Network Science **6** (2): 265–280, 2018. © Cambridge University Press 2018
doi:10.1017/nws.2018.4

Simulating network intervention strategies: Implications for adoption of behaviour

JENNIFER BADHAM, FRANK KEE and RUTH F. HUNTER
UKCRC Centre of Excellence (Northern Ireland), Queens University Belfast
University Rd, Belfast BT7 1NN, United Kingdom
(*e-mail:* research@criticalconnections.com.au, f.kee@qub.ac.uk, ruth.hunter@qub.ac.uk)

Abstract

This study uses simulation over real and artificial networks to compare the eventual adoption outcomes of network interventions, operationalized as idealized contagion processes with different sets of seeds. While the performance depends on the details of both the network and behaviour adoption mechanisms, interventions with seeds that are central to the network are more effective than random selection in the majority of simulations, with faster or more complete adoption throughout the network. These results provide additional theoretical justification for utilizing relevant network information in the design of public health behavior interventions.

Keywords: *social contagion, network interventions, simulation*

1 Introduction

Research furthering our understanding of the structure and function of social networks has provided new opportunities for the design and implementation of behavior change interventions to improve the health of individuals and populations (Kim et al., 2015; Valente et al., 2015). Social network interventions involve purposeful efforts to use social network data to help generate social influence, accelerate behavior change, and/or achieve desirable outcomes among individuals, communities, organizations, or populations (Valente, 2012). For example, social network data can be used to reach new participants, select individuals on the basis of some network property who may have greater roles in providing information or support, or focus the intervention on certain groups of people; therefore, improving the efficiency and effectiveness of public health interventions.

Such interventions work by the spreading or diffusion of knowledge and behavior across interpersonal ties (Valente & Davis, 1999). Mechanisms that might explain the effect of networks on health behavior change include conformity to group norms, social facilitation, social learning, social comparison, social support, coercion, and competition (Berkman et al., 2014; Latkin & Knowlton, 2015).

Designing interventions that deliberately foster such mechanisms requires (1) some understanding of the network over which the intervention is to be applied, (2) the identification of certain individuals or groups within a network to spread or diffuse the knowledge and/or behavior, and (3) that these individuals or groups are willing to take part in the intervention and implement the intervention processes.

A recent review (Valente, 2012) identified four categories of approaches in which social networks could be used to change behaviors within public health interventions:

- Individuals: identify participants with specific network properties to act as behavior change agents.
- Segmentation: identify communities for the intervention to be applied to all members.
- Induction: encourage additional use of the network, for example by encouraging participants to talk about specific issues with their friends or asking participants to nominate potential participants.
- Alteration: change the network, for example, by assigning a support person for each participant, introducing new people and edges into the network.

The "Individuals" category provides the most basic intervention approach and includes methods to identify seed participants based on their network position, with the objective that these participants will then promote the desired health behavior to their social network via different mechanisms of behavior change. A common real-world intervention of this type involves using peer nominations to identify leaders to promote behavior change, and has been shown to increase behavior adoption in both real-world studies and fitted simulations (Kim et al., 2015; Zhang et al., 2015). Other network informed intervention approaches, such as identifying peripherals or bridges to act as seeds, have received only limited attention.

Some network intervention approaches in the other categories can also be conceptualized as identifying seeds for cascades of behavior adoption; segmentation identifies groups of people to change at the same time, and induction stimulates peer-to-peer interaction to create cascades in information/behavioral diffusion. However, the specific network properties to be used to identify such seeds so as to maximize the diffusion of the intervention, and thus the behavior, is not clear.

The purpose of this study is to extend the work by Valente (2012) in describing and classifying these network interventions, so as to investigate the potential impact of these approaches on the effectiveness of public health behavior, as measured by speed or reach of behavior adoption. We show that interventions seeded with people who are most central in the network would lead to greater and faster adoption than random initial participants over a variety of network structures under two general diffusion mechanisms.

2 Methods

We compare the effect of interventions using agent-based simulation, with a model implemented in NetLogo (Wilensky, 1999). Fifteen network interventions are included; seven from the Individuals class, two Segmentation, four Induction, and two random selection methods for comparison (Table 1). To ensure comparability across simulations, we have not included interventions that alter the network (distinct class in Valente, 2012) or cannot be operationalized in the form of identification of seed participants for behavior diffusion.

These interventions were used to select seed adopters in both real and generated networks. Required network properties were calculated using a combination of the NetLogo Network extension (Wilensky, 1999), NetLogo R extension (Thiele &

Table 1. Operationalization of network interventions.

Class and Name	Description*
Baseline	
Random Uniform	Each node has an equal chance of selection
Random by Degree	Nodes selected with probability proportional to degree
Individuals	
Ind: Degree	Descending order by degree – number of edges to other nodes
Ind: Closeness	Descending order by closeness – number of edges that must be traversed to reach other nodes
Ind: Betweenness	Descending order by betweenness values – number of shortest paths between pairs of nodes that pass through the given node
Group: Degree	Nodes which, between them, have edges to the most others
Group: Closeness	The nodes which, between them, are closest to all others
Group: Betweenness	The nodes which have the most paths between other nodes passing through at least one
Peripherals	Ascending order by closeness
Segmentation	
Community	Randomly selected from largest community(s) by breaking high betweenness edges until maximum modularity achieved
Clique	Select entire cliques in descending size
Induction	
Persuasive	Random (uniform) selection, but those selected have twice the effect of other nodes during contagion (that is, twice the transmission probability for simple, or count as two neighbors for complex)
Random Walker	Start with a random node, then select a random network neighbor, then a random neighbor of that node and so on until the required number of nodes is obtained; allows backtracking
Friends of Popular	Select highest degree node, then from those at distance 1, then from distance 2 and so on
Community Leaders	Highest degree from each community

*For those interventions where nodes are selected according to rank order, random selection is used for equally ranked nodes if they cannot all be included.

