Modelling of Vaccination Strategies for Epidemics using Evolutionary Computation

Michael Dubé
Dept. of Computer Science
Brock University
St. Catharines, ON, Canada
Email: md12ol@brocku.ca

Sheridan Houghten
Dept. of Computer Science
Brock University
St. Catharines, ON, Canada
Email: shoughten@brocku.ca

Daniel Ashlock
Dept. of Mathematics and Statistics
University of Guelph
Guelph, ON, Canada
Email: dashlock@uoguelph.ca

***Abstract*—Personal contact networks that represent social interactions can be used to identify who can infect whom during the spread of an epidemic. The structure of a personal contact network has great impact upon both epidemic duration and the total number of infected individuals. A vaccine, with varying degrees of success, can reduce both the length and spread of an epidemic, but in the case of a limited supply of vaccine a vaccination strategy must be chosen, and this has a significant effect on epidemic behaviour.**

In this study we consider four different vaccination strategies and compare their effects upon epidemic duration and spread. These are *random vaccination*, *high degree vaccination*, *ring vaccination*, and the base case of no vaccination. All vaccinations are applied as the epidemic progresses, as opposed to in advance. The strategies are initially applied to static personal contact networks that are known ahead of time. They are then applied to personal contact networks that are evolved as the vaccination strategy is applied. When any form of vaccination is applied, all strategies reduce both duration and spread of the epidemic. When applied to a static network, random vaccination performs poorly in terms of reducing epidemic duration in comparison to strategies that take into account connectivity of the network. However, it performs surprisingly well when applied on the evolved networks, possibly because the evolutionary algorithm is unable to take advantage of a fixed strategy.

I. Introduction

A *personal contact network* represents connections within a community of individuals along which a disease can spread. These networks can be generated through the use of demographic information, historical epidemic data, or by using an existing network such as YouTube watch history [18].

Evolutionary computation has been used in previous work to induce networks that satisfy chosen problems, i.e. acting as a plausible personal contact network [21]. Personal contact networks are evaluated by simulating an epidemic on the network, observing the behaviour of a disease as it spreads through the simulated community. This process is known as *graph induction* — locating graphs which best satisfy a problem under study through exploration of the solution space. The test problems presented in this paper include maximizing *epidemic duration* [5] and maximizing *epidemic spread*, this being the total number of infected individuals over the course of an outbreak. Previous work also included the epidemic *profile matching* problem: searching for networks which behave most like an epidemic profile comprised of new infections in each time step over the course of an epidemic [7].

The Susceptible-Infected-Removed (SIR) model of infection [14] [16] divides a population into three mutually exclusive groups of individuals. The *susceptible* group consists of individuals who may become infected. The *infected* group consists of individuals who are currently infected and can spread the epidemic to others. The *removed* group consists of those who cannot become infected; this is comprised of individuals who are immune (whether by vaccination or due to previous infection from which they have recovered) and individuals who have succumbed to the disease.

In the case of a limited supply of vaccine, selection of individuals for vaccination can have a significant impact on the severity of an epidemic. Four different vaccination strategies were considered in [22] and [20]. These included the simple strategy of choosing random individuals to be vaccinated, and three other strategies based on the structure of the personal contact network, which is known ahead of time.

The current study aims to evaluate four vaccination strategies to determine their relative effectiveness in reducing the length of an epidemic or cumulative number of infected individuals. Other than the baseline case in which no vaccination occurs, these strategies are *random vaccination*, *high-degree random vaccination*, and *ring vaccination*. The first two of these are also considered in [22] and [20] although we model the situation in which individuals are vaccinated *during* the time at which the epidemic is spreading. We measure the effect of the strategies upon epidemic duration and epidemic spread. Initially, the different strategies are applied to personal contact networks that are known ahead of time and static. The strategies are later applied to personal contact networks that evolve in reaction to vaccination, with these networks designed to maximize either epidemic duration or epidemic spread. We analyze the relative performance of the different strategies, as well as properties of the personal contact networks evolved.

A. Graph Theory

The personal contact networks used in this study are modelled as *combinatorial graphs*. Members within the network are the graph's *vertices* (or *nodes*) and connections between members are the *edges* of the graph. The term network and

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

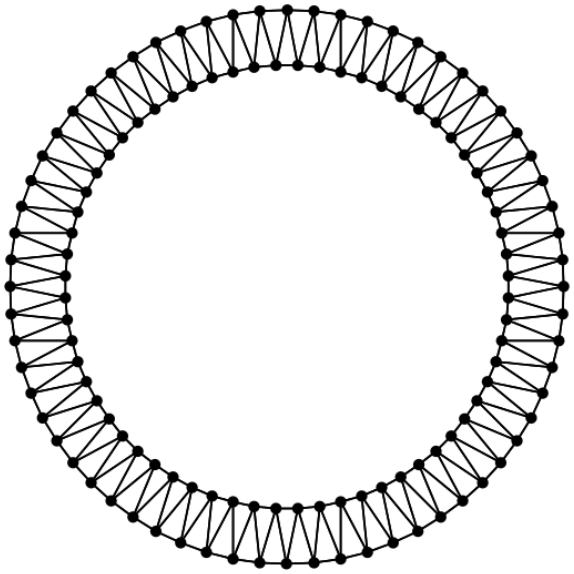

Fig. 1. Initial graph with 128 vertices for transformation with editing commands. Each vertex is the end of four edges, two ahead and two behind in the circular layout.

graph are treated as interchangeable within this paper. A graph G comprises a set of vertices V and a set edges E and is denoted $G(V, E)$. An edge between two vertices p and q from V can be represented as the unordered pair $\{p, q\}$. All edges are undirected, allowing the infection to spread both directions along the edge. A path from vertex p to vertex q on graph G is a sequence of edges from E which connect p and q. The *distance* from p to q is the length of the shortest path which connects p and q.

