the following explanation, "The purpose of data security is similar to the task of testing and contact tracing organizations. The sheer amount of daily user activity in IT database systems prevents testing and logging every action". Grushka et al. showed that reported contact with a confirmed case was the dominant feature for determining the probability that an individual will be tested positive. Again, a Markov process was not assumed, and probabilities were not updated over time.

Another line of previous work [1, 4, 17–19] did model epidemic progression as a Markov process. However, such models assume full observability regarding the susceptible, infected, and recovered subgroups. The resulting statistical inference relates to the infection distribution for the entire population and not per individual.

4 SICKNESS LIKELIHOOD INFERENCE PROCEDURE

We address the sickness likelihood inference problem as a hidden Markov model [9] where the sickness probabilities define the state space. At each time step, the sickness probabilities are updated according to test results, symptoms, and contact observations. At each state transition the infection probabilities are updated according to a recovery probability.

Applying the Markov property to the affiliated state space seems counter-intuitive since updating the sickness probabilities for the current state impacts the infection probabilities in previous time steps. For example, if s_i tested positive today, the sickness probability for its previous physical contacts should be increased. In order to maintain the Markov property, we include a compressed representation of the contact history in each state. Such a compressed representation also complies with the desired feasibility and privacy restrictions that prohibit explicitly storing the contact history. Algorithm 1 details our proposed solution.

At every time step, three data-fields are updated and stored per individual, s_i. These are, sickness probability ($B[i]$), double decayed contact history ($C_{\gamma^2}[i]$), and triple decayed contact history ($C_{\gamma^3}[i]$). Each of the decayed contact histories is stored as a symmetric matrix with an entry per (unordered) pair of individuals in the community. We denote these matrices as double and triple decayed contact histories since they are decayed by factors γ^2 and γ^3 respectively. The intuition behind these decay factors is not straightforward. This need is derived from the mathematical representation of the problem under a set of assumptions that are discussed later in the "Theoretical Analysis" section. Both contact matrices are initialized as a zero matrix in Lines 4 and 5. The diagonals of both contact matrices are set as constant at $C_{\gamma^2}[i, i] = 1$ and $C_{\gamma^3}[i, i] = 0$. The specified data-fields define the state space in the affiliated HMM representation. A decay rate, γ, is set such that an initial probability of 1 would decay to 0.5 after d days (Line 3). This decay rate represents a recovery probability of 0.5 by the median infection period (d). Such a decay rate assumes a constant per day probability of recovering (see Assumption 1 in the "Theoretical Analysis" section). Line 16 updates both (double and triple) decayed contact history to include the contacts reported for the current time step (C_t). Recall that the entries of matrices C_{γ^2}, C_{γ^3}, and C_t represent probabilities. As a result, every entry is capped at 1.

Algorithm 1: Sickness Likelihood Inference

Input: daily contacts, test results, and symptoms report
Result: Daily individual sickness probabilities, B

1 **Initialization:**
2 Init belief state as a vector of probabilities: $B = \mathbb{R}^{|S|}$ where $\forall s_i \in S$, $B[i] = |I|/|S|$;
3 Set the decay rate based on the sickness median length (d): $\gamma = \sqrt[d]{0.5}$;
4 Init double decay contact history as a symmetric matrix of probabilities: $C_{\gamma^2} = \mathbb{R}^{|S|\times|S|}$;
5 Init triple decay contact history as a symmetric matrix of probabilities: $C_{\gamma^3} = \mathbb{R}^{|S|\times|S|}$;
6 **foreach** *step t, (day)* **do**
7 Decay the sickness probabilities: $B = \gamma B$;
8 Update contact matrices based on today's reported contacts (C_t): $C_{\gamma^2} = \gamma^2 C_{\gamma^2} + C_t$ and $C_{\gamma^3} = \gamma^3 C_{\gamma^3} + C_t$;
9 **foreach** *individual,* $s_i \in S$ **do**
10 **If** tested positive, **then:** $B[i] = 1 - (1 - B[i])\, F^{+}_{test}$;
11 **Else if** tested negative, **then:** $B[i] = B[i] F^{-}_{test}$;
12 **If** showing symptoms, **then:** $B[i] = 1 - (1 - B[i])\, F^{+}_{symp}$;
13 **Else** (no symptoms), **then:** $B[i] = B[i] F^{-}_{symp}$;
14 **end**
15 **foreach** *individual,* $s_i \in S$ **do**
16 $B_{next}[i] = 1 - \prod_j \left(1 - B[j]\left(C_{\gamma^2}[i,j] - B[i] C_{\gamma^3}[i,j]\right)\right)$;
17 **end**
18 Normalize B_{next} ;
19 $B = B_{next}$;
20 **end**

The individual sickness probabilities are updated based on reported test results and symptoms observations. For instance, if individual s_i tested positive at the current time step, we update its sickness probability to equal the complementary event for both not-being sick **and** falsely testing positive (Line 10). Note that this algorithm makes a simplifying assumption that test observation and symptoms observation are conditionally independent given sickness. If this assumption is violated, as suggested for COVID-19 [10], then lines 10-13 should specify unique and mutually excluding cases per outcome combination. For example, **if** s_i tested positive **and** is showing symptoms, then $B[i] \leftarrow 1 - (1 - B[i]) F^{++}_{test\&symp}$ where $F^{++}_{test\&symp}$ is the probability of not being sick when both test results and symptoms indicate sickness (false positive-positive rate). Similarly, $F^{+-}_{test\&symp}$, $F^{-+}_{test\&symp}$, and $F^{--}_{test\&symp}$ will need to be defined.

