Group Testing on a Network

Arlei Silva, Ambuj Singh
University of California, Santa Barbara
{arlei,ambuj}@cs.ucsb.edu

Abstract

Group testing—where multiple samples are tested together using a single test kit and individual tests are performed only for samples in positive groups—is a popular strategy to optimize the use of testing resources. We investigate how to effectively group samples for testing based on a transmission network. We formalize the group assembling problem as a graph partitioning problem, where the goal is to minimize the expected number of tests needed to screen the entire network. The problem is shown to be computationally hard and thus we focus on designing effective heuristics for it. Using realistic epidemic models on real contact networks, we show that our approaches save up to 33% of resources—compared to the best baseline—at 4% prevalence, are still effective at higher prevalence, and are robust to missing transmission data.

Introduction

Cost-efficient, timely and massive testing is a key challenge in many domains. Population-scale testing during a pandemic is likely the first scenario with such a challenge that comes to mind (Taipale, Romer, and Linnarsson 2020). However, one could argue that testing plays an even bigger role in manufacturing, for products ranging from resistors to drugs (Du, Hwang, and Hwang 2000).

A popular approach to increase testing capacity is *group testing* (Dorfman 1943), which was originally proposed to test American soldiers for syphilis during WWII. The idea was to pool multiple individuals into groups and test each group using a single test kit. A negative test for the group would indicate that none of the group members had the pathogen. In case the group test returned positive, each individual in the corresponding group would be tested alone. This strategy is provably effective when the prevalence (or frequency) of the pathogen is low, as most groups will be negative. Group sizes have to be optimized according to the prevalence—large groups might be uninformative. In practice, larger groups also make it harder to detect the pathogen at the group level due to dilution (Ghosh et al. 2020). Today, group testing is also applied to problems outside of healthcare, such as in computer networks and production lines (Du, Hwang, and Hwang 2000). There are more recent approaches for group testing (Aldridge 2020; Broder and Kumar 2020) but Dorfman's design remains attractive due to its easy implementation. In particular, higher savings can be achieved with more rounds and stages of testing at the cost of more intricate procedures and/or longer wait times.

Dorfman assumed that groups were assembled at random. Later studies have identified the importance of the grouping step to minimize the use of testing resources. For instance, when screening for a pathogen, a common guideline is to put together family members and other closely connected subgroups (Fang et al. 2020; Augenblick et al. 2020). However, we know that people tend to maintain a complex network of contacts (Newman 2002; Salathé et al. 2010) and thus effectively assembling the groups becomes a challenge.

This paper investigates the problem of designing groups for testing based on a transmission network—e.g., discovered based on contact tracing. Individual tests are performed only for samples belonging to positive groups. We formalize our problem as a graph partitioning problem, where the goal is to select groups that will likely minimize the number of tests required to screen the entire network. Figure 1 illustrates how groups are assembled and tested.

COVID-19. A major motivation for this work is the coronavirus disease (COVID-19), which has become one of the worst healthcare crises in history. While many experts agree that a vaccine is the only effective long-term solution, a combination of social-distancing, testing and contact-tracing has been advocated as a viable alternative to control the spread of the virus (Taipale, Romer, and Linnarsson 2020). However, most countries have failed to provide large-scale and rapid testing for SARS-COVID-2 (the virus), as cases have increased faster than the testing infrastructure. The most widely used testing protocol, the qRT-PCR (Corman et al. 2020), costs approximately $100 per test and requires: (1) swab collection, (2) RNA purification, and (3) reverse transcription and quantitative PCR. Steps 2 and 3 are performed in a lab, requiring a trained technician, reagents, and specialized equipment. Therefore, there is a significant effort by researchers and healthcare providers to speed-up and decrease the cost of testing for SARS-COVID-2. Several countries are using group testing with this purpose.[1][2]

[1] https://www.scientificamerican.com/article/coronavirus-test-shortages-trigger-a-new-strategy-group-screening2/

[2] https://www.nytimes.com/2020/05/07/opinion/coronavirus-group-testing.html

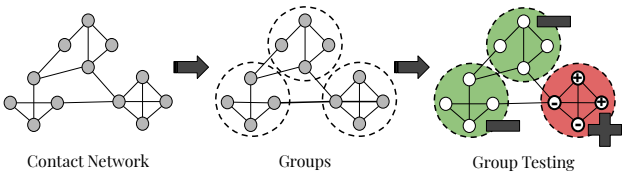


Figure 1: Group testing on a network. Groups are assembled based on a transmission network. Tests are first performed for groups and nodes are tested if their group is positive. This network can be screened with 7 tests, instead of 12.

We summarize the contributions of this paper as follows:

- We introduce the problem of group testing on a network, formulate it as a graph partitioning problem, and provide a characterization of its computational hardness.
- We propose efficient heuristics for the problem. Topology-based approaches only assume knowledge of the transmission network, while sampling-based ones minimize the expected number of tests over samples from the transmission process.
- We evaluate our heuristics using simulations of epidemics on real contact networks. Results show that the proposed solutions: (a) lead to savings in testing resources of up to 33% compared to the best baseline for a 4% prevalence; (b) are still effective when prevalence values reach 32%; and (c) are robust to missing transmission links.

Background

We describe processes on networks and Dorfman's design.

Processes on Networks

We start formalizing the notion of a transmission network.

Definition 1. ***Transmission network:*** *Graph $G(V, E, W)$ where V is the set of nodes and E is the set of edges. Edge weights $W : E \to \mathbb{R}$ are such that $w_{u,v} = W(u, v)$.*

An example of a transmission network is shown in Figure 1. There is an extensive literature on network processes and we refer to (Barrat, Barthelemy, and Vespignani 2008) for an overview. Here, we assume a generic process with parameters θ that can be sampled from a distribution $\mathcal{P}$—see Problem Definition for details. The transmission process will also define the role played by edge weights W in the transmission network. As an example, the network SIR process (Kiss, Miller, and Simon 2017; Eubank et al. 2004; Keeling and Eames 2005) is defined as follows.