B. Generating Personal Contact Networks

The Local THADS-N generative representation [11] will be used in this paper to generate the personal contact networks which then are assessed for performance against the test problems. Existing applications of generative representations include [9], [13], [17]. Evidence of the effectiveness of generative solutions is demonstrated in [15].

The representation of a personal contact network is comprised of two parts: an initial graph and an ordered list of edge-editing operations which are performed upon the initial graph to produce the personal contact network. These operations can add, delete, or move edges. A generative representation was chosen as it allows the researcher to provide an appropriate starting point in the space of personal contact networks. The graph in Figure 1 was used as the initial graph because graphs with four or five edges ending at each vertex performed well for a class of epidemic modelling problems in [4], [6]. Another advantage of generative representations is that they employ a simple linear structure, a list of editing commands, that permits evolutionary computation to be applied in a simple manner.

This representation has itself evolved since the first edge operation, swap, was introduced in [3]. The operation was thought to allow for a universal operator to perform network induction with a fixed number of edges and vertices, though other operations would need to be added and have since proven to be more useful. Next, operations to allow for the addition and deletion of edges to the graph were added in [1]. These added operations improved performance, leading to exploration of several new operations. Beneficial ones were incorporated into the representation [8], [10]–[12]. The newest operations added are *local* edge operations as they ensure that the triples of vertices involved within the operation will maintain a minimum distance of two from one another before and after the editing operation is applied. Local operations are valuable when there is a global distance structure that can be disrupted by non-local operations.

1) Edge Operations from Existing Representation: Given a graph $G(V, E)$ and the vertices p, q, r, and s from the set V the existing operations are defined below.

- **Toggle(p, q):** If edge $\{p, q\}$ is in E then remove $\{p, q\}$ from E, otherwise add $\{p, q\}$ to E.
- **Local Toggle(p, q, r):** If edge $\{p, q\}$ and $\{q, r\}$ are in E then **Toggle**(p, r).
- **Hop(p, q, r):** If edge $\{p, q\}$ and $\{q, r\}$ are in E and edge $\{p, r\}$ is not in E then remove edge $\{p, q\}$ from E and add edge $\{p, r\}$ to E.
- **Add(p, q):** If $\{p, q\}$ is not in E then add $\{p, q\}$ to E, otherwise do nothing.
- **Local Add(p, q, r):** If edge $\{p, q\}$ and $\{q, r\}$ are in E then **Add**(p, r).
- **Delete(p, q):** If $\{p, q\}$ is in E then remove $\{p, q\}$ from E, otherwise do nothing.
- **Local Delete(p, q, r):** If edge $\{p, q\}$ and $\{q, r\}$ are in E then **Delete**(p, r).
- **Swap(p, q, r, s):** If $\{p, q\}$ and $\{r, s\}$ are the only edges between p, q, r and s then remove $\{p, q\}$ and $\{r, s\}$ from E and add $\{p, s\}$ and $\{q, r\}$ to E.
- **Null():** Do nothing.

C. Unleashing an Epidemic

In accordance with the SIR model of infection, all individuals in the population are initially set to the *susceptible* state except for one individual (*patient zero*) who is chosen to be infected with the disease. The status of members of the population is represented by integers in an array; the first element of the array, vertex zero, is set to be patient zero for all epidemics simulated within this paper. Furthermore, the epidemic is then permitted to spread throughout the personal contact network along edges from infected individuals to those who are susceptible. Every infected member can infect each of their neighbours with a probability of $\alpha = 0.5$, calculated independently. This simplification allows us to analyze the impact of a vaccination strategy on an outbreak; in real life situations this value could vary depending on level and duration of contact between individuals.

II. Evolutionary Computation

In order to generate the personal contact networks the strings of edge operations must first be generated to apply

TABLE I
PARAMETER SETTINGS FOR PROBABILITY DENSITIES OF THE EDGE OPERATIONS. CREATED USING A POINT PACKING WITH MINIMUM DISTANCE OF 0.535 FROM [12]. THE HEADER ROW IS POPULATED WITH PS MEANING PARAMETER SETTING, FOLLOWED BY THE EDGE OPERATIONS IN THE ORDER LISTED IN SECTION I-B1