Grimm, 2010) to access R (R Core Team, 2015, v3.2) and the R packages "igraph" (Csardi & Nepusz, 2006) and "keyplayer" (An & Liu, 2016), which implements group selection (Borgatti, 2006).

The interventions were simulated over eight different networks: four real-world networks and four generated networks (including the hypothetical network used by Valente (2012) to demonstrate the different approaches). The properties of the networks used are summarized at Table 2. These networks are comparable on different properties, thereby supporting an analysis of whether the simulation results are network specific.

The four-real networks were selected for similarity in size to the hypothetical network, with friendship or social interaction as their key relationship: ham radio communication (Bernard et al., 1980), nominated friends in a prison (MacRae, 1960), observed social interaction in a clothing factory (Kapferer, 1972), observed regular interaction at social activities of a karate club (Zachary, 1977). These were obtained from the Pajek reposity of UCINet datasets (BKHAMB, PRISON, KAPFTI1, and

Table 2. Properties of networks used in simulations.

Network	Nodes	Edges	Mean degree	Gini degree	Transitivity	Mean path
Fixed degree	44	86–88	3.9–4.0	0.00–0.03	0.00–0.18	2.7–3.2
Random graph	44	63–141	2.9–6.4	0.13–0.36	0.00–0.22	2.2–4.1
Preferential attachment	44	85	3.9	0.27–0.41	0.03–0.20	2.2–2.9
Hypothetical	44	63	2.9	0.30	0.21	4.5
Ham radio	41	153	7.5	0.50	0.50	2.0
Prison	67	142	4.2	0.26	0.29	3.4
Tailor shop	35	76	4.3	0.36	0.29	2.5
Karate club	34	78	4.6	0.39	0.26	2.4

ZACHE, respectively, from `http://vlado.fmf.uni-lj.si/pub/networks/data/UciNet/UciData.htm`). To ensure comparability of simulation saturation results within and between networks, only the largest component was retained and each network was symmetrized if required.

Three network algorithms were used to generate multiple instances of networks with 44 nodes (size of the hypothetical network) and average degree of approximately 4 (integer similar to real networks). The three algorithms provide different levels of structure in their degree distribution. The fixed degree networks were generated by iterating through the nodes, randomly selecting other nodes to pair with until all had achieved the target degree, except for occasional instances where a self-loop would have been created (Molloy & Reed, 1995). The random graph networks were generated by iterating through pairs of nodes, with an edge created with probability 0.093 (Erdös & Rényi, 1960). The preferential attachment networks were generated from an initial complete graph of size 4, adding the other 40 nodes to the existing network with four edges each and connection selected with probability proportional to degree (Barabási & Albert, 1999).

Two rules representing simple and complex contagion (Valente, 1996; Centola & Macy, 2007) were used in separate sets of simulations to spread behavior adoption to the remainder of the network. For both rules, behavior is maintained once adopted.

Simple contagion was operationalized probabilistically, where each person who has already adopted the behavior has a fixed probability of triggering adoption by each network neighbor. The simulations were run until all people had adopted the behavior. This rule is an idealized representation of information provision interventions, for example, where trained peer educators may pass on their new knowledge informally to network members.

In the complex contagion simulations, each person adopts a behavior once some threshold proportion of their neighbors has already adopted it (Valente, 1996). These simulations were run until there were no new adoptions. This rule is an idealized representation of more sophisticated public health interventions that involve peer support.

As well as the network and the intervention, two other parameters were varied for the simulation experiments. The size of the seed group was set to 10%, 15%, or 20% of the network (rounded up), reflecting the common recruitment target of 15% to establish a critical mass for diffusion of information and peer support (Kelly &

Table 3. Experimental design: simulation parameters and number of runs.

Parameter	Values
Network	8 types: 4 real, 1 hypothetical, 3 generated
Intervention	15 rules
Transmission	2 types: Simple (probabilistic) or Complex (threshold)
Seed group size	10%, 15%, 20% of network
Transmission or Threshold	0.2 to 0.7 by 0.1
Repetitions	100 for most simulation sets
	1000 for simple contagion on 3 generated network types with any of 4 interventions (Random Uniform, Random by Degree, Persuasive, Random Walker).

Stevenson, 1995). The probability of transmission (simple) or threshold (complex) was varied between 0.2 and 0.7 in increments of 0.1.

Overall, there were 2,160 parameter combinations tested (8 networks by 15 interventions by 3 seed group sizes by 6 transmission/threshold values). For many simulation sets, potential variability was limited because the network was given or the intervention tightly constrained the selection of starting nodes. In addition, complex contagion is a deterministic process. For the limited variability simulation sets, 100 simulations were conducted for each parameter combination. However, 1,000 simulations were conducted for simple contagion on the generated networks for four interventions due to the combined variability of network and behavior adoption, with a new network instance for each run.