Next, sickness probabilities are updated according to both decayed contact matrices (Line 16). Specifically, Line 16 computes the probability that s_i was infected by some individual s_j over the past days and did not recover since. The derivation of this update formula is provided later in the "Theoretical Analysis" section. Setting the diagonals of the two contact matrices $C_{\gamma^2}, C_{\gamma^3}$ as, ones and zeros respectively results in self infection probability of 1 from the previous day. That is, unless recovered, a sick individual will remain sick. The reader can verify that for these values, $B[i](C_{\gamma^2}[i,i] - B[i]C_{\gamma^3}[i,i])$ results in $B[i]$ (Line 16). Note that, Lines 15-16 can, and should, be computed more efficiently using matrix operations. The iterative form is provided for ease of presentation.

Finally, the set of probabilities is normalized to fit the estimated disease spread in the community (Line 18). Specifically, B is scaled such that $\sum B = |\hat{I}|$ where $|\hat{I}|$ is the estimated number of sick individuals. We assume that I can be evaluated using positivity test rates and random serological/PCR tests.

5 THEORETICAL ANALYSIS

The following simplifying **assumptions** are used for justifying the update rule in Line 16 of Algorithm 1.

(1) A constant per day recovery probability $(1 - \gamma)$ for sick individuals.
(2) If s_i was infected on some day then it cannot get infected on subsequent days (events of infection are mutually exclusive over days).
(3) The probability that any individual was previously sick and recovered is practically zero.
(4) For any individual, s_i, the daily a priori infection probability is equal between past days.

The reader should note that, in many real-world scenarios, these assumptions are not guaranteed to hold. For example, evidence regarding the COVID-19 pandemic [16] do not support Assumption 1. Furthermore, Assumption 3 is mainly relevant during the initial stages of the outbreak. Nonetheless, the reader should keep in mind that the early stages of the outbreaks are exactly those where inference is most important, as it enables reducing the maximal number of concurrent active cases (the epidemic spread peak) by applying better confinement strategies.

Let p^t_i be the probability that individual s_i is sick on day t. Similarly, p^{t-k}_i is the probability the individual s_i was sick k days before t. Following Assumption 1, we get:

PROPOSITION 1. *The probability that s_j infected s_i on day $t-k$ and s_i did not recover since is:*

$$p^{t-k}_j C_{t-k}[i,j]\gamma^k \tag{1}$$

When considering Assumption 2, Proposition 1 must be updated to include the requirement that s_i was not already sick at day $t-k$.

PROPOSITION 2. *The probability that s_j infected s_i on day $t-k$ **and** s_i was not already sick **and** s_i did not recover since is:*

$$p^{t-k}_j C_{t-k}[i,j](1 - p^{t-k}_i)\gamma^k \tag{2}$$

Note that Proposition 2 does not take into account a case where s_i was infected and fully recovered before day $t-k$. Recall that such scenarios have a probability of 0 according to Assumption 3. Also note that Proposition 2 defines infection events that are mutually exclusive over days. It is well known that the probability that no

mutually exclusive event happens is one minus the sum of the events probabilities. Following Proposition 2, we can write the probability that individual i is currently sick as one minus the probability that no unrecovered infection occurred between s_i and any other individual, s_j, at any past day, k. And so:

Proposition 3. *The probability that individual i is currently (time step t) sick is:*

$$1 - \prod_j \left(1 - \sum_k p_j^{t-k} \gamma^k C_{t-k}[i,j](1 - p_i^{t-k})\right) \tag{3}$$

Lemma 1. *Assumptions 1-4 imply:*

$$p_i^{t-k} = p_i^t \gamma^k \tag{4}$$

Proof. Let p_i^t+ represent the event where s_i is sick at time step t. Similarly, p_i^t- represent the event where s_i is **not** sick at time step t. Let $P(A|B)$ be the conditional probability of event A given event B.

$$\begin{aligned} p_i^{t-k} &=^{(1)} p_i^t \cdot P(p_i^{t-k} + | p_i^t +) + (1 - p_i^t)(p_i^{t-k} + | p_i^t -) \\ &=^{(2)} p_i^t \cdot P(p_i^{t-k} + | p_i^t +) \\ &=^{(3)} p_i^t \cdot P(p_i^t + | p_i^{t-k} +) =^{(4)} p_i^t \gamma^k \end{aligned}$$