PS	Togg.	Hop	Add	Del.	Swap	L-Togg.	L-Add	L-Del.	Null
1	0.0312	0.3528	0.2545	0.2835	0.0065	0.0151	0.0011	0.0324	0.0229
2	0.3988	0.0250	0.4167	0.0122	0.0380	0.0016	0.0325	0.0569	0.0182
3	0.0107	0.0358	0.7801	0.0243	0.0614	0.0563	0.0130	0.0057	0.0128
4	0.0438	0.0188	0.0281	0.0499	0.0003	0.4322	0.0051	0.0170	0.4047
5	0.0085	0.0104	0.0232	0.0177	0.4035	0.0008	0.0613	0.4326	0.0421
6	0.7592	0.0305	0.0024	0.0060	0.0068	0.0344	0.0106	0.1176	0.0326
7	0.0122	0.0196	0.3838	0.0942	0.0070	0.0067	0.0303	0.0173	0.4288
8	0.0020	0.0727	0.0010	0.0435	0.0355	0.0123	0.7946	0.0201	0.0184
9	0.0073	0.0261	0.3406	0.0015	0.0195	0.3741	0.0077	0.2037	0.0195
10	0.4303	0.0847	0.0105	0.0086	0.4158	0.0042	0.0025	0.0208	0.0226
11	0.0108	0.0072	0.0030	0.0617	0.4359	0.0010	0.0129	0.0223	0.4451
12	0.3682	0.0191	0.0319	0.3670	0.0191	0.1107	0.0127	0.0466	0.0247
13	0.0199	0.0032	0.3243	0.1881	0.3991	0.0012	0.0262	0.0269	0.0110
14	0.0180	0.0491	0.0002	0.7646	0.0277	0.0174	0.0661	0.0047	0.0522
15	0.0334	0.4149	0.0113	0.0005	0.0323	0.4130	0.0213	0.0373	0.0360
16	0.0129	0.3896	0.0193	0.0034	0.0038	0.0001	0.1483	0.4096	0.0130
17	0.0093	0.0211	0.0098	0.0020	0.8084	0.0532	0.0000	0.0519	0.0442
18	0.0341	0.0192	0.0044	0.2095	0.0586	0.3328	0.3233	0.0047	0.0134
19	0.0531	0.7894	0.0394	0.0261	0.0114	0.0203	0.0150	0.0164	0.0289
20	0.0038	0.0457	0.4066	0.0768	0.0226	0.0061	0.3952	0.0268	0.0164
21	0.0032	0.3982	0.0129	0.0426	0.0336	0.0115	0.0070	0.0539	0.4371
22	0.0259	0.0140	0.0032	0.2220	0.0056	0.0160	0.0203	0.3623	0.3308
23	0.0053	0.0229	0.0372	0.0192	0.0110	0.0134	0.0435	0.0045	0.8430
24	0.0176	0.3765	0.0125	0.0156	0.3637	0.0094	0.1757	0.0099	0.0190
25	0.3572	0.2016	0.0045	0.0079	0.0026	0.0088	0.3556	0.0510	0.0109
26	0.0027	0.0091	0.0011	0.0628	0.0071	0.8545	0.0484	0.0028	0.0115
27	0.0038	0.0156	0.0221	0.0021	0.0285	0.0267	0.4183	0.0396	0.4432
28	0.0265	0.0456	0.0144	0.0109	0.0179	0.0102	0.0087	0.7966	0.0691
29	0.4238	0.0343	0.0021	0.0319	0.0021	0.0066	0.0487	0.0037	0.4469

to the initial graph in Figure 1. This is accomplished using a steady state genetic algorithm [19].

The population consists of 1000 chromosomes which are themselves comprised of 256 edge operations from the list of operations described in Section I-B1. These edge operations combined with the initial graph result in a personal contact network and thus a candidate solution to the problem. To initialize these chromosomes, they are randomly assigned edge operations according to the probability densities in Table I. These probabilities were generated using a point packing [2], a procedure which allows for the creation of parameter sets with the explicit intent of best exploring the parameter space of the problem. To run this algorithm, a researcher must choose a value for the minimum allowable distance between parameter settings. The 29 parameter settings used here come from [12], and correspond to a minimum distance of 0.535.

After generating the initial chromosomes and calculating their fitness according to a fitness function, they undergo 500 000 mating events, with statistical reporting being conducted every 5 000 mating events. A mating event is made up of three procedures: tournament selection, crossover and mutation. Size-7 tournament selection is used in this paper. First, seven chromosomes are chosen at random. The two chromosomes with the worst fitness are replaced by copies of the two with the best fitness. These copied chromosomes are known as *parents* and undergo two-point crossover resulting in two *children* which overwrite their parents' chromosomes. Next, 1-3 mutations are applied to the children, chosen uniformly at random. New edge operations are chosen according to the probability densities from the parameter set being used. Lastly, the fitness of the children is calculated. After evolution the chromosome resulting in a solution with best fitness from the whole population is saved. This process is repeated 30 times for each experiment being tested. One change to the tournament selection used in this paper is the re-evaluation of the fitness of the two chromosomes after being copied to make the parent chromosomes. This procedure, known as *skeptical tournament selection*, was first described in [21] and has been shown to be useful in [10]–[12].

This process does not indicate the *absolute* quality of a network, but instead its *relative* quality, permitting successive solutions to converge to graphs which are more likely to create epidemics satisfying the problem at hand.

III. EXPERIMENTAL DESIGN

In this study we consider the modelling of an epidemic with respect to two different fitness functions, each of which is combined with four different vaccination strategies.

A. Epidemic Duration

The *duration* or *length* of an epidemic is the number of time steps elapsed from the time at which it is unleashed, up until the time at which there are zero infected individuals within the population. The current study applies epidemics both to static graphs that are known ahead of time (see Section III-D) and to graphs that evolve in reaction to the vaccination strategy (see Section III-E). To evaluate epidemic length when applied to a given graph, we use *skeptical tournament selection* from [21].

B. Epidemic Spread

The *spread* of an epidemic is the total number of individuals infected by the epidemic over its entire course, i.e. from the time at which it is unleashed up until the time at which there are zero infected individuals within the population. As with epidemic length, this is applied both to static graphs and to graphs that evolve in reaction to the vaccination strategy. Evaluation of epidemic spread is a simple sum of the number of newly infected individuals during each time step. Additionally, in order to prevent evolution from coalescing around networks with ever-increasing edge counts the fitness function evaluates to zero whenever the total number of edges is greater than five times the number of vertices. Skeptical tournament selection is used from [21].