The experimental design is summarized at Table 3, requiring 626,400 simulations overall in a full factorial design. These simulations were managed with BehaviorSpace, the batch simulation tool in NetLogo, and results were analyzed using R (R Core Team, 2015), particularly the packages dplyr (Wickham & Francois, 2016, v 0.5.0) and ggplot2 (Wickham, 2009, v 2.1.0).

3 Results

3.1 Simple Contagion: probabilistic transmission

For the simple probabilistic contagion mechanism, all nodes eventually adopt the behavior. The measure of intervention effectiveness is therefore the speed of saturation, or fewest steps. The *Group: Closeness* intervention is expected to be the most effective as, by definition, this intervention starts with the set of nodes that require the smallest average number of successful transmissions to reach all other nodes.

The mean steps over the set of simulations for each parameter combination (100 or 1,000 repetitions) is shown in Figure 1. It is immediately clear that faster saturation (dark green) is associated with higher transmission probabilities, as expected. In contrast, the relative size of the seed group has little impact on the time taken for full adoption (see supplementary materials Figure 3). For each network and intervention combination, there are 18 sets of simulations (6 probability values and 3 seed group sizes); the average over these 18 sets of the mean steps value provides an initial indication of relative effectiveness and is displayed at Tables 4 and 5. The distribution of steps to saturation is available at supplementary materials Figure 4 for selected simulation sets (15% seeds with 0.4, 0.5, or 0.6 transmission probability).

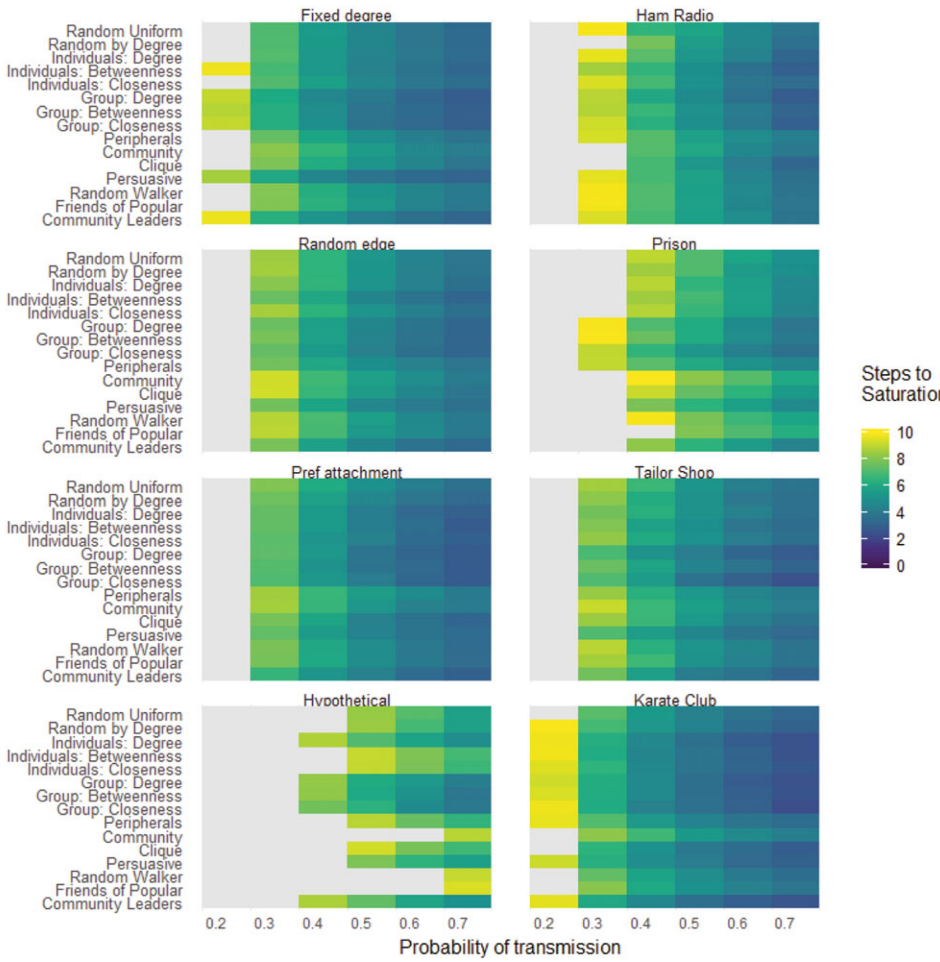


Fig. 1. Steps to saturation with simple contagion: simulation results. Each colored cell indicates the mean (over 100 or 1,000 simulations) time steps until all nodes have adopted the behavior. The number of steps is truncated; values greater than 10 are removed. Each panel includes all the results for simulations with a specific network or networks generated by the nominated algorithm and proportion of the network in the seed group. Within each panel, interventions are compared (row) for six different transmission probabilities (column). (Color online)

Saturation is much slower in the hypothetical network than in the other artificial networks (Table 4). This likely arises from the longer mean shortest path (see Table 2), which in turn reflects the lower edge density (because there are fewer edges for behavior to be transmitted along). The Prison network (Table 5) shows a similar, but weaker, effect.