$=^{(1)}$ by definition.
$=^{(2)}$ follows from Assumption 3.
$=^{(3)}$ follows from Bayes Theorem and Assumption 4.
$=^{(4)}$ follows from Assumption 1. □

Combining Proposition 3 with Lemma 1, we get:

Proposition 4.

$$\begin{aligned} p_i^{t+1} &= 1 - \prod_j \left(1 - \sum_k p_j^t \gamma^k \gamma^k C_{t-k}[i,j](1 - p_i^t \gamma^k)\right) \\ &= 1 - \prod_j \left(1 - p_j^t \left(\sum_k \left(\gamma^{2k} C_{t-k}[i,j]\right) - p_i^t \sum_k \left(\gamma^{3k} C_{t-k}[i,j]\right)\right)\right) \end{aligned} \tag{5}$$

$\sum_k \gamma^{2k} C_{t-k}[i,j]$ is an exponentially decayed moving average that is stored as C_{γ^2} in Algorithm 1. That is, there is no need for explicitly storing the full contact history. The same goes for $\sum_k \gamma^{3k} C_{t-k}[i,j]$ that is stored as C_{γ^3}.

The reader can verify that Line 16 from Algorithm 1 follows from Equation 5 in Proposition 4.

6 EXPERIMENTAL STUDY

To complement our theoretical analysis, we evaluate the effectiveness of the proposed approach via experiments in a custom-built agent-based pandemic simulator. Although some of the simulator's design have been motivated by the COVID-19 pandemic, we expect that our main experimental results are generally applicable. Note that in the simulator, the simplifying assumptions upon which the theoretical analysis relied do not hold. Namely, the recovery probability is not constant (in contrast to Assumption 1), but rather a function of the sickness duration and individual attributes; the probability that any individual was previously sick and recovered grows as time progresses (in contrast to Assumption 3); and for any individual, the daily a priori infection probability changes as a function of contact with sick individuals (in contrast to Assumption 4).

Our empirical study addresses the following questions.

(1) Can the proposed inference approach proactively identify sick individuals better than existing approaches?
(2) When comparing to existing approaches, can the proposed inference approach reduce epidemic progression when combined with a simple testing and quarantine policy?
(3) How do different levels of observability regarding contact tracing affect the efficiency of the proposed approach?

The reported results support the following answers: yes, yes, and better contact tracing leads to better inference.

6.1 Experimental Settings

Our experiments are conducted in a novel open-source agent-based simulator [15], written in python, that models the interactions between individual people at specific locations within a community. The simulator was developed with input from leading epidemiologists. It is fully configurable to allow modeling of the demographics of a specific community, as well as the number of grocery stores, schools, and other businesses. The simulator follows previous work [13] which suggested social mirror identities of daily-contact networks for purposes of performing epidemiological simulations. On top of that, the simulator includes configurable parameters to model the degree to which each individual within the population complies to social distancing guidelines. It also models other known properties of COVID-19 spread, such as the facts that some people are asymptomatic or spread the virus while presymptomatic [3], and that some infected individuals barely spread the virus, while others spread it widely [22]. Crucially for this paper, the simulator models testing, including false positive and false negative tests, and allows for contact tracing and subsequent isolation and quarantining.

From a high-level perspective, the simulator models a population of individuals who are assigned to houses. The simulator advances in 24 steps per day (each representing 1 hour). During each step, individuals execute a stochastic behavior, which can be simplistically summarized as follows: they stay home (at night or if retired), go to school (if minor), or go to work at one of several types of business places (if of working age). People also go to a local grocery store once a week, and a hair salon roughly once a month. In the evenings, people occasionally go to social gatherings.

The spread of the virus is modeled using a standard SEIR model [2] represented as a finite state machine such that *susceptible* individuals probabilistically transition to the *exposed* state depending on how many sick individuals they interacted with during the day. The probability of this transition is also governed by whether individuals are following social distancing measures, individuals' spread rates (drawn from a distribution), and the contact rates associated with different locations (for example to represent that the virus is more likely to spread at a hair salon than a grocery store, due to the different type of contacts). From the *exposed* state, individuals can transition to *infected-asymptomatic* or *infected-symptomatic*, which can evolve to *infected-critical*. People in the infected-critical state go to a hospital if there is sufficient capacity. Infected individuals

can transition to *recovered*; infected-critical individuals can also transition to a *deceased* state.

To contain the spread of the virus, available government actions include limiting the size of social gatherings, limiting the capacity or hours of certain businesses, and enacting social distancing measures (which are obeyed probabilistically). Most pertinent to this paper, the government can test symptomatic and asymptomatic individuals at different rates and can implement contact tracing to decide which individuals to quarantine. Quarantined individuals refrain from going to school, work, businesses, or social gatherings; they are completely isolated for a duration of 14 days. Isolated individuals are not considered for being tested. Contact tracing can be implemented at three levels.