C. Vaccination Strategies

One of the goals of this study is to evaluate different vaccination strategies with respect to their effect upon (a) epidemic length and (b) epidemic spread. Four strategies are applied to each of these, as follows:

- **No vaccine:** No individuals are vaccinated at any point.
- **Random:** At each time step, one individual from the *susceptible* category is selected for vaccination. Selection is performed uniformly at random.
- **High degree:** At each time step, one individual from the *susceptible* category is selected for vaccination. Selection is performed uniformly at random amongst all of the

individuals (vertices) that have the highest degree of all nodes in the graph.

- **Ring:** At each time step, one individual from the *susceptible* category is selected for vaccination. Selection is performed uniformly at random from amongst all of the individuals (vertices) that are neighbours of infected individuals.

As stated earlier, according to our model infection can pass from one individual to a neighbouring individual within one time step. The length of time before a vaccine becomes effective can vary from one vaccine to another. In the current study, the vaccine becomes effective and the individual is added to the *removed* category immediately upon being selected to receive the vaccine according to the current vaccination strategy.

D. Static Graphs

The epidemic duration (ED) problem [4] seeks to find graphs which promote longer-lasting epidemics. In considering the epidemic duration problem, previous work [12] generated a large number of graphs. As a first step in the current study, we take the best 30 of these, i.e. those for which epidemics had the longest duration. Using these graphs, we then evaluate the length of the epidemic when each of the four vaccination strategies listed in Section III-C are applied.

The epidemic spread (ES) problem seeks to find graphs which promote more widespread epidemics, i.e. those in which more individuals are infected over the course of the epidemic. As this problem has not been previously addressed, we generate graphs designed to maximize epidemic spread, using the process defined in Section II. Equivalently to the ED problem, we take the best 30 of these, i.e. those for which the maximum number of individuals were infected. We then evaluate the spread of the epidemic when each of the four vaccination strategies listed in Section III-C are applied.

Each of the above represents a situation in which a personal contact network is known ahead of time and does not change throughout the course of the epidemic, regardless of vaccine strategy applied.

E. Evolving Graphs

To represent a situation in which a personal contact network changes in reaction to vaccination, we also evolve graphs. To evolve the graphs, we follow the process defined in Section II. Graphs are evolved with respect to (a) maximizing epidemic length and (b) maximizing epidemic spread. In each case, the graphs are evolved while a chosen vaccination strategy is being applied. As before, we use the four vaccination strategies described in Section III-C.

IV. Results and Discussion

This section provides the results for the performance of the four vaccination strategies. These are evaluated for their effect on both epidemic duration and epidemic spread, and for both static graphs and evolved graphs.

Note that the evolutionary algorithm is on the side of the epidemic, i.e. works towards providing personal contact networks which maximize the fitness (duration or spread). From the other side, the goal of the vaccination strategies is to reduce the length or spread of an epidemic. Therefore, in terms of vaccination strategy lower values are better, even though both fitness functions have the goal of maximizing their value during evolution.

A. Static Graphs

1) Epidemic Duration: The results for the 29 parameter settings under the epidemic duration problem, applied to a static environment with the four vaccination strategies, are available in Figure 2. Recall that the epidemic is not permitted to evolve in response to the chosen vaccination strategy in a static environment. Therefore, the epidemic length without any vaccination is expected to be much higher than with any vaccination strategy as the personal contact network was explicitly evolved to achieve a maximal epidemic length without any vaccinations. This turns out to be true: for the 29 parameter settings the mean epidemic length without vaccination ranged between $33.233 - 44.1667$ compared to a drastically reduced $7.3 - 19.5667$ for the other three vaccination strategies. Regardless of strategy, within a static environment the introduction of vaccines has a significant impact on overall length of the epidemic within the population for all parameter settings.

Other than when there is no vaccination, the random vaccination strategy performs worst for all parameter settings except experiments 5, 17, and 20. However, the confidence intervals also typically overlap between these strategies and are thus not statistically significant. The random high-degree vaccination strategy performs exceptionally well in experiment 18 with a mean length of 11.9 ± 1.7319 compared to 17.5333 ± 3.195 for ring vaccination and 36.4333 ± 1.2142 without vaccination. Ring vaccination and high-degree vaccination perform similarly for the majority of experiments, though in experiments 5, 10, and 17 ring vaccination outperforms the other strategies. Lastly, the standard deviation for ring vaccination is larger than the other vaccination strategies leading to more variability in outcome when that strategy is chosen.

2) Epidemic Spread: The results for the 29 parameter settings under the epidemic spread problem, applied to a static environment with the four vaccination strategies, are available in Figure 3. The lack of a vaccine results in the totality of the population becoming infected during the course of an epidemic for each of the 30 runs performed on all parameter settings except for 2, 3, and 6. In experiments 2 and 3 the number of edges in the network happened to be higher than the cutoff to make the fitness evaluate to zero for all vaccination strategies. Recall that the epidemic spread fitness is zero whenever the total number of edges is more than 5 times the total number of nodes for a personal contact network. Also, in experiment 6, this was also the case for 16 of the 30 networks tested. These 16 fitness values of zero were removed from experiment 6 for all strategies before the mean fitness was calculated in order to not dramatically skew the results towards zero.