Relative effectiveness of different interventions is most clear in the hypothetical network, where there is substantial variation in duration. In that network, the most effective interventions are those involving central nodes – by degree, betweenness or closeness– with the ensemble correction (group version of central nodes). Two interventions involving high degree nodes without group correction are also effective,

Table 4. Steps to saturation* with simple contagion: Artificial networks

Intervention	Fixed degree	Random edge	Pref attachment	Hypothetical
Random uniform	5.8	6.9	6.5	11.4
Random by degree	5.8	7.0	6.2	11.2
Ind: Degree	5.8	6.8	5.8	9.8
Ind: Betweenness	5.6	6.2	5.6	12.0
Ind: Closeness	6.0	6.8	5.8	12.6
Group: Degree	5.1	6.0	5.5	8.9
Group: Betweenness	5.1	6.2	5.6	8.8
Group: Closeness	5.1	6.0	5.6	8.6
Peripherals	6.3	6.4	7.1	13.5
Community	6.8	7.5	6.9	16.8
Clique	6.7	7.3	6.1	13.5
Persuasive	5.0	6.2	5.9	10.6
Random walker	6.7	7.3	6.3	16.5
Friends of popular	6.6	7.2	6.4	17.0
Community leaders	5.5	6.2	5.7	9.6

*Calculated by first finding the mean value for each simulation set (probability and seed group size combination), and then reporting the mean of those 18 results. These values therefore represent the average performance of the intervention over the network, but may not reflect simulations with the same parameter settings.

whether dispersed across communities (*Community Leaders*) or based solely on degree (*Individuals: Degree*). In contrast, those interventions with seeds that are connected to each other (*Community*, *Clique*, *Friends of Popular*) are relatively ineffective, likely because many edges from the seed adopters are "wasted" as they connect to other seeds. While there are indications of similar relative effectiveness in the other artificial networks, the pattern is weaker and has some inconsistencies. In particular, the additional benefit of the group correction for the interventions involving central nodes is much smaller over the preferential attachment network, and betweennees (rather than degree) is the effective network property for the random edge network.

The *Persuasive* intervention is also relatively effective; it randomly selects seed nodes but then those nodes have double probability of transmission. For example, in the simulations with 0.4 as the transmission parameter, the seeds trigger adoption in their network neighbors with 0.8 probability each time step. This intervention is not strictly comparable to the others, but it represents those real-world interventions that focus on opinion leaders within the network, regardless of their network characteristics. It is likely, of course, that such leaders in the real world will also be popular, but that combined effect is not simulated.

The broad pattern of intervention efficiency also occurs over real networks; interventions that involve central nodes result in shorter times to saturation than interventions with seeds from a single community (Table 5). As with the artificial networks, the need for the group correction is network dependent, with a larger impact for the Prison and Tailor shop networks in comparison to the Ham radio and Karate club networks. The *Community Leaders* intervention has no performance benefit over *Random Uniform* selection for the Ham Radio network. As

Table 5. Steps to saturation* with simple contagion: Real world networks.

Intervention	Ham radio	Prison	Tailor shop	Karate Club
Random uniform	7.6	9.2	6.8	5.9
Random by degree	7.5	9.4	6.7	5.6
Ind: Degree	7.2	9.1	6.8	5.1
Ind: Betweenness	7.0	9.1	6.5	5.0
Ind: Closeness	7.3	9.4	6.6	5.1
Group: Degree	6.9	7.9	5.8	4.9
Group: Betweenness	7.3	8.3	6.0	4.9
Group: Closeness	6.9	7.5	5.7	4.9
Peripherals	7.5	7.9	7.0	5.9
Community	7.8	10.6	7.5	7.0
Clique	7.4	10.2	6.9	5.6
Persuasive	7.4	8.5	6.1	5.3
Random walker	7.7	10.5	6.9	6.2
Friends of popular	7.6	11.0	6.9	6.6
Community leaders	7.6	9.0	6.2	4.9

*Calculated by first finding the mean value for each simulation set (probability and seed group size combination), and then reporting the mean of those 18 results. These values therefore represent the average performance of the intervention over the network, but may not reflect simulations with the same parameter settings.

for the artificial networks, *Community*, *Clique,* and *Friends of Popular* are relatively ineffective; except that *Clique* is effective over two networks, Ham radio and Karate Club.

While the average durations for the 18 parameter combinations provide a broad pattern, the relative performance must be examined for comparable simulation sets: each specific combination of transmission probability and seed proportion. From Figure 1 (and supplementary materials Figure 4), the general pattern observed in the mean durations (Tables 4 and 5) is a reasonable indicator of relative effectiveness for comparable simulations, with some noise in the patterns. The *Peripherals* intervention is consistently ineffective, as expected because the seeds are selected to be as far from other nodes as possible (and will therefore require the maximum number of steps to saturate the network). Those that use a connected subnetwork (*Community*, *Random Walker*, *Friends of Popular*) have similar results as *Random Uniform*; seeds are generally even less effective than *Peripherals*, perhaps indicating a tendency for the behavior adoption to be "trapped" in part of the network. The *Clique* intervention could be expected to have the same difficult but does not, perhaps because cliques are small so each simulation was generally seeded with more than one clique. The group centrality (*Group: Degree*, *Group: Closeness,* and *Group: Betweenness*), *Community Leaders* and *Persuasive* interventions are the most effective. Except over the Hypothetical and Prison networks, the individual centrality interventions are also effective.