Network Susceptible, Infected, Recovered (SIR) Process ($\theta = [\tau, \gamma, N]$): At any discrete time $t \in [1, T]$, each vertex $v \in V$ can be in one of the following states/sets: S (Susceptible), I (Infected), and R (Recovered). At $t = 0$, $S = V - N$ and $I = N$ and $R = \emptyset$, where $N \subseteq V$ is the set of initial infections. A vertex v moves from S to I with probability $\tau \times w_{u,v}$ after one of its neighbors $u \in I$. Vertices move from I to R a probability γ after infection.

We will apply SIR simulations on synthetic and real transmission networks in our experiments.

Dorfman's Group Testing

We focus on the two-stage design originally proposed by Dorfman, where individual tests are performed in the second stage (Dorfman 1943). Algorithm 1 formalizes the group testing procedure for a given group C_i. It returns a result (true or false) for each member of C_i while using a varying number of tests ranging from 1 to $|C_i| + 1$.

Algorithm 1 Dorfman's Group Test

Require: Group C_i
Ensure: Test results $R_v, \forall v \in C_i$
1: Apply test to group C_i
2: **if** group test is **positive then**
3: **for** member $v \in C_i$ **do**
4: $R_v \leftarrow$ individual test for v
5: **end for**
6: **else**
7: $R_v \leftarrow$ **negative**, $\forall v \in C_i$
8: **end if**

A group test (line 1) returns positive if at least one of the members of C_i is positive. Notice that this approach is fully accurate given our assumption that tests are noiseless (see further discussion in the Related Work). Our work is focused on the problem of assembling the groups for testing.

Problem Definition

The Group Testing on a Network (GTN) problem consists of partitioning the set of nodes into (non-overlapping) groups $C = \{C_1, C_2, \dots C_m\}$ of bounded size as to minimize the expected number of tests needed to screen the entire network. Let $\mathcal{P}(G, \theta)$ be a probability distribution over possible outcomes for a transmission in G given parameters θ (e.g., infection rate, seed nodes). Each outcome maps the set of nodes in G to a set of states $X : V \to \{0, 1\}^{|V|}$ (negative or positive). Moreover, let $r(X, C_i)$ be a function that computes the number of tests required by Algorithm 1 for a group C_i given an outcome X. The total number of tests for X given groups C is given by:

$$R(X, C) = \sum_{i=1}^{m} r(X, C_i) \tag{1}$$

We define $\sigma(G, \theta, C)$ as the expected number of tests for C under the distribution $\mathcal{P}(G, \theta)$:

$$\sigma(G, \theta, C) = \mathbb{E}_{X \sim \mathcal{P}(G,\theta)}[R(X, C)] \tag{2}$$

Notice that we do not assume that an analytical expression for $\mathcal{P}$ exists, instead we approximate σ using Monte Carlo simulations. We are now able to define our problem:

Definition 2. ***Group Testing on a Network (GTN):*** *Given a transmission network G, epidemic parameters θ and a maximum group size k, partition the vertices V into groups $\{C_1, \dots C_m\}$ such that $|C_i| \leq k$, $\forall i$, and the expected number of tests $\sigma(G, \theta, C)$ to screen V is minimized.*

We emphasize that the number of groups m is not an input of the problem. Moreover, group sizes are only upper-bounded by k. In the next section, we will expand Equation 2, which is the objective to be minimized in GTN.

A Probabilistic Model for Group Tests

Based on our problem definition, we will formalize how the groups affect the expected number of tests. Intuitively, under a fixed prevalence, the group design should maximize the probability of co-infection within relatively small groups. As positive cases are transmitted through the network, its structure can be used to optimize the group design.

Let the outcomes of a process $X = \langle X_1, X_2, \dots X_n \rangle$, $n = |V|$, be a multivariate Bernoulli random variable (Teugels 1990) with parameters $p_v = \mathbf{Prob}(X_v = 1)$ for all v. Given a fixed set of m groups C, we have that:

$$\sigma(G, \theta, C) = m + \sum_{j=1}^{m} |C_j| \times \mathbf{Prob}(\sum_{v \in C_j} X_v \geq 1) \quad (3)$$

$$= n + m - \sum_{j=1}^{m} |C_j| \times \mathbb{E}[\prod_{v \in C_j} (1 - X_v)] \quad (4)$$

The above equation shows the trade-off between the number of groups and the probability of a positive case within each group in minimizing the σ. The expectation is a monotonic non-decreasing function of the group sizes and $m \leq \sigma \leq m + n$. While small groups lead to many tests at the first stage, large groups might lead to many positive groups and thus a large number of tests in the second stage.

The original analysis of Dorfman's design assumed that groups were assembled at random and with fixed size $|C_j| = k$ (i.e. $m = n/k$) and that infections were i.i.d with $p_v = p$ for all v. In such setting, one can show that:

$$\mathbb{E}[\prod_{v \in C_m} (1 - X_v)] = (1 - p)^k \quad (5)$$

Here, we are interested in the case where positive cases are transmitted through the network and thus we need to account for the correlation/covariance between variables X_v. For instance, in the case of two variables, X_u and X_v:

$$\mathbb{E}[(1 - X_u)(1 - X_v)] = 1 - p_u - p_v + p_u p_v + \mathbf{cov}_{u,v} \quad (6)$$

where $\mathbf{cov}_{u,v}$ is the covariance between variables X_u and X_v. For k variables, the expectation is a polynomial of degree k—i.e. a function of moments of the Bernoulli distribution with order up to k.

The key idea of our network-based group testing design is to minimize the objective in Equation 3 by exploiting our knowledge of the transmission network structure—the driver of the correlations between node outcomes X_v—and also the transmission process. In the next section, we will focus on characterizing the computational hardness of GTN.

The Hardness of Group Testing on a Network

We will show that the Group Testing on a Network is NP-hard, as is the case for many other graph partitioning problems (Fortunato 2010; Andreev and Racke 2006). This result forces us to search for approximate algorithms and heuristics to solve GTN for large graphs, which is the focus of the next section. Moreover, we also show that computing the expected number of tests for a given set of groups in a network is #P-hard (Valiant 1979; Arora and Barak 2009).

