Can Social Contagion Spread Without Key Players?

Gizem Korkmaz, Chris J. Kuhlman
Biocomplexity Institute of Virginia Tech
Email: {gkorkmaz, ckuhlman}@vbi.vt.edu

Fernando Vega-Redondo
Department of Decision Sciences, Bocconi University
Email: fernando.vega@unibocconi.it

***Abstract*—Contagion models have been used to study the spread of social behavior among agents of a population, such as information diffusion, social influence, and participation to collective action (e.g., protests). Key players, which are typically high-degree, -k-core or -centrality agents in a networked population, are considered important for spreading social contagions. In this paper, we ask whether contagions can propagate through a population that is void of key players. We use Erdős-Rényi random graphs as a representation of unstructured populations that lack key players, and investigate whether complex contagions—those requiring reinforcement—can spread on them. We demonstrate that two game-theoretic contagion models that utilize common knowledge for collective action can readily spread such contagions, which is a significant difference from classic complex contagion models. We compare contagion dynamics results on unstructured networks to those on more typically-studied, structured social networks to understand the role of network structure. We test a total of 14 networks. The two common knowledge models are also contrasted to understand the effects of different modeling assumptions on dynamics. We show that under a wide range of conditions, these two models produce markedly different results.**

***Keywords*—common knowledge; collective action; key players; social modeling; complex contagion**

I. Introduction

A. Background and Motivation

Contagion processes that include the transmission of emotions, information, and influence among members of networked populations have been empirically identified on Digg, Twitter, and Facebook [1], [2], [3]. By and large, these processes have been modeled using **unilateral** models of contagion spreading, such as independent cascade, threshold, linear threshold, and variants [1], [4], [3]. In these models, which we refer to throughout as **classic diffusion** (CD) models, agents (vertices in a social network) are influenced by their immediate neighbors (with whom they form edges). Each agent unilaterally makes an individual decision whether to change state from an *inactive* state 0 to an *active* state 1 based on the states of his/her neighbors. We focus on **complex contagions**, where an agent v requires multiple neighbors to be in the active state in order to change to state 1 [5]. The minimum number of influencing neighbors in state 1 required by v to change state is given by its **threshold** θ_v.

Social networks over which these contagions spread are often characterized by heavy-tailed degree distributions (e.g., scale-free (SF) or exponential decay (ED) distributions) [6], [7]. In the social networks domain, these highly connected agents are called **hubs**, a type of **key player**. Other types of key players are agents with high betweenness or closeness centralities, or k-cores [8]. McAdams [9] and others have argued that contagions initiate with these key players, and algorithms exist to identify them in social networks [10]. Most of the above types of contagion studies have been performed on networks with one or more types of key players.

The central question driving our work is the following: ***can complex contagions initiate and propagate through a population that is void of key players?***

We use agent-based simulation to study two classes of contagion models: a class of CD models and a class of game-theoretic models that we refer to as **common knowledge** (CK) models. CD models, as stated previously, are unilateral models. CK models, in contrast, are **coordination**-based models in which each agent makes a decision, with the potential to achieve mutual benefits only if her decision is consistent with those of others. That is, agents may **cooperate** to change state **jointly**. This collective behavior requires not only that agents know about each other (e.g., their thresholds and states) but also that this information is *common knowledge*. We evaluate two CK models: the Chwe [11] model and the common knowledge on Facebook (CKF) [12] model. All models are described in Section III below.

Although we study heavy-tailed social networks, we primarily investigate contagion spreading on Erdős-Rényi (ER) graphs. This is because ER networks are arguably the canonical network type that is void of structure and key players. ER networks exhibit the following trait: a large fraction of agents possess maximum (or near maximum) degree, betweenness centrality, and k-core (e.g., [13] and Figure 1 below). Hence, key players—distinguished agents with appreciably greater degrees, centralities, or k-cores than most other agents—are absent in ER networks. ER networks represent those that might form among random people who meet in a public area (e.g., city center, park) to interact.

The question then becomes whether complex contagions can propagate through populations represented as ER networks. A simple argument shedding light on threshold-based CD models is the following. Suppose the probability p_e that an edge exists between two arbitrary agents of an ER network is $p_e = 0.0004$, which is realistic for many reasonable numbers of agents and average degrees of social networks. Then, in a unilateral threshold model [5], suppose a person u in state 0 has $\theta_u = 2$, meaning that it must form edges with two agents already in state 1 in order to change to state 1. Since edges in an ER graph independently exist, the probability that u is

attached two agents (regardless of state) is $p_e^2 \approx 10^{-7}$. Hence, it can be difficult to propagate complex CD contagions in ER networks. Experimental data support this argument [14]. These results motivate the study of CK models as alternative contagion models.