(1) **Passive tracing** - home and work/school addresses are known.
(2) ***x*% tracing** - $x\%$ of the population are actively traced (e.g., by a relevant cellphone app). When two such individuals occupy the same building, a contact event is registered along with the contact duration.
(3) **Active tracing** - denotes 100% tracing.

For passive tracing, the daily reported contacts $C_t[i, j]$ is set to equal 0 and then +0.5 if s_i, s_j shared the same house and +0.1 for sharing the same school/work place. For active tracing, $C_t[i, j]$ was set proportional to the contact length between s_i and s_j in hours divided by 24, i.e., the fraction of time steps that they were in the same location. For $x\%$ tracing, $C_t[i, j]$ was set according to the active tracing rule if both s_i and s_j are actively traced; otherwise it was set according to the passive tracing rule.

For our experiments we followed the default simulator settings. That is, a small-town scenario of 1000 individuals with an age distribution based on U.S. demographics. Physical locations throughout the town are defined by types, assigned number of employees, and a visitors bound. The town consist of 300 homes to which all 1000 people were assigned. Adults are assigned to work at one of the locations occupying an employee slot. Minors are assigned as visitors to schools. The locations were configured using the following settings. 6 Grocery stores, each with 3 employees and up to 20 concurrent visitors (customers). 30 offices, each with 35 employees and no visitors. 8 school sections (classes), each with 5 employees and up to 37 concurrent visitors (students). 2 hospital sections each with 15 employees and up to 5 concurrent visitors (patients). 6 retail stores, each with 3 employees and up to 20 concurrent visitors (customers). 4 barber shops each with 3 employees and up to 5 concurrent visitors (customers).

The simulator's hyper-parameters were set based on data reported for the COVID-19 pandemic. Decay rate is set based on a median sickness length of 5 days. False positive F^+_{test} and false negative F^-_{test} test rates were set to 0.0165 and 0.0332 respectively following average outcomes reported for 3,524 PCR tests [5] in Brazil.

The false positive symptomatic rate (showing symptoms yet are not sick), F^+_{symp}, was set to 0.0655 following the average workdays loss ratio (pre COVID-19) due to sickness in Japan [7]. The false negative symptomatic rate (sick but asymptomatic), F^-_{symp}, was set to 0.6 following the "Current Best Estimate" (September-8, 2020) of the US Centers for Disease Control and Prevention [6]. Further, it was assumed that not all of those whom are symptomatic would report their status. In our experiments 30% of the population dutifully reported their symptomatic status.

6.2 Baseline

Our baseline for comparison follows the risk score method presented by Grushka et al. [10] for ranking COVID-19 positive testing probabilities. According to the reported correlations, the following ranking is inferred (lower rank number implies higher sickness probability).

(1) Individuals who tested positive.
(2) Individuals who were in contact with a confirmed case during the last 7 days (exposed) and are showing symptoms.
(3) Exposed individuals.
(4) Individuals showing symptoms.
(5) All others.

Each individual in the community is assigned to the highest rank (where 1 is the highest and 5 the lowest) that fits their status. That is, if an individual is both showing symptoms and was exposed, it is assigned to rank 2. We consider four unique baselines that are derived from the above categories.

- **Exposure+symptoms**, sickness probability ordering follows the above ranking as suggested by Grushka et al. i.e., $1 \prec 2 \prec 3 \prec 4 \prec 5$.
- **Exposure**, sickness probability ordering follows the ranking $1 \prec (2 \equiv 3) \prec 4 \prec 5$.
- **Symptomatic**, sickness probability ordering follows the ranking $1 \prec (2 \equiv 4) \prec 3 \prec 5$.
- **Random,** sickness probability ordering follows no ranking (excluding positive test results), i.e., $1 \prec 2 \equiv 3 \equiv 4 \equiv 5$.

When querying for the n most probable sick individuals, each baseline method returns the n highest ranked individuals while breaking ties randomly.

6.3 Inference Accuracy

The first set of experiments aims to address research question #1: Can the proposed inference approach proactively identify sick individuals better than the baseline methods?

In order to allow fair comparison between the baselines and the inference approach, no quarantine operations were used and the testing policy was purely random (sampling 1% of the population each day). The baselines and the inference approach were provided the exact same information (test results and contact tracing) within the exact same run. Doing so allowed us to compare how accurately each approach managed to guess the subset of sick individuals. It is important to note that the compared prediction approaches did not influence the simulation progression in any way (they simply observed and reported predictions).

Let I_t be the set of actively sick individuals at day t. Let $S^n(B_t)$ be the set of n individuals with the highest infection probability according to belief state B_t. Define hit-ratio for day t as $\frac{|I_t \cap S^n(B_t)|}{|I_t|}$ with $n = |I_t|$.