Theorem 1. *The Group Testing on a Network (GTN) problem (Definition 2) is NP-hard.*

Proof. The proof is by a reduction from 3-partition (3P), for which an instance contains a set of integers $A = \{a_1, a_2, \dots a_n\}$, where $n = 3k'$, and a threshold h such that:

$$\frac{h}{4} < a_i < \frac{h}{2}, \forall a_i \in A \quad (7)$$

$$\sum_{i=1}^{n} a_i = k'.h \quad (8)$$

The problem asks whether elements of A can be partitioned into triplets such that each triplet sums to h and is known to be strongly NP-complete (Garey and Johnson 1979).

We reduce an instance of 3P to an instance of GTN as follows. Create a graph G that is a union of n cliques, one for each element $a_i \in A$, where clique i has a_i vertices. Edge weights $W_e = 1$ and the maximum group size $k = h$. The last step is to define a process for which we must guarantee that the prevalence is small enough so that each group has exactly k members. This is achieved with a process that selects a seed node from V uniformly at random with probability $1/(|V| + 1)$. The seed will infect the entire clique it belongs to with probability 1—e.g. as for an independent cascade process with edge probability 1 (Goldenberg, Libai, and Muller 2001; Kempe, Kleinberg, and Tardos 2003).

It follows that $3P$ has a solution iff its corresponding GTN solution is such that the expected number of tests σ is $m + k/(|V| + 1)$. That is due to the fact that each clique will be contained in exactly one group, which guarantees that infections do not leave the group from which they originated. □

We also analyze the counting complexity of computing the expected number of tests σ (see Equations 2 and 3).

Theorem 2. *Computing the expected number of tests $\sigma(G, \theta, C)$ for a group assignment C is #P-hard.*

Proof. We show a reduction from the influence function of *influence maximization* under the Linear Threshold (LT) process, which is known to be #P-hard (Chen, Yuan, and Zhang 2010). Given a graph $G'(V', E', W')$, with edge weights $W' : V' \times V' \to [0, 1]$, vertex thresholds $\Lambda : V' \to [0, 1]$, and a seed set of vertices $S \subseteq V'$, the influence function $\psi(S)$ gives the expected number of vertices in V' activated by an LT process with S as seeds. According to LT, a vertex v is activated at a discrete time t if the weighted sum of its activated neighbors satisfies:

$$\sum_{u \text{ is active at t}} W(u, v) \geq \Lambda(v)$$

We convert the problem of computing $\psi(S)$ to the problem of computing σ for an instance of GTN as follows. Let the graph G and the transmission process $\mathcal{P}$ be the same as in the influence maximization problem—same seed set S. Finally, let each group in C contain a single vertex from V (i.e., $k = 1$). It follows that $\sigma(G, \theta, C) = |V| + \psi(S)$. □

Notice that Theorems 1 and 2 are not redundant, as they reflect the hardness of different aspects of the problem.

Algorithm	Pseud.	Δ	Time (O)
Greedy-Top.	Alg. 2	Eq. 9	$\|V\|(\log(\|V\|)+kd)$
KL-Top.	Alg. 3	Eq. 10	stm^2k^2d
Greedy-Samp.	Alg. 2	Eq. 11	$\|V\|(log(\|V\|)+kqz)$
KL-Samp.	Alg. 3	Eq. 12	stm^2k^2qz

Table 1: Summary of our algorithms, where **Pseud.** is the pseudo-code for the high-level approach, Δ is the scoring function, **Time** is the running time complexity, d is the max degree, q is the max vertex prevalence, k is the max group size, z is the number of samples, m is the number of groups, and s and t are numbers of iterations for Kernighan-Lin.

Algorithm 2 Greedy

Require: Graph G, group size k, process parameters θ
Ensure: Groups C
1: $C = \bigcup_{v \in V} \{v\}$
2: $(C_i, C_j) = \arg\max_{C_a, C_b \in C} \Delta^g(C_a, C_b, \theta, k)$
3: **while** $\Delta^g(C_i, C_j, \theta, k) \geq 0$ **do**
4: $\quad C = (\bigcup_{q \neq i,j} C_q) \cup (C_i \cup C_j)$
5: $\quad (C_i, C_j) = \arg\max_{C_a, C_b \in C} \Delta^g(C_a, C_b, \theta, k)$
6: **end while**

Algorithms for Group Testing on Networks

We propose two types of approaches for group testing on networks. The first type consists of partitioning algorithms based on the graph topology with the group size constraint. For the second, we minimize the expected number of tests by sampling from the network process. Because both the topology and sampling-based algorithms apply the same high-level strategies, we start by describing these strategies.

Algorithm 2 (Greedy) is a bottom-up scheme that starts with singleton groups and merges smaller groups by maximizing a generic score function Δ^g. Similarly, Algorithm 3 (Kernighan-Lin) starts with groups $C^{(0)}$ and then swaps members between pairs of groups while maximizing a generic score function Δ^{kl} (Kernighan and Lin 1970). Table 1 summarizes our algorithms and their time complexities.

Topology-based Algorithms

For the topology-based methods, we will assume the transmission network G to be undirected. Our first approach, *Greedy-Topology*, combines Algorithm 2 with a density-based scoring function. More specifically, we apply a function Δ^g as to maximize the total weight of edges inside the groups while bounding the size of new groups by k:

$$\Delta^g(C_i, C_j, \theta, k) = \begin{cases} \sum\limits_{u \in C_i, v \in C_j} w_{u,v}, & |C_i \cup C_j| \leq k \\ -1, & \text{otherwise} \end{cases} \quad (9)$$

We also propose *KL-Topology*, which combines Algorithm 3 with another scoring function also based on density. Let $\delta(u, C_i, C_j)$ be the change in within-group weight for moving vertex u from C_i to C_j, $\sum_{v \in C_i} w_{u,v} - \sum_{v \in C_j - \{u\}} w_{u,v}$. Then Δ^{kl} for *KL-Topology* is defined as:

$$\Delta^{kl}(u, v, C_i, C_j, \theta) = \delta(u, C_i, C_j) + \delta(v, C_j, C_i) \quad (10)$$