The overall pattern is not reproduced exactly at a more detailed level; some reversals occur for specific parameter sets. For example, the *Individuals: Closeness* intervention is relatively effective over the Ham Radio network except for the simulations with transmission probability of 0.7.

Table 6. Proportion of network adopting* with complex contagion: Real world networks.

Intervention	Ham radio	Prison	Tailor shop	Karate Club
Random uniform	0.44	0.57	0.55	0.54
Random by degree	0.69	0.61	0.67	0.67
Ind: Degree	0.95	0.70	0.78	0.82
Ind: Betweenness	0.94	0.67	0.79	0.79
Ind: Closeness	0.95	0.66	0.78	0.75
Group: Degree	0.88	0.65	0.75	0.80
Group: Betweenness	0.95	0.69	0.79	0.86
Group: Closeness	0.86	0.62	0.74	0.79
Peripherals	0.17	0.32	0.36	0.32
Community	0.65	0.47	0.70	0.54
Clique	0.85	0.47	0.74	0.66
Persuasive	0.63	0.70	0.70	0.69
Random walker	0.71	0.53	0.66	0.66
Friends of popular	0.62	0.51	0.69	0.60
Community leaders	0.41	0.69	0.61	0.84

*Calculated by first finding the mean value for each simulation set (threshold and seed group size combination), and then reporting the average of those 18 results. These values therefore represent the average performance of the intervention over the network, but may not reflect simulations with the same parameter settings.

3.2 Complex contagion: adoption with neighborhood threshold

For the complex threshold contagion mechanism, the simulation ends when no new nodes adopt the behavior. The measure of intervention effectiveness is therefore the proportion of the network that eventually adopt the behavior. For each network and intervention combination, there are 18 sets of simulations (six threshold levels and three seed group sizes); the average over these 18 sets of the mean proportion adopted is displayed at Tables 7 and 6. The distribution of proportion adopted is available at supplementary materials Figure 6 for selected simulation sets (15% seeds with 0.4, 0.5, or 0.6 threshold proportion).

For many combinations of network and intervention, the median result is that all nodes adopt (not shown specifically, but visible for some simulation sets in supplementary materials Figure 6). That is, for each threshold and seed proportion combinations, more than half the simulations lead to adoption by all nodes. Nevertheless, there is sufficient differentiation to suggest that the pattern for complex contagion has much in common with that for simple contagion.

The individual and group centrality interventions are most effective but, unlike simple contagion, the group versions do not lead to higher levels of adoption than the individuals versions. Weighting random selection by degree (*Random by Degree*) provides some of the benefit of central seeds (except for the Fixed Degree network, where such weighting has no effect). The *Persuasive* intervention is consistently relatively effective (with similar performance as the centrality interventions except over the Ham Radio network), and the *Peripherals* intervention consistently ineffective.

Community Leaders, *Community,* and *Clique* interventions have mixed results. For example, *Community* and *Clique* perform well over the Ham Radio and Tailor Shop networks, but poorly on the Prison network, while Community Leaders performs

Table 7. Proportion of network adopting* with complex contagion: artificial networks.

Intervention	Fixed degree	Random edge	Pref attachment	Hypothetical
Random uniform	0.59	0.55	0.56	0.54
Random by degree	0.59	0.60	0.69	0.61
Ind: Degree	0.60	0.72	0.81	0.77
Ind: Betweenness	0.52	0.70	0.80	0.63
Ind: Closeness	0.47	0.68	0.79	0.62
Group: Degree	0.52	0.68	0.78	0.79
Group: Betweenness	0.53	0.72	0.81	0.79
Group: Closeness	0.53	0.67	0.77	0.77
Peripherals	0.43	0.36	0.44	0.42
Community	0.37	0.53	0.56	0.35
Clique	0.37	0.57	0.73	0.56
Persuasive	0.73	0.71	0.71	0.66
Random walker	0.38	0.55	0.69	0.42
Friends of popular	0.39	0.55	0.67	0.38
Community leaders	0.59	0.69	0.78	0.79

*Calculated by first finding the mean value for each simulation set (threshold and seed group size combination), and then reporting the average of those 18 results. These values therefore represent the average performance of the intervention over the network, but may not reflect simulations with different parameter settings.

well on the Prison and Karate Club networks but poorly on the Ham Radio and Tailor Shop networks.

For detailed analysis of effectiveness, the mean proportion adopted over the set of simulations for each parameter combination (100 repetitions) is shown in Figure 2. As expected, a higher proportion of the network as seeds (see supplementary materials Figure 5) and a lower threshold are both associated with higher saturation; the former because there are more nodes already adopted to provide network neighbors that contribute to the pressure to adopt, and the latter because less pressure is required. At the lowest threshold tested, every intervention on every network leads to all nodes adopting except for the Peripherals intervention on the Ham Radio network, which has no diffusion from the seeds. At the other extreme with a 0.7 threshold, only some interventions are successful in diffusing to new nodes; with only one set of simulations able to reach saturation, the *Individuals: Betweenness* intervention over the Ham Radio network with 20% of the nodes as seeds. The largest difference in effect between interventions, therefore, arises for intermediate thresholds, where only some are able to trigger potentially several steps in an adoption cascade. At a threshold of 0.5, some of the interventions induce more than three times the adoption level than is achieved by uniform random selection over almost all networks and seed proportions (not *Fixed Degree*, and only for starting proportion of 0.1 for Hypothetical and Prison networks).