Figure 1 presents the hit-ratio over time for our inference approach and the four baseline approaches. Three scenarios are considered with regards to contact tracing, namely, active, 50%, and

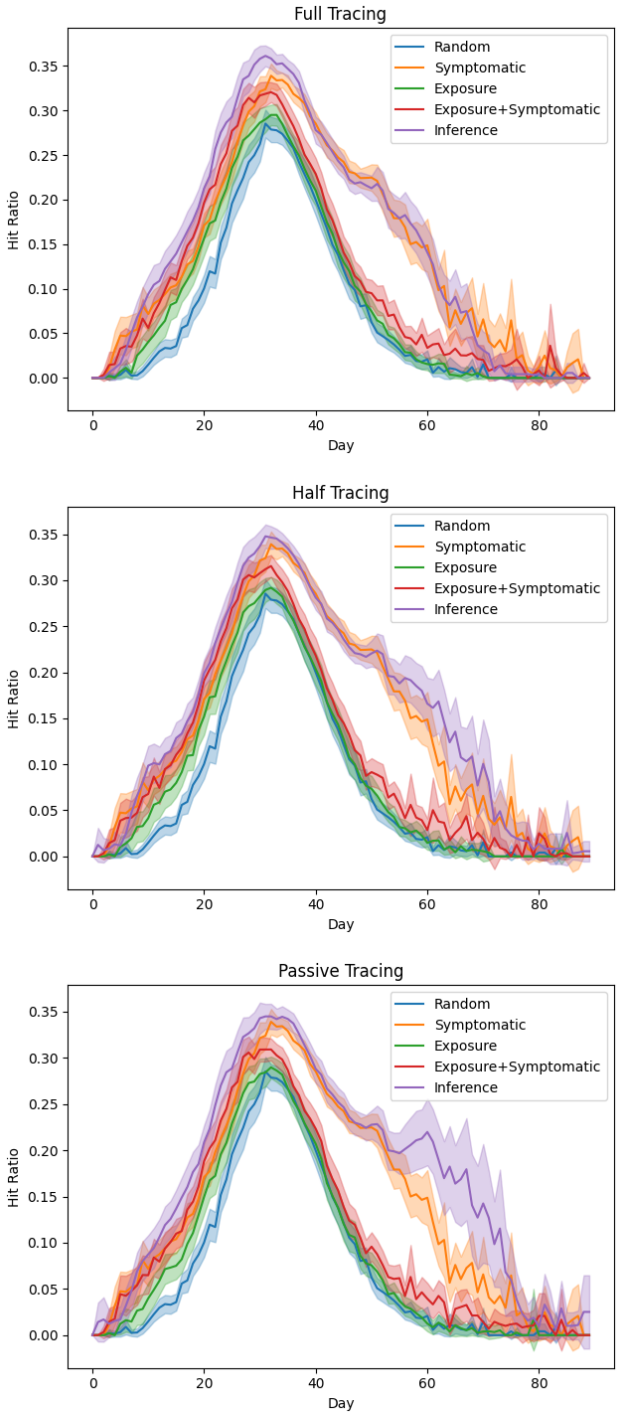

Figure 1: Hit-ratio as a function of time. Shaded areas represent 95% confidence interval over 60 trials.

passive. Note that over different contact tracing scenarios (between the subfigures) the 'Random' curve is showing the exact same trend and this is also true for the 'Symptomatic' curve. Neither the 'Random' or 'Symptomatic' approaches consider contact tracing, so all the scenarios are the same from their perspective. However, several trends regarding our proposed inference approach can be observed.

First, in all scenarios and all time steps the proposed inference approach performs at least as well as all the baseline approaches (equal or higher hit-ratios). The advantage reported for our inference approach is not a clear winner and, in many times steps, its advantage is not statistically significant over all the baseline methods. Nonetheless, even small advantages in prediction accuracy can result in a significant advantage once combined with a suitable testing and quarantining policy, as shown in the next section.

Another observed trend is the significant advantage for the 'Symptomatic' and 'Inference' curves after the 'Hit Ratio' peak (around day 37). Blindly prioritizing exposed individuals ('Exposure', 'Exposure+Symptomatic') does not perform well in such cases due to "herd immunity", i.e., most contacts are with recovered individuals who do not get infected. However, the reader should note that post-peak sickness prediction has little to no impact when aiming to "flatten the curve", i.e., reduce the peak's magnitude with respect to number of sick individuals.

6.4 Impact on Test and Quarantine Policies

The second set of experiments aims to address research question #2: Can the proposed inference approach reduce epidemic progression when combined with an appropriate testing and quarantine policy?

A simple testing policy was implemented where tests are assigned to the most probable sick individuals (excluding those previously teased positive) using either our inference approach or the baseline approaches. The number of tests per day was set to 1% of the population or 10 in total. A complimentary quarantine policy was implemented where the most probable sick individuals were isolated and had no active contacts in successive days. For the *Random* baseline, those tested positive were isolated. Isolation lasts 14 days after which normal behavior is resumed. In order to allow a fair comparison, the number of individuals that are sent to be isolated per day is capped at 2% of the simulated population (20 individuals). Note that more than 20 individuals can be isolated simultaneously if they were initially sent to isolation on different days.