Algorithm 3 Kernighan-Lin

Require: Graph G, group size k, process parameters θ, initial groups $C^{(0)}$, numbers of iterations s and t
Ensure: Groups C
1: $C = C^{(0)}$
2: **for** $1, \dots s$ **do**
3: $\quad$ **for** $C_i, C_j \in C$ **do**
4: $\quad\quad (u, v) = \arg\max\limits_{a \in C_i, b \in C_j} \Delta^{kl}(a, b, C_i, C_j, \theta)$
5: $\quad\quad$ **for** $1, \dots t$ **do**
6: $\quad\quad\quad$ **if** $\Delta^{kl}(u, v, C_i, C_j, \theta) > 0$ **then**
7: $\quad\quad\quad\quad C_i = (C_i - \{u\}) \cup \{v\}$
8: $\quad\quad\quad\quad C_j = (C_j - \{v\}) \cup \{u\}$
9: $\quad\quad\quad\quad (u, v) = \arg\max\limits_{a \in C_i, b \in C_j} \Delta^{kl}(a, b, C_i, C_j, \theta)$
10: $\quad\quad\quad$ **end if**
11: $\quad\quad$ **end for**
12: $\quad$ **end for**
13: **end for**

Sampling-based Algorithms

The solutions described in the previous section are independent of the network process, which leads to two limitations: (1) group sizes cannot be adaptive to different regions of the transmission network and (2) prior information about the process (e.g., edge directions) cannot be easily accounted for. Here, we avoid these limitations by sampling from the process and minimizing our objective (Eq. 3) over samples.

We estimate the expected number of tests σ from process samples $X^{(1)}, X^{(2)}, \dots X^{(z)} \sim \mathcal{P}(G, \theta)$ as:

$$\sigma'(G, \theta, C) = m + \sum_{j=1}^{m} \frac{|C_j|}{z} \sum_{i=1}^{z} \mathbf{1}\Big\{ \sum_{v \in C_j} X_v^{(i)} \geq 1 \Big\}$$

From bounds on sampling proportions (Ott and Longnecker 2015), the error of σ' is bounded by:

$$|\sigma(G, \theta, C) - \sigma'(G, \theta, C)| \leq \epsilon$$

where $\epsilon = O(n/\sqrt{z})$ and $n = |V|$. Although computing σ exactly is $\#$-P hard, σ' can be made arbitrarily close.

To describe *Greedy-Sampling*, let us define $f(C_i, C_j)$ as:

$$f(C_i, C_j, \theta) = Z(C_i) + Z(C_j) - Z(C_i \cup C_j) + 1$$

where $Z(C_j) = |C_i| \sum_{i=1}^{z} \mathbf{1}\{\sum_{v \in C_i} X_v \geq 1\}$ is the expected number of tests for C_i and can be computed in time $O(kqz)$ for prevalence q. The function Δ^g is defined as:

$$\Delta^g(C_i, C_j, \theta, k) = \begin{cases} f(C_i, C_j, \theta), & |C_i \cup C_j| \leq k \\ -1, & \text{otherwise} \end{cases} \quad (11)$$

Our last algorithm is *KL-Sampling*, which applies Algorithm 3 and the following scoring function:

$$\begin{aligned} \Delta^{kl}(u, v, C_i, C_j, \theta) = & Z(C_i) + Z(C_j) \\ & - Z((C_i - u) \cup \{v\}) \\ & - Z((C_j - v) \cup \{u\}) \end{aligned} \quad (12)$$

To allow the KL algorithm to adaptively discover group sizes, we also consider node transfers (besides swaps) between groups in Algorithm 3 (lines 6-9).

	$\|V\|$	$\|E\|$	**time**
Primary School (PS)	242	2,242	2 days
High School (HS)	326	2,141	1 week
Company (CP)	212	1,428	2 weeks
Conference (CF)	393	2,334	2 days
Erdos-Renyi (ER)	500	2,500	-
Gaussian Rand. Part. (GRP)	400	2,200	-
Gowalla (GW)	1,899	3,565	7 months

Table 2: Summary of transmission networks.

Experiments

We evaluate our algorithms and baselines using epidemic processes on real and synthetic contact networks. [3] [4]

Experimental Settings

Evaluation metrics: We compare testing approaches in terms of number of tests per person/vertex—the lower the better. We also analyze the running time of the methods.

Transmission networks (Table 2): We apply five real contact networks and two synthetic ones as transmission networks in our experiments. Table 2 summarizes the main statistics of each dataset. *Primary School (PS), High School (HS), Company (CP), and Conference (CF)* are real face-to-face contact networks over varying periods of time (from 2 days to 2 weeks) from *sociopatterns* (Génois and Barrat 2018).[5] We set edge weights for these networks as the (max) normalized total contact time between two people. Erdos-Renyi (ER) and Gaussian Random Partition (GRP) are unweighted synthetic graphs generated with the respective models (Erdős and Rényi 1959; Brandes, Gaertler, and Wagner 2003). For GRP, we set the average cluster size to 10, the variance in cluster sizes to 5, and the intra and inter cluster edge probabilities to .8 and .01, respectively. Gowalla (GW) is a co-location network based on user *check-ins* from the (now extinct) Gowalla social network (Liu et al. 2013). Edges in GW indicate that one user visited a place within 1 minute after another. Different from the previous datasets, we make GW directed to simulate an indirect transmission of a virus—e.g. via touching a contaminated surface.

Monte Carlo epidemic simulations: Given a transmission network G, we run multiple network SIR processes (see Background Section) with a single seed node selected at random (i.i.d.). The other parameters of the process are set for each network, according to their weight distribution, as to produce a slow progression of the infection, with γ varying from .1 and 1.5 and τ varying from 1 to 40. Each simulation is stopped when $q.|V|$ nodes in the network are infected or recovered, where $q \in [0, 1]$ is the prevalence value. A node is considered to be positive if it is either infected or recovered—i.e. as in an antibody test. Our simulations are implemented using the open-source EON Python module[6]. We report the average and standard deviation of the number of tests over 10,000 simulations.