Comparing the individual and group centrality interventions with the same parameter combination, one common pattern is that the group corrected betweenness (*Group: Betweenness*) and individual-based degree interventions are the most effective, generally followed by the other individual centrality interventions then the other two group corrected interventions. While this occurs for most thresholds and

https://doi.org/10.1017/nws.2018.4 Published online by Cambridge University Press

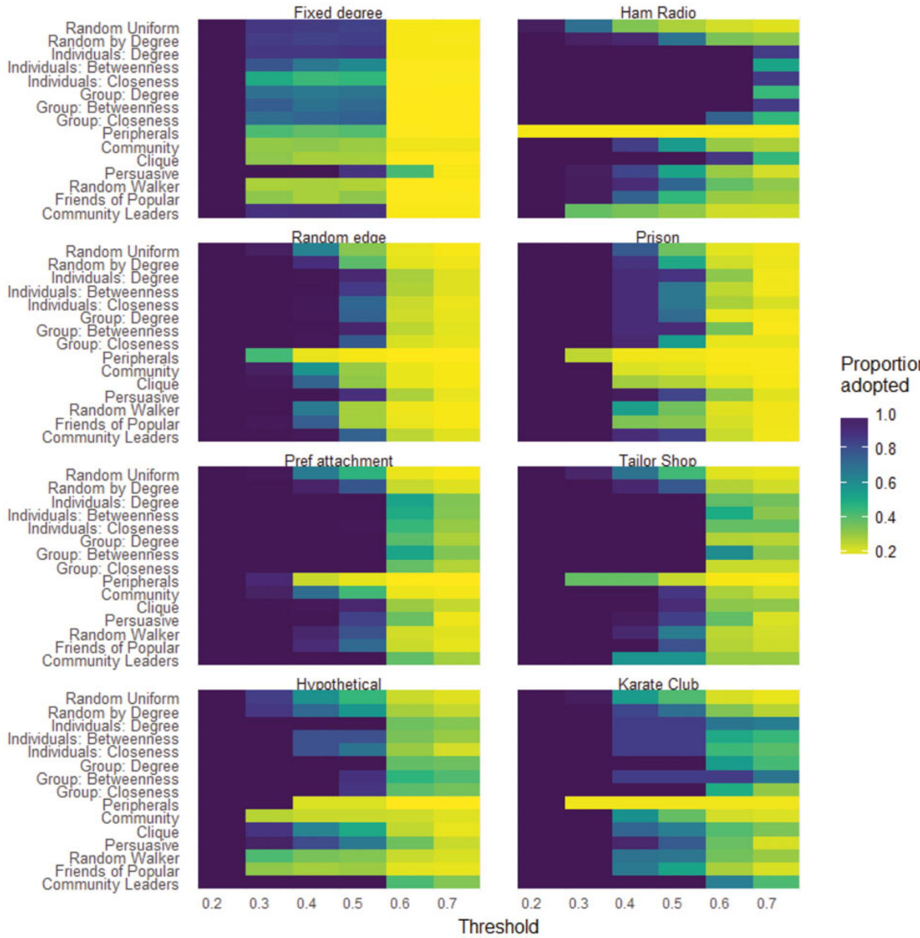


Fig. 2. Proportion of network adopting with complex contagion: simulation results. Each colored cell indicates the mean (over 100 simulations) proportion of nodes that have adopted the behavior when no further nodes will adopt. Each panel includes all the results for simulations with a specific network or networks generated by the nominated algorithm and proportion of the network as seed adopters. Within each panel, interventions are compared (row) for six different thresholds that represent the proportion of network neighbors that must have already adopted for the nodes to adopt the behavior (column). (Color online)

networks, the *Group: Betweenness* intervention is less effective than the other group centrality interventions for the moderate thresholds on the Karate Club network. The Karate Club network also has different patterns depending on the seed proportion (see supplementary materials Figure 5): *Group: Closeness* and *Group: Degree* are always more effective than the individual centrality interventions once saturation is not being achieved, but *Group: Betweenness* is similar to the latter for seed proportion of 0.1 or 0.15, and is at least as effective as the other group centrality interventions for seed proportion of 0.2.

The mixed effectiveness across different networks of the *Community Leaders* intervention that was apparent for average results (Tables 7 and 6) holds across parameter sets. While it is generally as effective as the centrality interventions, it is only of similar effectiveness as *Random Uniform* for the Tailor Shop network, and on the Ham Radio network is relatively ineffective for the 0.15 and 0.2 seed proportion simulations.

There are four interventions that select seeds that are close to each other: *Community*, *Clique*, *Random Walker,* and *Friends of Popular*. These potentially have more edges to the same nodes, and therefore could be relatively effective under the complex contagion condition, at least for an initial cascade. However, this did not generally occur. They were effective over the Preferential Attachment, Ham Radio, and Tailor shop networks, with *Clique* achieving similar performance as the centrality interventions. Otherwise, however, these interventions achieved similar or lower levels of adoption as *Random Uniform.*

4 Discussion

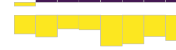
There is clear benefit in designing interventions that utilize social network structures, at least for the idealized behavior transmission processes simulated in this study (see summary at Table 8). Results demonstrated that interventions using network information to identify seeds are able to deliver substantial gains compared to random seeds. While the size of the potential benefit or loss varied considerably across networks and for different simulation parameters, there are some consistent patterns.