Figure 2 presents the number of sick individuals over time for our inference approach and the four baseline approaches. As in Figure1, three scenarios are considered and presented regarding the contact tracing, namely, active, 50%, and passive. When full contact tracing is considered, our inference approach shows a significant advantage over the baseline approaches (the error intervals are not overlapping). The advantage is apparent when aiming to "flatten the curve", that is, when seeking to reduce the maximal number of concurrent sick individuals. On the other hand, when passive tracing is considered, our inference approach shows little to no advantage. For half tracing, our inference approach shows a significant advantage however it is not as prominent as in the full tracing case.

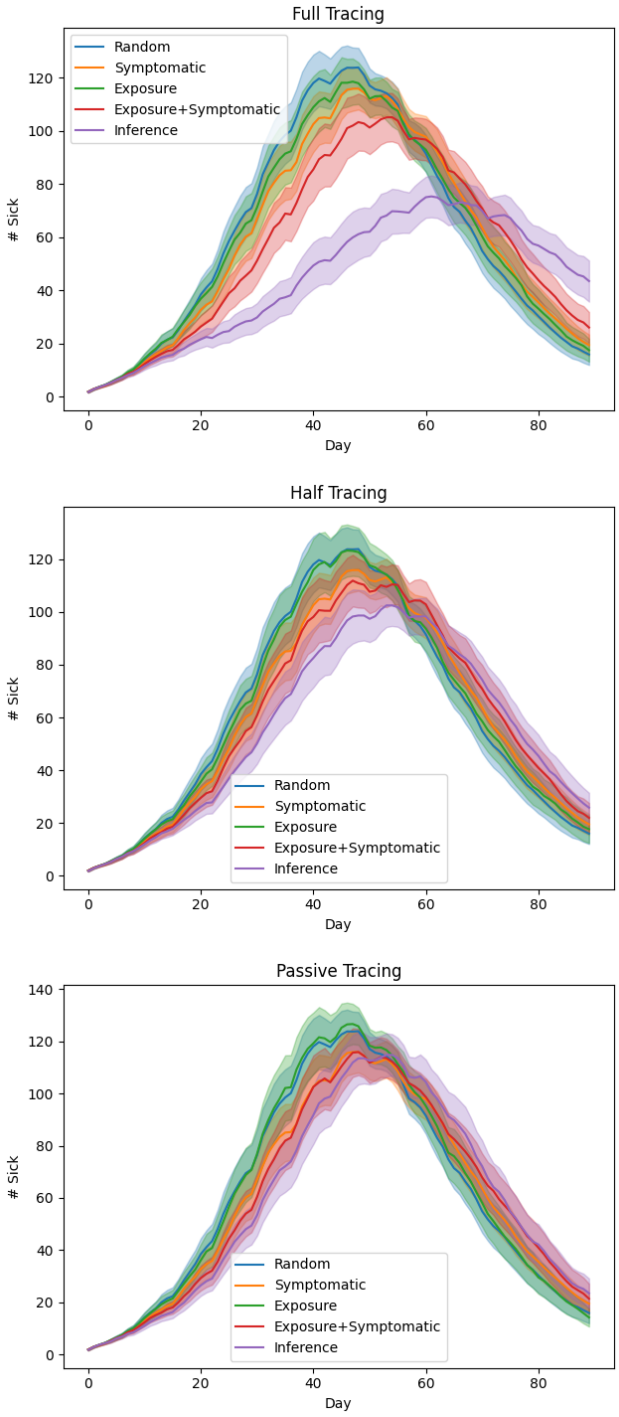

Figure 2: Number of actively sick individuals as a function of time. Shaded areas represent 95% confidence interval over 60 trials.

6.5 Sensitivity Analysis

Next we examine the inference procedure's performance sensitivity to the number of available tests and available quarantine orders per day. This set of experiment is motivated by the assumption that extensive quarantine orders cripple the economy and should, thus, be minimized/capped by the authorities.

Table 1 compares the best performing baseline approach (Exposure+Symptomatic) with our inference approach for different tests and quarantine caps. The table entries report the maximal number of concurrent sick individuals over the entire simulation run for the baseline as well as the ratio of improvement over the baseline for our inference approach ("improvement ratio"). Results are averaged over 36 runs with similar random seeds for both the baseline and our inference. 95% statistically significant, using a paired t-test, is denoted by an asterisk.