Group Testing Approaches: We divide the approaches into three groups: I) those that do not consider the transmission network (*Random* and *Origami*); II) those that exploit knowledge of the network topology but not the epidemic process (*Modularity*, *Greedy-Topology* and *KL-Topology*); and III) those that sample directly from the epidemic process (*Greedy-Sampling* and *KL-Sampling*).

Random applies randomly selected groups as in (Dorfman 1943). *Modularity* (Girvan and Newman 2002) is a classical community detection method for which we build groups with nodes within the same community. *Origami* (Kainkaryam and Woolf 2008) is a non-adaptive method (single stage) that assigns each node to multiple groups according to precomputed publicly-available assays.[7] For a given prevalence, we generate infections that match the dimension of each assay. In case there are false negatives, we add a second stage of individual tests (similar to Dorfman's). Results for the best assay are reported. For the topology-based approaches, we search over a range of group sizes and pick the best values, while for the sampling-based ones we constraint the largest group size, k, to 64. Finally, we apply the greedy methods (*Greedy-Topology* and *Greedy-Sampling*) to initialize the Kernighan-Lin methods (*KL-Topology* and *KL-Sampling*, respectively). We found this strategy to work better than a random initialization. The number of samples, z, for the sampling-based methods was set to 1,000 in all experiments.

Testing Performance

Table 3 shows the performance of testing approaches for seven networks and a prevalence of 4%.[8] *Random* achieves the worst results but still leads to savings of at least 60%. *Origami* performs well, especially for CP and ER. These are transmission networks where groups are not good clusters. Conversely, for networks with community structure (PS, HS, CF, and GRP), topology-based approaches are among the top-performing ones. Notice that our methods consistently outperform *Modularity*. However, the *KL* steps do not lead a noticeable improvement over *Greedy*. We have also tried an initialization using *Random* and found that results are not better. Methods that sample the epidemic process achieve the best performance for most of the datasets, outperforming *Random* and *Origami* by up to 40% and 33%, respectively. Moreover, results for GW illustrate the importance of process sampling in case prior information (in this case, edge directions) about the process is available.

Figure 2 shows the testing performance of some of the approaches (*Random, Origami, KL-Topology*, and *KL-Sampling*) using four networks (PS, HS, CF, and GW) for values of prevalence varying from 2% to 32%.[9] Results for the remaining datasets follow a similar pattern and are omitted due to space constraints. *Origami* achieves good results for very low values of prevalence (below 3%), but its perfor-

[3] Code: https://github.com/arleilps/group-testing

[4] See supplementary material for details and extra experiments.

[5] http://www.sociopatterns.org/datasets/

[6] https://epidemicsonnetworks.readthedocs.io

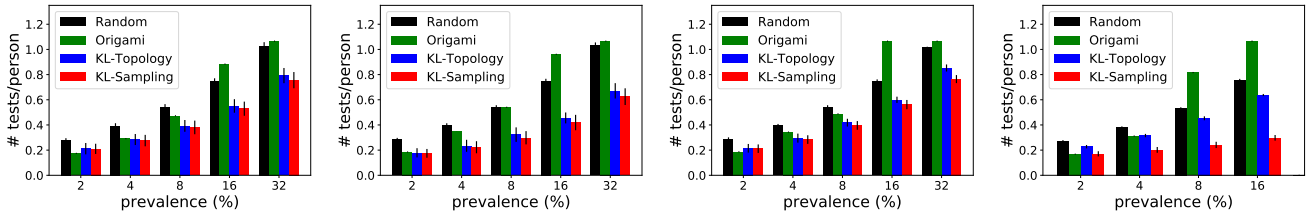
[7] https://www.smarterbetter.design/origamiassays/

[8] See the appendix for SARS-COVID-2 prevalence in the US.

[9] 16% for GW as the number of reachable vertices is often small.

	Method	PS	HS	CP	CF	GW	ER	GRP
No network	Random	.39 ± .02	.40 ± .01	.40 ± 0.2	.39 ± .02	.38 ± .01	.39 ± .01	.38 ± .01
	Origami	.29 ± .0	.35 ± .00	**.28 ± .00**	.34 ± .00	.31 ± .00	**.33 ± .00**	.32 ± .04
Topology-based	Modularity	.32 ± .05	.27 ± .04	.32 ± .06	.33 ± .04	.36 ± .01	.38 ± .02	.32 ± .04
	Greedy	**.28 ± .04**	.23 ± .05	.29 ± .05	.29 ± .04	.32 ± .01	.36 ± .02	.31 ± .04
	KL	**.28 ± .04**	.23 ± .05	.29 ± .05	.29 ± .04	.32 ± .01	.36 ± .02	.31 ± .04
Sampling-based	Greedy	**.28 ± .05**	.23 ± .05	.29 ± .05	.29 ± .04	.21 ± .03	.37 ± .02	.32± .03
	KL	**.28 ± .04**	**.22 ± .05**	**.28 .± .05**	**.28 ± .04**	**.20 ± .02**	.36 ± .02	**.30 ± .03**

Table 3: Comparison of testing schemes in terms of tests/person for the datasets described in Table 2 and a prevalence of 4%. Results for varying values of prevalence are shown in Figure 2. Our best approach (*KL-Sampling*) outperforms *Random* and *Origami* for most of the datasets and by up to 40% and 33%, respectively.



(a) Primary School (PS) (b) High School (HS) (c) Conference (CF) (d) Gowalla (GW)

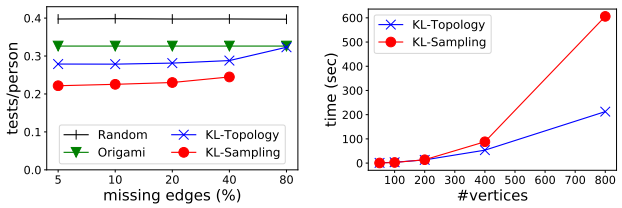
Figure 2: Testing performance of our approaches (with standard deviation) and some of the baselines at varying prevalence levels using the PS, HS, CF and GW datasets. Our best method (KL-Sampling) outperforms all competing approaches for values of prevalence beyond 2% and is still effective (for PS, HS and CF) when prevalence reaches 32%.

mance quickly degrades as the prevalence grows due to the increase in the number of false positives. When the prevalence reaches 32%, both *Random* and *Origami* become as effective as individual testing. However, the network-based approaches are still able to save up to 38% (HS) of testing resources at such a high prevalence rate.