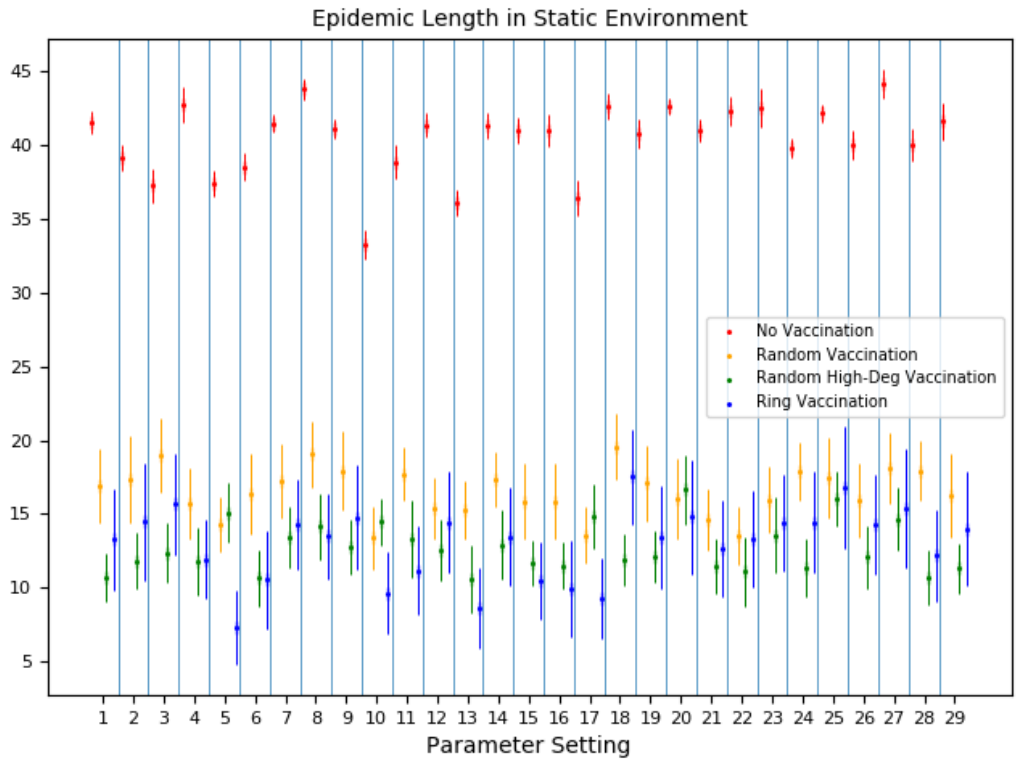

Fig. 2. The mean epidemic length, with 95% confidence interval, of epidemics unleashed on the 30 best personal contact networks for each of the 29 parameter settings with the specified vaccination strategy. The personal contact networks are from [12].

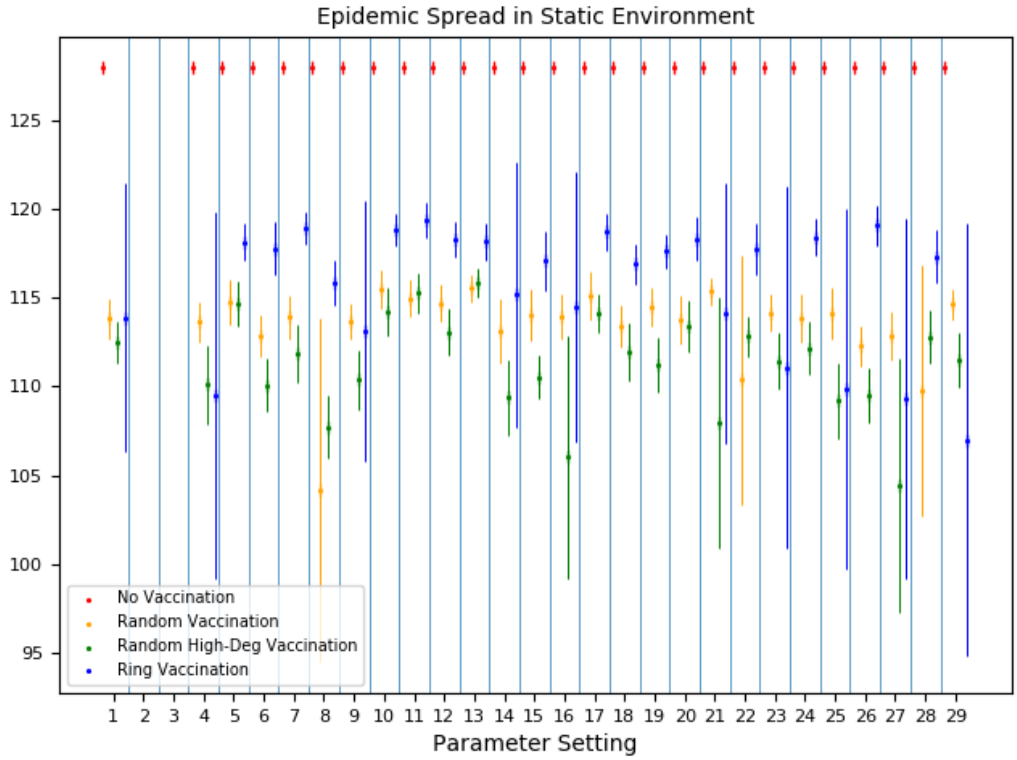

Fig. 3. The mean epidemic spread, with 95% confidence interval, of epidemics unleashed on the 30 best personal contact networks for each of the 29 parameter settings with the specified vaccination strategy. The personal contact networks are from [12].