We observe that, in general, more testing yields a greater advantage to our inference approach. This trend, however, is not apparent when the baseline method can halt the disease progression (values ≤ 10). In such cases the baseline is sufficient to stop the disease progression, meaning that it does not spread over the entire community. As a result, the advantage from our (better) inference method is limited. This phenomenon is more apparent in the 4% cap quarantine/day. Such an aggressive quarantining policy results in slightly more than 50% of the population concurrently isolated (as opposed to 14% for a 1% cap). Consequently, the epidemic dies out in most runs. Nonetheless, our inference approach can still provide significant advantage when paired with active tracing by stopping the epidemic progression earlier.

The reported results for passive tracing 2% quarantine/day and 0.5%, 1% test caps suggest that our inference approach can perform worse than the baseline (improvement ratio < 1). However, the reader should note that these results are not statistically significant. As an example, consider 2% quarantine/day and 1% test caps. Table 1 places the baseline at a factor of 0.88 from the inference approach yet in Figure 2 the inference approach seems to perform on par or even slightly better than the baseline for the same scenario. Further note that the results in Figure 2 are expected to be more accurate as they are based on more trials, 60, as opposed to 36 for Table 1.

6.6 Conclusions

Several general conclusions are drawn from our empirical study.

- The proposed inference approach can better predict the sick set of individuals prior to the epidemic peak when compared to the baseline approaches.
- When paired with a simple testing and quarantine policy, the proposed inference approach can significantly reduce the number of concurrent active sick cases (flatten the curve). This advantage can reach a factor of 58/17 = 3.41 (for testing and quarantine caps of 4% and 2% respectively).
- In all of our experiments, other than the aggressive 4% quarantine/day policy, incorporating the inference with Active contact tracing resulted in significant improvement over half tracing, which significantly improved on passive tracing.

	Active tracing					Passive tracing				
Test cap (%)	0.5	1	2	4	10	0.5	1	2	4	10
	cap 1% quarantine/day									
Baseline	197	188	165	161	64	199	194	185	160	108
Improvement ratio	1.09	*1.11	1.11	*1.44	*1.88	1.05	1.03	*1.08	1.15	*1.86
	cap 2% quarantine/day									
Baseline	100	97	75	58	10	104	105	106	86	25
Improvement ratio	*1.28	*1.54	*1.79	*3.41	1.25	0.88	0.95	*1.18	*1.39	1.67
	cap 4% quarantine/day									
Baseline	10	8	7	8	8	10	10	7	9	6
Improvement ratio	*1.11	*1.14	1.00	1.00	*1.14	1.11	1.43	1.00	1.13	1.20

Table 1: Maximal number of sick individuals in a single day for different caps of tests and quarantine orders per day. Asterisk in front of a value denotes a 95% statistically significant difference over 36 trials.

- When applying an aggressive quarantine policy (4% quarantine/day), identifying sick individuals has little to no advantage as most of the population ends up isolated and the epidemic dies out.

It is important to note that these conclusions are relevant to the simulator utilized in this study. Discrepancies between the simulated model and the real-world might influence these general conclusions. An important direction for future work is to examine the extent to which the above conclusions hold in other simulation models. Ultimately, the reported trends ought to be examined in a real-world scenario.

7 SUMMARY

We present and justify an inference approach for detecting infected individuals during an epidemic outbreak. The proposed inference method makes use of observed infection symptoms and infection test results that are applied to a subset of the community. We show that, under a set of simplifying assumptions, the statistical inference can be accurately performed without the need to store the entire history of test results and symptoms presence. We report a comprehensive experimental study in a custom-built agent-based simulator that considers inter-agent contacts. The reported results suggests that our proposed inference approach is beneficial for more realistic scenario where the simplifying assumptions do not hold. Our study further suggests that the more detailed the contact tracing is, the better the inference performs when paired with a straightforward testing and quarantine policy.

ACKNOWLEDGMENTS

A portion of this work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CPS-1739964, IIS-1724157, NRI-1925082), ONR (N00014-18-2243), FLI (RFP2-000), ARO (W911NF-19-2-0333), DARPA, Lockheed Martin, GM, and Bosch. Peter Stone serves as the Executive Director of Sony AI America and receives financial compensation for this work. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