Robustness to Missing Transmission Links

So far, we have assumed full knowledge of the transmission network in our experiments. However, in practice, the collection of such transmission data (e.g., from contact tracing) is subject to errors. Here, we evaluate the robustness of our group testing approaches to missing edges in the transmission network. For the HS dataset, we first run the epidemics on the entire network and then remove a random fraction of the edges before applying our group testing approaches.

Figure 3a shows the number of tests/person achieved by *KL-Topology* and *KL-Sampling* for a rate of missing edges varying from 0-80%. We also show the performance for *Random* and *Origami* for comparison. Results show that the savings are quite robust to missing edges. That is evidence that the groups discovered by our approaches—of high school classmates in this case—are quite dense. Even with 40% of edges missing, our approaches are still able to save up to 38% and 25% of test kits on average compared to *Random* and *Origami*, respectively. However, notice that *KL-Sampling* is unable to sample at a prevalence of 4% in the extreme case where 80% of edges are missing, which is due to the fragmentation of the network.



(a) Missing edges (b) Scalability

Figure 3: Testing performance for varying amounts of missing edges in the HS network (a) and scalability of the methods in terms of number of vertices (b). When 80% of edges are missing, *KL-Sampling* is unable to sample at the needed prevalence as the network becomes fragmented.

Scalability

Figure 3b shows the scalability of *KL-Topology* and *KL-Sampling* for networks of increasing sizes generated by the ER model. We set the number of edges in the graph and samples used by *KL-Sampling* both as five times the number of vertices. Results show that sampling the transmission process leads to an overhead in running time compared to the topology-based approach. For 800 vertices, *KL-Sampling* and *KL-Topology* take approximately 600 and 200 secs to finish, respectively. In practice, we expect our approaches to be applied to networks with a few thousand vertices—e.g. a large office, retirement home or university campus. Moreover, the Kernighan-Lin methods are easily paralelizable.

Related Work

In this paper, we consider the group testing procedure originally proposed in (Dorfman 1943). Dorfman's design is adaptive, as the outcome of tests in the first stage decides the tests in the second. Group testing can also apply more than two stages, by recursively partitioning groups at the cost of longer wait times for results (Hwang 1972). There are also non-adaptive schemes, where each individual sample is assigned to multiple groups (e.g., Origami) (Kainkaryam and Woolf 2008; Aldridge, Johnson, and Scarlett 2019) and mixed schemes (Aldridge 2020). Moreover, we have assumed that the tests are noiseless but there are methods that account also for test errors (Atia and Saligrama 2012).

Group testing has regained popularity recently due to SARS-COVID-2 (Mallapaty 2020; Brault, Mallein, and Rupprecht 2020; Gollier and Gossner 2020; Taipale, Romer, and Linnarsson 2020; Broder and Kumar 2020). While some of these recent studies follow the lines of earlier, more theoretical or simulation-based work (Beunardeau et al. 2020; Cuturi, Teboul, and Vert 2020; Gajpal et al. 2020; Broder and Kumar 2020), others are based on laboratory experiments with real SARS-COVID-2 patients (Schmidt et al. 2020; Ghosh et al. 2020). Three recent studies very related to ours are (Fang et al. 2020), (Augenblick et al. 2020) and (Nikolopoulos et al. 2020), which demonstrate the benefit of correlation to reduce the costs in group testing. However, these works assume that correlated groups are known a priori. Here, we focus on how such groups can be discovered based on a transmission network, which we show to be a computationally hard problem. In (Cheraghchi et al. 2012), a graph structure constrains the pooling procedure, playing a different role than in our problem.

This work is also related to the study of epidemic (Kiss, Miller, and Simon 2017) and other diffusion processes on networks (Granovetter 1978; Rogers 2010; Domingos 2005; Adar and Adamic 2005). For instance, influence maximization in social networks is a graph combinatorial problem defined in terms of influence processes (Kempe, Kleinberg, and Tardos 2003). Here, we focus on a graph partitioning problem, which is similar to (Barbieri, Bonchi, and Manco 2013). However, their problem definition does not take into account partition size constraints.

There is extensive literature on graph partitioning and community detection (Girvan and Newman 2002; Karypis and Kumar 1998; Fortunato 2010). The k-partition problem (Andreev and Racke 2006), of which the classical graph bisection is a special case (Garey, Johnson, and Stockmeyer 1974), also searches for balanced partitions. However, these problems do not have the partition size as a hard constraint and thus bi-criteria solutions might be acceptable, which is not our case. Still, we are able to apply ideas from existing heuristics for balanced partition to our problem (Kernighan and Lin 1970; Fiduccia and Mattheyses 1982).

Conclusion

This paper proposes a group testing design based on knowledge of an underlying transmission network. We have formalized our problem, characterized its computational hardness, and proposed heuristics for it. In our experiments, we have evaluated these heuristics, and other alternatives, using simulated epidemics on real and synthetic transmission networks. Results have shown that our approaches can save up to 33% of testing resources for a prevalence of 4%. Moreover, we are able to achieve savings even for higher values of prevalence (32%), for which competing approaches are ineffective, and when part of the transmission edges are missing.

While the evaluation of our methods using real infections would provide further evidence of their effectiveness, large-scale network infection data is not publicly available. Still, there are some promising studies that might release such type of data in the near future (Gudbjartsson et al. 2020; Klepac, Kissler, and Gog 2018).

In an epidemic, the transmission network is based on highly sensitive contact tracing information (Ahmed et al. 2020; Chan et al. 2020; Cho, Ippolito, and Yu 2020; Troncoso et al. 2020). Contact tracing apps have been developed recently to address the COVID-19 crisis (Ahmed et al. 2020) and there is a growing effort to guarantee the privacy of their users (Chan et al. 2020; Cho, Ippolito, and Yu 2020; Troncoso et al. 2020). Implementing our algorithms on a private network is an interesting direction for future research.