The inclusion of vaccination does an adequate job at limiting the spread of the epidemic throughout the population, regardless of which strategy is chosen. First, the random vaccination strategy does not outperform the other strategies nor does it under-perform them. Rather, the random strategy typically falls between the performance of high-degree and ring vaccination with the confidence intervals overlapping or one being contained within another. However, random vaccination can achieve superior fitness, though these scenarios greatly increase the variance of the fitness as can be seen in parameter settings 8, 22, and 28. Similarly, in experiments 16, 21, and 27 the random high-degree vaccination strategy also achieves fitness values that are less than the other strategies; this is at the cost of consistency of outcome when the strategy is applied. Finally, the ring vaccination strategy behaves in two distinct ways: a large variance and thus inconclusive results (e.g. for experiments 4 and 25) or a small variance and poor performance (e.g. experiments 11 and 26).

B. Evolving Graphs

1) Epidemic Duration: The results for the 29 parameter settings under the epidemic duration problem, applied to an evolving environment with the four vaccination strategies, are available in Figure 4. Recall that in an evolving environment a strategy for vaccination is chosen and the evolutionary algorithm is then permitted to find personal contact networks which maximize fitness while vaccines are present. Therefore, it is not surprising that the lengths achieved in environments with vaccine present perform much better than the static results from above. Most obviously, the ring vaccination strategy performs poorly compared to the random vaccination strategies. This could be because of the less random selection of individuals to vaccinate, permitting the evolutionary algorithm to develop strategies against the fixed strategy employed by ring vaccination. In comparison, an entirely random strategy leaves little intuition to be gained by the evolutionary algorithm across generations.

The random vaccination strategies perform similarly to one another for all parameter sets, with the confidence intervals overlapping for all but experiments 15, 19, and 22. Random high-degree vaccination performs best in experiments 15 and 19 while the fully random vaccination strategy performs best in experiment 22. Furthermore, the overall reduction in the length of an epidemic is not consistent across parameter settings. For example, in experiment 10 the mean epidemic length went from 33.233 without a vaccine to 29.3667, 29.5333, and 30.3 for random vaccines, random high-degree vaccines, and ring vaccines respectively; this is a reduction of around three time steps. In contrast, in parameter setting 27 the mean lengths go from 44.1667 to 33.0667, 33.6667, and 34.4333, a reduction of around 10 time steps. Therefore, the overall impact of a vaccine on epidemic length is more dependent on the underlying personal contact network than the vaccination strategy being applied in this case.

2) Epidemic Spread: The results for the 29 parameter settings under the epidemic spread problem, applied to an evolving environment with the four vaccination strategies, are available in Figure 5. As with the static epidemic spread, parameter settings 2 and 3 both universally failed the fitness functions requirement that the total number of edges cannot exceed 5 times the total number of vertices, and are thus omitted to better see the contrast between the more interesting results from the other experiments. Also, experiment 6 once again has 16 of the 30 runs fail this requirement for all four vaccination strategies and these were removed before calculating the mean for experiment 6.

The ring vaccination strategy performs worse than both random strategies by a significant margin for all the experiments. Once again, this is likely a product of the evolutionary algorithm devising contact networks which counteract the effect of the ring vaccination strategy. Both random and high-degree strategies' results once again overlap in all but a few experiments, namely experiments 1, 15, and 24. The random high-degree strategy outperforms the random strategy for these parameter settings making them the optimal strategy in this case.

C. Static vs Evolving Graphs

The static environment does not allow for the personal contact network to respond to the vaccination strategy being applied to the population. Therefore, the outcomes of the strategies within the static environment have a much larger variance compared to those who undergo evolution. Thus, the result of applying the strategy varies greatly between runs. Whereas, when the networks are permitted to evolve the fitness values increase and variance decreases. This phenomenon is also what leads to the vaccine strategies performing better in a static environment as compared to one which is evolving.

D. Parameters

The parameter settings have a great deal of influence over the final networks being generated and thus they also greatly impact the performance of the various vaccine strategies above. Recall, experiment 2 and 3 failed when tested against the epidemic spread fitness function. Both of these settings had an abundance of *add* density with 0.4167 and 0.7801 respectively, which would result in an abundance of edges to be added to the graph likely leading to the epidemic spread fitness to evaluate to zero for all of the runs. Parameter setting 2 also had the *toggle* density set to 0.3988. Experiment 5 had a value of 0.4035 for *swap* and 0.4326 for *local delete*. *Local delete* would go about removing edges from triples of vertices which remain connected after the edge is deleted. This reduction in local connections can lead to a more restricted path for the epidemic to spread, improving the likelihood that the ring vaccination strategy immunizes a node necessary to access subsections of the graph. Experiment 5 accomplishes this on the epidemic length problem in a static environment. Parameter setting 6 also fails for 16 of the 30 runs when epidemic spread fitness is used which is likely the result of the density for *toggle* being 0.7592. As *toggle* can add or remove edges and the edges chosen are at random, it seems reasonable