The ranking of intervention approaches by relative effectiveness is reasonably consistent across networks and simulation parameters. Those interventions with more central seeds are generally the most effective, with the redundancy correcting group versions outperforming the individuals versions for simple contagion only, but not for complex contagion. The *Persuasive* intervention is also relatively very effective, but some care must be taken in interpreting this result as there is no theoretical basis for the operationalization used, with persuasive individuals considered to have twice the impact as other individuals.

These results have important implications for real-world interventions. The full network structure must be known to calculate betweenness and closeness, and to apply the group correction. However, it is relatively straightforward to identify those individuals with high degree in a real-world intervention. Simply asking a uniform random sample of individuals to each nominate one of their friends generates a sample that is biased by degree (and will be weighted by degree with enough steps, see Noh & Rieger, 2004), which achieves some effectiveness gains for complex contagion. Further, such a process is relatively robust with in-degree highly correlated with the number of nominations for 30% samples in small networks and smaller proportions in large networks (Costenbader & Valente, 2003; Leskovec & Faloutsos, 2006). This approach has been shown to be effective in public health interventions (Kim et al., 2015). Alternatively, nominations can be used to identify leaders, capturing elements of both degree and persuasive interventions (Campbell et al., 2008).

The two idealized contagion processes simulated represent extremes, relying only on personal factors (simple) or only on social factors (complex). Real-world behavior

Table 8. Effectiveness of interventions across simulation sets, relative to uniform random selection.

	Simple contagion*			Complex contagion†		
	Gain	Loss	Duration	Gain	Loss	Adoption
Random uniform	0	0		0	0	
Random by degree	2	1		79	0	
Individuals: Degree	48	3		89	0	
Individuals: Betweenness	53	4		85	7	
Individuals: Closeness	37	7		82	10	
Group: Degree	113	0		84	6	
Group: Betweenness	112	0		88	7	
Group: Closeness	122	0		83	7	
Peripherals	19	36		2	98	
Community	0	86		37	33	
Clique	13	44		62	17	
Persuasive	60	0		86	0	
Random walker	0	54		58	21	
Friends of popular	0	67		58	25	
Community leaders	77	1		75	8	

* The "Gain" column is the number of simulation sets (from 144: 8 networks, 6 transmission probabilities, 3 seed group sizes), where the mean steps to saturation is at least 10% higher than the mean steps for the random uniform intervention with the same simulation parameters. Similarly, the "Loss" column is the number of simulation sets where mean saturation is at least 10% slower. The "Duration" figure displays the difference in steps between the intervention and "Random Uniform" results by network (values from Tables 4 and 5). The network order is "Fixed Degree," "Random edge," "Preferential attachment," "Hypothetical," "Ham radio," "Prison," "Tailor shop," "Karate club," with gains (fewer steps) above the line and purple, and losses (more steps) below the line and yellow. †The "Gain" column is the number of simulation sets (from 144: 8 networks, 6 thresholds, 3 seed group sizes), where the mean proportion of nodes adopted at the end of the simulation is at least 10% higher than the mean proportion for the random uniform intervention with the same simulation parameters. As the random uniform intervention leads to network saturation for some simulations, the potential maximum is less than 144. Similarly, the "Loss," column is the number of simulation sets where mean adoption is at least 10% lower. The "Adoption" figure displays the difference in adoption between the intervention and "Random Uniform" results by network (values from Tables 7 and 6). The network order is "Fixed Degree," "Random edge," "Preferential attachment," "Hypothetical," "Ham radio," "Prison," "Tailor shop," "Karate club," with gains above the line and purple, and losses below the line and yellow.

transmission is likely to have elements of both. As the centrality interventions are relatively effective with both processes, it is reasonable to expect that they would also be effective with more realistic behavior adoption mechanisms that combine individual and social factors.

https://doi.org/10.1017/nws.2018.4 Published online by Cambridge University Press

Those interventions that extract subnetworks with the intention of creating neighborhoods with high levels of adoption (*Community*, *Clique*, *Random Walker*, *Friends of Popular*) generally perform poorly, even for complex contagion where the neighborhood effect could be expected to trigger a cascade.

The relative effectiveness of other interventions varies between networks and type of contagion. In particular, *Community Leaders* could be expected to be effective as it uses nodes with high degree as seeds, but they may be in different communities and potentially far apart so the benefit may be dispersed. This intervention performed poorly over two real-world networks under both contagion conditions (Ham Radio for both, Prison for simple, and Tailor Shop for complex). The variation suggests that specific structural aspects of the particular network are important for these interventions.

Further work is required with a selection of networks with similar and dissimilar properties to potentially derive rules for intervention selection based on specific network properties. One of the properties to be varied should be the number of nodes, to assess whether differences in the effectiveness of interventions are affected by network size.

Some prior level of knowledge of the network structure is required to optimize the network intervention approach and most effectively focus intervention resources on a relatively small number of seed participants. Innovative methods are being used for data collection in large networks (Perkins et al., 2015; Shakya et al., 2017), but mapping whole social networks is costly and may not be feasible in real-world interventions. If we are to implement interventions that purposefully utilize inherent networks to inform intervention design, then we must also develop simple, low-cost methods to estimate relevant structural properties.