Appendix

Discussion on Hardness of Approximation

We have proposed heuristics for Group Testing on Networks (GTN), which is NP-hard (Theorem 1). However, we have not discussed whether our heuristics produce any approximation guarantee or if GTN can be approximated at all. We claim that there is no polynomial-time algorithm with a meaningful approximation guarantee for GTN.

First, let us define an edge sampling process in an undirected graph, where each vertex $v \in V$ becomes a seed with probability proportional to its degree and the process dies after the seed vertex infects one of its neighbors. This simple process selects edges $e \in E$ uniformly. Moreover, let us assume that group sizes are exactly k—this is the case if the prevalence is small enough. For a given group assignment C, it follows that $\sigma = m + k(1 + \phi(C)/|E|)$, where $\phi(C)$ is the number of edges connecting different groups—also known as the graph cut induced by C as partitions. Moreover, given an input network, m, k and E are constants, and thus one could directly minimize $\phi(C)$ under the constraint that groups have size k. This problem is equivalent to the balanced (a.k.a., k-partition) problem, which is NP-hard to approximate (Andreev and Racke 2006).

Notice that the above claim holds because we have removed the constant part $k(m + 1)$ from the objective. Returning to the original objective σ, because $\phi(C) \leq E$, then any algorithm can achieve an $m + 2k$ solution, which is a constant factor approximation but also an upper bound on σ.

COVID Prevalence in the US (May, 2020)

New York City 7%, Louisiana 6%, Connecticut 5%, Philadelphia 3%, San Francisco 1% (Havers et al. 2020). Our experiments cover prevalence values in the same scale.

Acknowledgements

This research is partially funded by grants NSF IIS-1817046, DTRA HDTRA1-19-1-0017 and a UCSB Office of Research VCR COVID-19 Seed Grant. We thank Sourav Medya and the anonymous reviewers for their helpful comments on this paper.

References

Adar, E.; and Adamic, L. A. 2005. Tracking information epidemics in blogspace. In *The 2005 International Conference on Web Intelligence*, 207–214. IEEE.

Ahmed, N.; Michelin, R. A.; Xue, W.; Ruj, S.; Malaney, R.; Kanhere, S. S.; Seneviratne, A.; Hu, W.; Janicke, H.; and Jha, S. 2020. A survey of covid-19 contact tracing apps. *IEEE Access* .

Aldridge, M. 2020. Conservative two-stage group testing. *arXiv preprint arXiv:2005.06617* .

Aldridge, M.; Johnson, O.; and Scarlett, J. 2019. Group testing: an information theory perspective. *arXiv preprint arXiv:1902.06002* .

Andreev, K.; and Racke, H. 2006. Balanced graph partitioning. *Theory of Computing Systems* 39(6): 929–939.

Arora, S.; and Barak, B. 2009. *Computational complexity: a modern approach*. Cambridge University Press.

Atia, G. K.; and Saligrama, V. 2012. Boolean compressed sensing and noisy group testing. *IEEE Transactions on Information Theory* 58(3): 1880–1901.

Augenblick, N.; Kolstad, J. T.; Obermeyer, Z.; and Wang, A. 2020. Group Testing in a Pandemic: The Role of Frequent Testing, Correlated Risk, and Machine Learning. Technical report, National Bureau of Economic Research.

Barbieri, N.; Bonchi, F.; and Manco, G. 2013. Cascade-based community detection. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 33–42.

Barrat, A.; Barthelemy, M.; and Vespignani, A. 2008. *Dynamical processes on complex networks*. Cambridge university press.

Beunardeau, M.; Brier, É.; Cartier, N.; Connolly, A.; Courant, N.; Géraud-Stewart, R.; Naccache, D.; and Yifrach-Stav, O. 2020. Optimal Covid-19 Pool Testing with a priori Information. *arXiv preprint arXiv:2005.02940* .

Brandes, U.; Gaertler, M.; and Wagner, D. 2003. Experiments on graph clustering algorithms. In *European Symposium on Algorithms*, 568–579. Springer.

Brault, V.; Mallein, B.; and Rupprecht, J.-F. 2020. Group testing as a strategy for the epidemiologic monitoring of COVID-19. *arXiv preprint arXiv:2005.06776* .

Broder, A. Z.; and Kumar, R. 2020. A Note on Double Pooling Tests. *arXiv preprint arXiv:2004.01684* .

Chan, J.; Gollakota, S.; Horvitz, E.; Jaeger, J.; Kakade, S.; Kohno, T.; Langford, J.; Larson, J.; Singanamalla, S.; Sunshine, J.; et al. 2020. Pact: Privacy sensitive protocols and mechanisms for mobile contact tracing. *arXiv preprint arXiv:2004.03544* .

Chen, W.; Yuan, Y.; and Zhang, L. 2010. Scalable influence maximization in social networks under the linear threshold model. In *2010 IEEE international conference on data mining*, 88–97. IEEE.

Cheraghchi, M.; Karbasi, A.; Mohajer, S.; and Saligrama, V. 2012. Graph-constrained group testing. *IEEE Transactions on Information Theory* 58(1): 248–262.

Cho, H.; Ippolito, D.; and Yu, Y. W. 2020. Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs. *arXiv preprint arXiv:2003.11511* .

Corman, V.; Bleicker, T.; Brünink, S.; Drosten, C.; and Zambon, M. 2020. Diagnostic detection of 2019-nCoV by real-time RT-PCR. *World Health Organization, Jan* 17.

Cuturi, M.; Teboul, O.; and Vert, J.-P. 2020. Noisy Adaptive Group Testing using Bayesian Sequential Experimental Design. *arXiv preprint arXiv:2004.12508* .

Domingos, P. 2005. Mining social networks for viral marketing. *IEEE Intelligent Systems* 20(1): 80–82.

Dorfman, R. 1943. The detection of defective members of large populations. *The Annals of Mathematical Statistics* 14(4): 436–440.