B. Contributions

A summary of our major contributions follows.

1. Common knowledge models (Chwe and CKF) provide mechanisms to initiate and propagate contagions in ER networks (that lack key players) for much greater thresholds than classic diffusion (CD) models. This finding answers affirmatively the fundamental question posed in this work—whether CK models can produce wide-spread contagion transmission in networks without key players. As an example, in ER graphs with $n = 10^4$ vertices and $p_e = 0.001$, the maximum agent thresholds that can produce cascades for CD, CKF, and Chwe models are 2, 22, and > 1000, respectively. This is the first paper that uses CK models to address key player problems. For two of the three sets of simulation results (Section V), on both ER and heavy-tailed networks, there is no contagion spread based on the CD model.

2. There are significant differences in the behaviors of the Chwe and CKF models over three classes of networks: scale-free (SF), exponential decay (ED), and Erdős-Rényi (ER) graphs. We demonstrate through simulations and explain via model mechanisms why the Chwe model more readily propagates contagions: the model assumes direct edges are formed between a node and its distance t edges at times $t \geq 1$. Differences between the two models are smallest for power law or SF networks, and greatest for ER networks, for the conditions examined in this work. Furthermore, the Chwe model is essentially insensitive to network structure over all three network classes, in terms of final fraction of active agents. By comparison, contagion with the CKF model is affected by different network classes.

3. The maximum threshold that can produce widespread contagion propagation in ER networks is relatively insensitive to changes in edge probability, p_e, for the CD and CKF models. However, for the Chwe model, the maximum threshold changes significantly. In the Chwe model, edge probability, p_e, for the ER networks affects the maximum thresholds that can generate widespread contagion early in the contagion process. This is related to the changing network assumption in the Chwe model, which we characterize in Section IV. The Chwe and CD models produce similar results if the network is assumed to be constant over time.

II. Related Work

There is a vast number of contagion models where agents make unilateral decisions to change state based only on the current states of their neighbors, which we refer to as classic diffusion (CD) models herein. Several models are given in [12], [14], [15]. Others include those based on burstiness [16] and other factors [3]. Multiple interacting contagions on Twitter are modeled in [15], which considers a URL in a post as a contagion.

Contagion models of Facebook include modeling news feeds, which is a broadcast mechanism similar to that of Twitter [17], [2]. Although Facebook walls are studied experimentally [18], we are not aware of any other study (besides the CKF model) that models the Facebook wall mechanism.

ED and SF degree distributions arise frequently in social networks [6]. Various CD studies evaluate network structure; e.g., lattice and SF networks [19], and highly clustered networks [20]. Here, we use networks with SF, ED, and ER degree distributions to study contagion dynamics that are driven by CK. We know of no studies of this kind for different CK models, or of studies that compare CK models on this range of graph classes.

Chwe studies the effects of strong and weak links in his CK model [21]. A strong link is one that is part of triangle subgraphs (triadic closure). However, his arguments are based on the *initially given* social network. We demonstrate in Section IV that in Chwe's model, all edges evolve into strong edges; even the ones that are initially weak.

III. Contagion Dynamics Models

We are interested in the contagion dynamics underlying the three theoretical models within the context of a collective action problem: the Chwe and CKF common knowledge models and a CD model. We begin with preliminaries for the models. Owing to the complexity of the CK models, complete formal descriptions are not possible here. We develop the basics and refer the reader to the provided references.

Suppose there is a finite set of people $N = \{1, 2, ..., n\}$ and each person $i \in N$ chooses a state $a_i \in \{0, 1\}$, where 0 (resp., 1) is the inactive (resp., active) state. The state of i at time t is a_{it}. Each person i has an idiosyncratic private threshold $\theta_i \in \{1, 2, ..., n\}$, which is the minimum number of people that must adopt the contagion (i.e., be in state 1) for i to adopt it. Individuals in N are connected by edges in the social network G, which denote pairwise interactions. Let N_i^1 be the set of distance-1 neighbors of i in G.

A. Classic Diffusion (CD) Models [5], [1]

In the CD model, an agent i changes from state 0 to 1 based on the number of her direct neighbors that are already in state 1 (if it is greater than or equal to her threshold θ_i). This is referred to as the network-based threshold. State transition of an agent i at time t is based on the following rule:

$$a_{it} = \begin{cases} 1 & \text{if } \#\{j \in N_i^1 : a_{jt-1} = 1\} \geq \theta_i \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

A person uses two pieces of information: (i) her own threshold, and (ii) the number of her distance-1 neighbors who have already changed to the active state (namely, state 1). Individuals, then, **unilaterally** decide whether to change state. In these models, the contagion requires "seeds" (agents that are in state 1 at $t = 0$) to initiate.