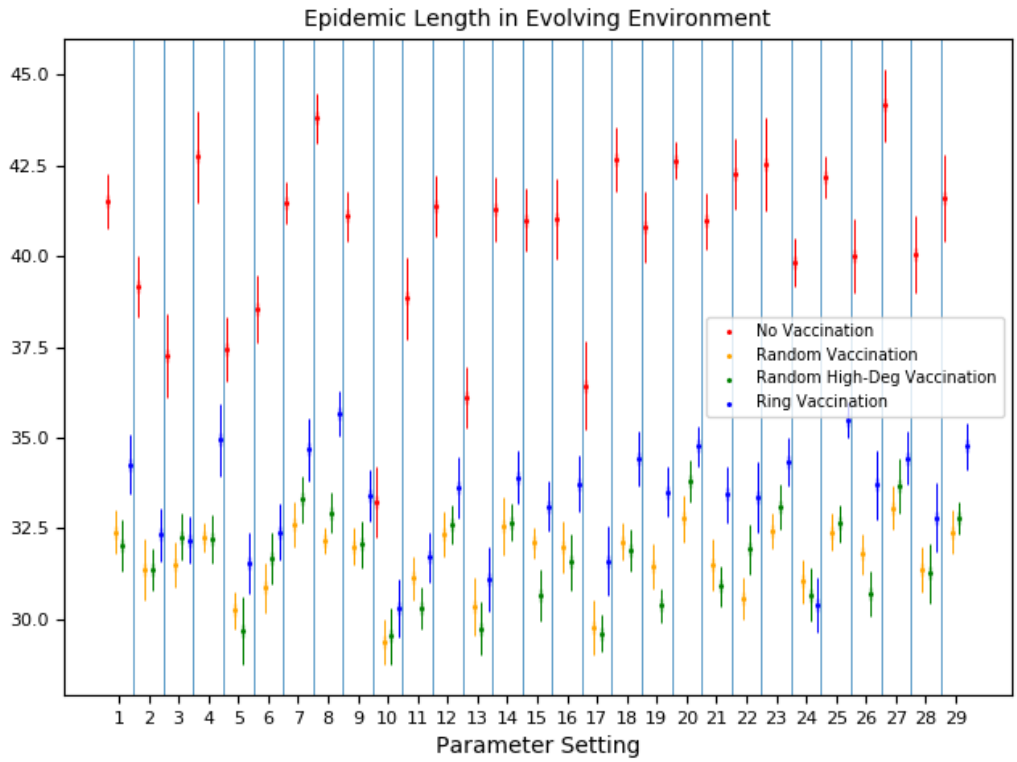

Fig. 4. The mean epidemic length, with 95% confidence interval, of the best fitness value achieved on 30 runs of the evolutionary algorithm for each of the 29 parameter settings with the specified vaccination strategy.

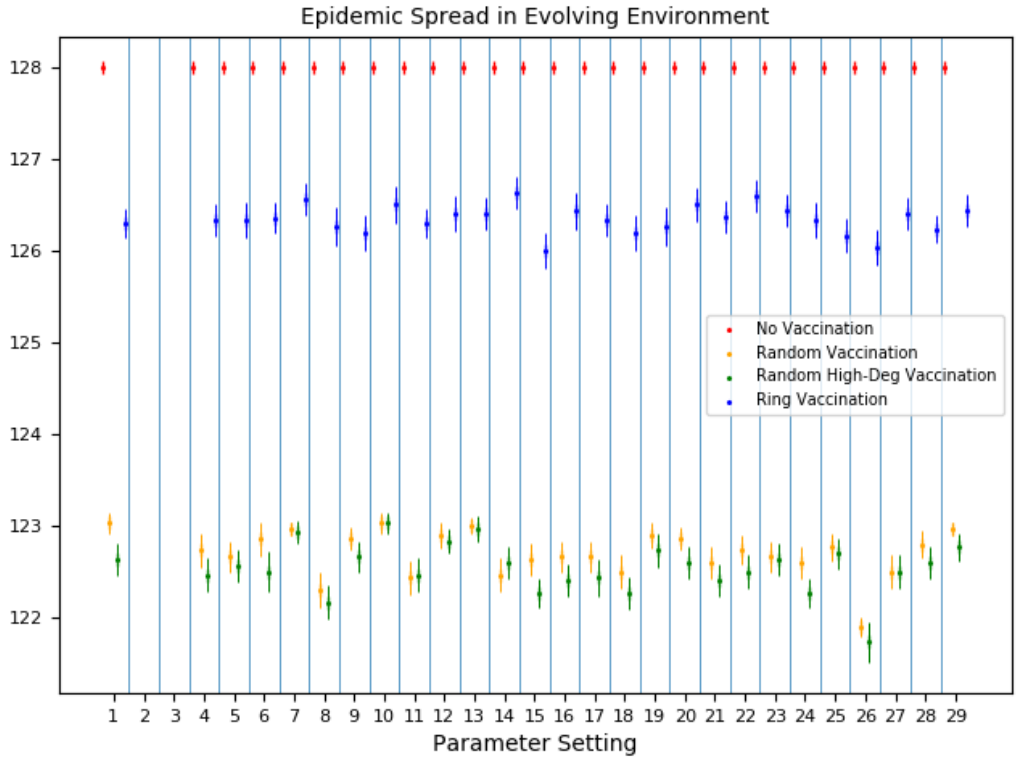

Fig. 5. The mean epidemic spread, with 95% confidence interval, of the best fitness value achieved on 30 runs of the evolutionary algorithm for each of the 29 parameter settings with the specified vaccination strategy.

that this reliance on *toggle* is the culprit for why the fitness calculation fails approximately half the time.

V. Conclusions and Future Work

Four different vaccination strategies were evaluated for their effect upon epidemic duration and epidemic spread. These were first applied to personal contact networks known ahead of time, and that had been created specifically to maximize the duration or spread of the epidemic. They were then applied to personal contact networks that were evolved with a goal of maximizing epidemic duration or spread *as the vaccination strategy was applied*. For all of the above, the vaccinations were applied as the epidemic progressed, as opposed to in advance. This is closer to a real-life situation than one in which all vaccines are applied prior to the start of the epidemic and on a known, static graph.