Du, D.; Hwang, F. K.; and Hwang, F. 2000. *Combinatorial group testing and its applications*, volume 12. World Scientific.

Erdős, P.; and Rényi, A. 1959. On random graphs I. *Publ. math. debrecen* 6(290-297): 18.

Eubank, S.; Guclu, H.; Kumar, V. A.; Marathe, M. V.; Srinivasan, A.; Toroczkai, Z.; and Wang, N. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988): 180–184.

Fang, L.; Ling, S.; Jing, B.-Y.; and Yang, Q. 2020. FAST: a Feasible, Accurate and Speedy Test Strategy for COVID-19. *medRxiv* .

Fiduccia, C. M.; and Mattheyses, R. M. 1982. A linear-time heuristic for improving network partitions. In *19th design automation conference*, 175–181. IEEE.

Fortunato, S. 2010. Community detection in graphs. *Physics reports* 486(3-5): 75–174.

Gajpal, Y.; Appadoo, S.; Shi, V.; and Liao, Y. 2020. Optimal Multi-Stage Group Partition for Efficient Coronavirus Screening. *Available at SSRN 3591961* .

Garey, M. R.; and Johnson, D. S. 1979. *Computers and intractability*, volume 174. freeman San Francisco.

Garey, M. R.; Johnson, D. S.; and Stockmeyer, L. 1974. Some simplified NP-complete problems. In *Proceedings of the sixth annual ACM symposium on Theory of computing*, 47–63.

Génois, M.; and Barrat, A. 2018. Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Science* 7(1): 11.

Ghosh, S.; Rajwade, A.; Krishna, S.; Gopalkrishnan, N.; Schaus, T. E.; Chakravarthy, A.; Varahan, S.; Appu, V.; Ramakrishnan, R.; Ch, S.; et al. 2020. Tapestry: A Single-Round Smart Pooling Technique for COVID-19 Testing. *medRxiv* .

Girvan, M.; and Newman, M. E. 2002. Community structure in social and biological networks. *Proceedings of the national academy of sciences* 99(12): 7821–7826.

Goldenberg, J.; Libai, B.; and Muller, E. 2001. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing letters* 12(3): 211–223.

Gollier, C.; and Gossner, O. 2020. Group testing against Covid-19. *Covid Economics* 2: 32–42.

Granovetter, M. 1978. Threshold models of collective behavior. *American journal of sociology* 83(6): 1420–1443.

Gudbjartsson, D. F.; Helgason, A.; Jonsson, H.; Magnusson, O. T.; Melsted, P.; Norddahl, G. L.; Saemundsdottir, J.; Sigurdsson, A.; Sulem, P.; Agustsdottir, A. B.; et al. 2020. Spread of SARS-CoV-2 in the Icelandic population. *New England Journal of Medicine* .

Havers, F. P.; Reed, C.; Lim, T.; Montgomery, J. M.; Klena, J. D.; Hall, A. J.; Fry, A. M.; Cannon, D. L.; Chiang, C.-F.; Gibbons, A.; et al. 2020. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020. *JAMA Internal Medicine* .

Hwang, F. 1972. A method for detecting all defective members in a population by group testing. *Journal of the American Statistical Association* 67(339): 605–608.

Kainkaryam, R. M.; and Woolf, P. J. 2008. poolHiTS: A Shifted Transversal Design based pooling strategy for high-throughput drug screening. *BMC bioinformatics* 9(1): 256.

Karypis, G.; and Kumar, V. 1998. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on scientific Computing* 20(1): 359–392.

Keeling, M. J.; and Eames, K. T. 2005. Networks and epidemic models. *Journal of the Royal Society Interface* 2(4): 295–307.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146.

Kernighan, B. W.; and Lin, S. 1970. An efficient heuristic procedure for partitioning graphs. *The Bell system technical journal* 49(2): 291–307.

Kiss, I. Z.; Miller, J. C.; and Simon, P. L. 2017. *Mathematics of epidemics on networks*, volume 598. Springer.

Klepac, P.; Kissler, S.; and Gog, J. 2018. Contagion! the bbc four pandemic–the model behind the documentary. *Epidemics* 24: 49–59.

Liu, X.; Liu, Y.; Aberer, K.; and Miao, C. 2013. Personalized point-of-interest recommendation by mining users' preference transition. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, 733–738.

Mallapaty, S. 2020. The mathematical strategy that could transform coronavirus testing. *Nature* .

Newman, M. E. 2002. Spread of epidemic disease on networks. *Physical review E* 66(1): 016128.

Nikolopoulos, P.; Guo, T.; Fragouli, C.; and Diggavi, S. 2020. Community aware group testing. *arXiv preprint arXiv:2007.08111* .

Ott, R. L.; and Longnecker, M. T. 2015. *An introduction to statistical methods and data analysis*. Nelson Education.

Rogers, E. M. 2010. *Diffusion of innovations*. Simon and Schuster.

Salathé, M.; Kazandjieva, M.; Lee, J. W.; Levis, P.; Feldman, M. W.; and Jones, J. H. 2010. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences* 107(51): 22020–22025.

Schmidt, M.; Hoehl, S.; Berger, A.; Zeichhardt, H.; Hourfar, K.; Ciesek, S.; and Seifried, E. 2020. Novel multiple swab method enables high efficiency in SARS-CoV-2 screenings without loss of sensitivity for screening of a complete population. *Transfusion* 1–7.

Taipale, J.; Romer, P.; and Linnarsson, S. 2020. Population-scale testing can suppress the spread of COVID-19. *medRxiv* .

Teugels, J. L. 1990. Some representations of the multivariate Bernoulli and binomial distributions. *Journal of multivariate analysis* 32(2): 256–268.

Troncoso, C.; Payer, M.; Hubaux, J.-P.; Salathé, M.; Larus, J.; Bugnion, E.; Lueks, W.; Stadler, T.; Pyrgelis, A.; Antonioli, D.; et al. 2020. Decentralized privacy-preserving proximity tracing. *arXiv preprint arXiv:2005.12273* .

Valiant, L. G. 1979. The complexity of computing the permanent. *Theoretical computer science* 8(2): 189–201.