Further, if network interventions are to meaningfully inform public health policy and practice, then a number of implementation factors must be overcome. For example, those people identified as the preferred seeds may not wish to participate in any trial intervention, may withdraw during the study period, or may participate but not respond to the intervention. In addition, unlike the simulations, people have relevant characteristics other than their network position, and seeds may be chosen or excluded for reasons such as access, greater need, presence of other health risks, or motivation.

5 Conclusion

Utilizing the social network to most effectively deliver a public health behavior intervention has the potential to increase the reach and sustainability of the intervention at minimal cost. The best intervention (as defined by selection of seed adopters) and the potential gain available depend fundamentally on characteristics of the network and the behavior adoption mechanisms.

For a broad range of networks with around 50 people, interventions that use those people who are most central in the network as seeds would lead to greater and faster adoption than random recruitment. Further work is required to test the results on larger networks, and also to isolate the properties of networks that influence the effectiveness of specific interventions.

Acknowledgements

The authors would like to thank the reviewers for their helpful comments and questions. This work is primarily supported by a NIHR Career Development Fellowship. It was undertaken under the auspices of the UKCRC Centre of Excellence for Public Health Research Northern Ireland, which is funded by the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the National Institute for Health Research, and the Wellcome Trust.

Conflicts of interest

The authors have nothing to disclose.

Supplementary materials

For supplementary material for this article, please visit https://doi.org/10.1017/nws.2018.4

References

An, W., & Liu, Y.-H. (2016). Keyplayer: An R package for locating key players in social networks. *The R Journal*, **8**(1), 257–268. R package version 1.0.1.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, **286**(5439), 509–512.

Berkman, L. F., Kawachi, I., & Glymour, M. M. (eds). (2014). *Social epidemiology*. (2nd ed.). New York, NY: Oxford University Press.

Bernard, H. R., Killworth, P. D., & Sailer, L. (1980). Informant accuracy in social network data IV: A comparison of clique-level structure in behavioral and cognitive network data. *Social Networks*, **2**(3), 191–218.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational and Mathematical Organization Theory*, **12**(1), 21–34.

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., Hughes, R., & Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): A cluster randomised trial. *The Lancet*, **371**(9624), 15951602.

Centola, D., & Macy, M. (2007). Complex contagion and the weakness of long ties. *American Journal of Sociology*, **113**(3), 702–734.

Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks*, **25**(4), 283–307.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, **Complex Systems**, 1695.

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Institute of Mathematics, Hungarian Academy of Science*, **5**, 17–60.

Kapferer, B. (1972). *Strategy and transaction in an African factory: African workers and Indian management in a Zambian town*. Manchester, UK, Manchester University Press.

Kelly, J.A., & Stevenson, L.Y. (1995). *Opinion leader HIV Prevention Training Manual*. Milwaukee, WI: Center for AIDS Intervention Research, Medical College of Wisconsin.

Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., OMalley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, **386**(9989), 145–153.

Latkin, C. A., & Knowlton, A. R. (2015). Social network assessments and interventions for health behavior change: A critical review. *Behavioral Medicine*, **41**(3), 90–97.

Leskovec, J., & Faloutsos, C. (2006). Sampling from large graphs. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*. ACM, pp. 631–636.

MacRae, D. (1960). Direct factor analysis of sociometric data. *Sociometry*, **23**(4), 360–371.

Molloy, M., & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, **6**(2–3), 161–180.

Noh, J. D., & Rieger, H. (2004). Random walks on complex networks. *Physical Review Letters*, **92**(Mar), 118701.

Perkins, J. M., Subramanian, S.V., & Christakis, N. A. (2015). Social networks and health: A systematic review of sociocentric network studies in low- and middle-income countries. *Social Science & Medicine*, **125**(Jan), 60–78.

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Shakya, H. B., et al. (2017). Exploiting social influence to magnify population-level behaviour change in maternal and child health: Study protocol for a randomised controlled trial of network targeting algorithms in rural Honduras. *BMJ Open*, **7**(3), e012996.

Thiele, J. C., & Grimm, V. (2010). NetLogo meets R: Linking agent-based models with a toolbox for their analysis. *Environmental Modelling & Software*, **25**(8), 972–974.

Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social Networks*, **18**(1), 69–89.

Valente, T. W. (2012). Network interventions. *Science*, **337**(6090), 49–53.

Valente, T. W., & Davis, R. L. (1999). Accelerating the diffusion of innovations using opinion leaders. *The ANNALS of the American Academy of Political and Social Science*, **566**(1), 55–67.

Valente, T. W., Palinkas, L. A., Czaja, S., Chu, K.-H., & Brown, C. H. (2015). Social network analysis for program implementation. *PLoS ONE*, **10**(6), e0131712.

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.

Wickham, H., & Francois, R. (2016). *Dplyr: A Grammar of Data Manipulation*. R package version 0.5.0.

Wilensky, U. (1999). *NetLogo http://ccl.northwestern.edu/netlogo/*. Tech. rept. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, **33**(4), 452–473.

Zhang, J., Shoham, D. A., Tesdahl, E., & Gesell, S. B. (2015). Network interventions on physical activity in an afterschool program: An agent-based social network study. *American Journal of Public Health*, **105**(S2), S236–S243.