When applied to a static network, it was demonstrated that all vaccination strategies greatly reduced both duration and spread of the epidemic. Random vaccination was tied to increased epidemic duration in comparison to both ring vaccination and high-degree vaccination, but generally slightly reduced spread in comparison to the others.

When applied to evolved networks, again all vaccination strategies reduced both duration and spread of the epidemic, but not as drastically. For these networks, random vaccination actually reduced epidemic duration and spread. This was possibly due to the fact that the evolutionary algorithm was unable to take advantage of a fixed strategy, and also the fact that the random vaccination strategy is more likely to choose a node further away from currently infected individuals, thereby allowing some sections of the graph to prevent infection passing through.

All of the above trends concerning relative performance of the vaccination strategies hold across all 29 parameter sets considered, with some minor variation. This is despite the fact that the graphs produced by the different parameter sets are quite different from one another.

In the future, it would be of interest to consider how to evolve the vaccination strategies themselves, using known, static, personal contact networks. These evolved strategies could then be compared to existing strategies, including those presented in this paper. One could also consider such factors as a sliding scale of high-degree vaccination. It would also be of interest to apply these ideas and techniques to related issues such as the design of quarantine strategies.

Also, in the current study we considered that vaccines would take effect immediately. In real life situations, this can depend upon the type of vaccine. Therefore we could consider the following vaccine effectiveness delays: (i) that the vaccine is effective immediately, (ii) that the vaccine takes one time step to become effective, and (iii) that the vaccine takes two (or more) time steps to become effective.

Acknowledgements

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] D. Ashlock and L. Barlow. A class of representations for evolving graphs. In *Proceedings of the 2015 Congress on Evolutionary Computation*, pages 1295–1302, 2015.

[2] D. Ashlock and S. Graether. Conway crossover to create hyperdimensional point packings, with applications. In *Proceedings of the 2016 Congress on Evolutionary Computation*, pages 1570–1577, Piscataway, NJ, 2016. IEEE Press.

[3] D. Ashlock and F. Jafargholi. Representations for the evolution of extremal epidemic networks. In *Proceedings of the 2008 World Congress on Computational Intelligence*, pages 660–667, Piscataway NJ, 2008. IEEE Press.

[4] D. Ashlock and C. Lee. Characterization of extremal epidemic networks with diffusion characters. In *Proceedings of the 2008 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 264–271, 2008.

[5] D. Ashlock, J. Schonfeld, L. Barlow, and C. Lee. Test problems and representations for graph evolution. In *Proceedings of the IEEE Symposium on the Foundations of Computational Intelligence*, pages 38–45, 2014.

[6] D. Ashlock and E. Shiller. Fitting contact networks to epidemic behavior with an evolutionary algorithm. In *Proceedings of the 2011 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 1–8, Piscataway NJ, 2011. IEEE Press.

[7] D. Ashlock, E. Shiller, and C. Lee. Comparison of evolved epidemic networks with diffusion characters. In *Proceedings of IEEE Congress on Evolutionary Computation*, pages 781–788, 2011.

[8] D. Ashlock and M. Timmins. Adding local edge mobility to graph evolution. In *Proceedings of the 2016 IEEE World Congress on Computational Intelligence*, Piscataway NJ, 2016. IEEE Press.

[9] A. Devert, T. Weise, and K. Tang. A study on scalable representations for evolutionary optimization of ground structures. *Evolutionary computation*, 20(3):453–472, 2012.

[10] M. Dubé. The local thads-n generative representation for epidemic profile matching. Undergraduate thesis, Brock University, 05 2018.

[11] M. Dubé, S. Houghten, and D. Ashlock. Parameter selection for modeling of epidemic networks. In *Proceedings of the 2018 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, 2018.

[12] Michael Dube, Sheridan Houghten, and Daniel Ashlock. Representation for evolution of epidemic models. In *2019 IEEE Congress on Evolutionary Computation (CEC)*, pages 2370–2377, 06 2019.

[13] P. F. Hingston, L. C. Barone, and Z. Michalewicz. *Design by Evolution: Advances in Evolutionary Design*. Springer, New York, NY, 2008.

[14] F.C. Hoppensteadt and C.S. Peskin. *Mathematics in medicine and the life sciences*, volume 10. Springer Science & Business Media, 2013.

[15] G.S. Hornby. Improving the scalability of generative representations for openended design. In R. Riolo, T. Soule, and B. Worzel, editors, *Genetic Programming Theory and Practice V*, Genetic and Evolutionary Computation Series, pages 125–142. Springer, Boston, MA, 2008.

[16] William O Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[17] T. Kowaliw, N. Bredeche, and R. Doursat. *Growing Adaptive Machines*. Springer, 2014.

[18] Robert A. Van Gorder * Rahil Sachak-Patwa, Nabil T. Fadai. Understanding viral video dynamics through an epidemic modelling approach. *Physica A*, 502:416–435, February 2018.

[19] Gilbert Syswerda. A study of reproduction in generational and steady state genetic algorithms. In *Foundations of Genetic Algorithms*, pages 94–101. Morgan Kaufmann, 1991.

[20] Punam R Thakare and SS Mathurkar. Modeling of epidemic spread by social interactions. In *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 1320–1324. IEEE, 2016.

[21] M. Timmins and D. Ashlock. Network induction for epidemic profiles with a novel representation. *Biosystems*, 162:205–214, 2017.

[22] Zhaoyang Zhang, Honggang Wang, Chonggang Wang, and Hua Fang. Modeling epidemics spreading on social contact networks. *IEEE transactions on emerging topics in computing*, 3(3):410–419, 2015.