REFERENCES

[1] E Almaraz and A Gómez-Corral. 2018. On SIR-models with Markov-modulated events: Length of an outbreak, total size of the epidemic and number of secondary infections. *Discrete & Continuous Dynamical Systems-B* 23, 6 (2018), 2153–2176.
[2] Joan L Aron and Ira B Schwartz. 1984. Seasonality and period-doubling bifurcations in an epidemic model. *Journal of theoretical biology* 110, 4 (1984), 665–679.
[3] Yan Bai, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen, and Meiyun Wang. 2020. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA* 323, 14 (04 2020), 1406–1407. https://doi.org/10.1001/jama.2020.2565
[4] Tom Britton and Etienne Pardoux. 2019. *Chapter 2 Inference for Markov Chain Epidemic Models.* Springer International Publishing, Cham, 343–362. https://doi.org/10.1007/978-3-030-30900-8_13
[5] Rodolfo Castro, Paula M Luz, Mayumi D Wakimoto, Valdilea G Veloso, Beatriz Grinsztejn, and Hugo Perazzo. 2020. COVID-19: a meta-analysis of diagnostic test accuracy of commercial assays registered in Brazil. *The Brazilian Journal of Infectious Diseases* (2020).
[6] CDC. 2020. COVID-19 Pandemic Planning Scenarios. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html. Accessed: 2020-09-8.
[7] Odgerel Chimed-Ochir, Tomohisa Nagata, Masako Nagata, Shigeyuki Kajiki, Koji Mori, and Yoshihisa Fujino. 2019. Potential Work Time Lost Due to Sickness Absence and Presence Among Japanese Workers. *Journal of occupational and environmental medicine* 61, 8 (2019), 682–688.
[8] Jon Cohen and Kai Kupferschmidt. 2020. Countries test tactics in 'war'against COVID-19.
[9] Sean R Eddy. 2004. What is a hidden Markov model? *Nature biotechnology* 22, 10 (2004), 1315–1316.
[10] Hagit Grushka-Cohen, Raphael Cohen, Bracha Shapira, Jacob Moran-Gilad, and Lior Rokach. 2020. A framework for optimizing COVID-19 testing policy using a Multi Armed Bandit approach. *arXiv preprint arXiv:2007.14805* (2020).
[11] Daniel F Gudbjartsson, Agnar Helgason, Hakon Jonsson, Olafur T Magnusson, Pall Melsted, Gudmundur L Norddahl, Jona Saemundsdottir, Asgeir Sigurdsson, Patrick Sulem, Arna B Agustsdottir, et al. 2020. Spread of SARS-CoV-2 in the Icelandic population. *New England Journal of Medicine* (2020).
[12] Jo EB Halliday, Katie Hampson, Nick Hanley, Tiziana Lembo, Joanne P Sharp, Daniel T Haydon, and Sarah Cleaveland. 2017. Driving improvements in emerging disease surveillance through locally relevant capacity strengthening. *Science* 357, 6347 (2017), 146–148.
[13] Chung-Yuan Huang, Chuen-Tsai Sun, Ji-lung Hsieh, Yi-Ming Chen, and Holin Lin. 2005. A Novel Small-World Model: Using Social Mirror Identities for Epidemic Simulations. *Simulation* 81 (10 2005), 671–699. https://doi.org/10.1177/0037549705061519
[14] Edward H Kaplan and Howard P Forman. 2020. Logistics of aggressive community screening for coronavirus 2019. In *JAMA Health Forum*, Vol. 1. American Medical Association, e200565–e200565.
[15] Varun Kompella, Roberto Capobianco, Stacy Jong, Jonathan Browne, Spencer Fox, Lauren Meyers, Peter Wurman, and Peter Stone. 2020. Reinforcement Learning for Optimization of COVID-19 Mitigation policies. *arXiv preprint arXiv:2010.10560* (2020).
[16] Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine* 172, 9 (2020), 577–582.

[17] Claude Lefèvre and Matthieu Simon. 2019. SIR-Type Epidemic Models as Block-Structured Markov Processes. *Methodology and Computing in Applied Probability* (2019), 1–21.

[18] Michael Li, Jonathan Dushoff, and Benjamin M Bolker. 2018. Fitting mechanistic epidemic models to data: A comparison of simple Markov chain Monte Carlo approaches. *Statistical Methods in Medical Research* 27, 7 (2018), 1956–1967. https://doi.org/10.1177/0962280217747054 arXiv:https://doi.org/10.1177/0962280217747054 PMID: 29846150.

[19] Yohana Maiga Marwa, Samuel Mwalili, and Isambi Sailon Mbalawata. 2018. Markov chain Monte Carlo analysis of cholera epidemic. *J. Math. Comput. Sci.* 8, 5 (2018), 584–610.

[20] Marcel Salathé, Christian L Althaus, Richard Neher, Silvia Stringhini, Emma Hodcroft, Jacques Fellay, Marcel Zwahlen, Gabriela Senti, Manuel Battegay, Annelies Wilder-Smith, et al. 2020. COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation. *Swiss medical weekly* 150, 11-12 (2020), w20225.

[21] Saar Shoer, Tal Karady, Ayya Keshet, Smadar Shilo, Hagai Rossman, Amir Gavrieli, Tomer Meir, Amit Lavon, Dmitry Kolobkov, Iris Kalka, et al. 2020. Who should we test for COVID-19? A triage model built from national symptom surveys. *medRxiv* (2020).

[22] Richard A Stein. 2011. Super-spreaders in infectious diseases. *International Journal of Infectious Diseases* 15, 8 (2011), e510–e513.

[23] Rochelle P. Walensky and Carlos del Rio. 2020. From Mitigation to Containment of the COVID-19 Pandemic: Putting the SARS-CoV-2 Genie Back in the Bottle. *JAMA* 323, 19 (05 2020), 1889–1890. https://doi.org/10.1001/jama.2020.6572