B. Chwe Common Knowledge (Chwe) Model [11], [21]

The game-theoretic models of contagion consider a **coordination game** in which two or more people each make a decision to participate, with the potential to achieve shared **mutual** benefits only if their decisions are consistent. Each individual must take into account what she expects the other actors to do. In game-theoretic contexts, coordination requires that people know each others' willingness to participate, and that this information is *common knowledge* among a sufficient number of people. Common knowledge among a set of people implies that: *they know each others' thresholds (and states) and they know that everyone knows their thresholds (and states).* Therefore, they can count on each other, and jointly participate. As these models allow for mutual participation, contagion can emerge where none previously existed (i.e., "seeds" are not required for the initiation of the contagion).

Chwe [11], [21] models social structure as a communication network through which every person i tells her neighbors her willingness to participate, i.e., her θ_i and state a_{it}. The communication network helps coordination by creating common knowledge at each discrete time.

Given person i's threshold θ_i and everyone's states $a_t = (a_{1t}, a_{2t}, \ldots, a_{nt})$, his utility U at time $t \in \{0, 1, \ldots, T\}$ can be formulated as

$$U_{it} = \begin{cases} 0 & \text{if } a_{it} = 0 \\ 1 & \text{if } a_{it} = 1 \ \wedge \ \#\{j \in N : a_{jt} = 1\} \geq \theta_i \\ -z & \text{if } a_{it} = 1 \ \wedge \ \#\{j \in N : a_{jt} = 1\} < \theta_i \end{cases} \quad (2)$$

where $-z < 0$ is the penalty he gets if he participates and not enough people join him. Thus, a person will participate as long as he is sure that there is a sufficient number of people (in the population) in state 1. A person always gets utility 0 by staying in state 0 regardless of what others do since we do not consider free-riding problems. When he participates, he gets utility 1 if the total number of people in state 1 is at least θ_i.

The Chwe model assumes that the network itself is common knowledge, so that agents know about all communication that occurs between all members of the population. The dynamic approach in Chwe [21] assumes that every individual connects to his/her neighbors at increasing geodesic distance at each time step; this eventually results in complete network, if it is originally connected. The Chwe model utilizes a *hierarchy of cliques* as the network substructure that produces common knowledge. That is, every set M^{clique} of vertices that forms a clique in G also forms a CK set.

C. Facebook Common Knowledge (CKF) Model [12]

In the CKF model [12], the utility function is the same as that of the Chwe model (Equation 2). However, the communication technology is "Facebook-type" communication in which people write on each others' "walls." In other words, neighbors at distance-1 write their thresholds and states on each others' walls. Posts on an agent's wall are observed by all of his/her neighbors in distance-1. Consequently, person i knows about the thresholds and states of the people in his '*ball*' which is denoted by $B_i = \{j \in N_i^2\}$, where N_i^2 is the set of all neighbors within distance-2 of i. Motivated by online social networks, this allows for friend-of-friend communication. It naturally follows that only local information is known to any agent i in the network.

The network substructure that produces common knowledge in the CKF model is the biclique, $M^{biclique}$; i.e., the complete bipartite graph [12]. Nodes in the biclique jointly change states if their thresholds are all less than the size of the CK set.

Another mechanism by which agents change state $0 \rightarrow 1$ in the CKF model is through *threshold inference*. Given an agent i in state 0, and a neighbor $j \in N_i^2$ in state 1 with threshold θ_j, i can infer from j that at least $\theta_j + 1$ agents are in state 1. If $\theta_i \leq \theta_j + 1$, then i changes to state 1.

TABLE I

NETWORK CHARACTERISTICS FOR 11 ER GRAPHS AND THREE NETWORKS THAT ARE MORE TYPICAL OF SOCIAL NETWORKS. HERE, n AND m ARE THE NUMBERS OF VERTICES AND EDGES, RESPECTIVELY; p_e IS THE PROBABILITY OF AN EDGE BETWEEN TWO VERTICES IN AN ER GRAPH; d_{ave} AND d_{max} ARE THE AVERAGE AND MAXIMUM DEGREES; AND $diam$ IS THE GRAPH DIAMETER.

Network	n	p_e	m	d_{ave}	d_{max}	*diam*
FB	43,953	–NA–	182,384	8.3	223	18
NRV	769	–NA–	4,551	11.8	20	7
$SF1$	4,956	–NA–	45,031	18.2	269	8
ER^*_{FB}	43,945	1.9×10^{-4}	182,650	8.3	21	9
$ER1$-$ER10$; 10 nets	10,000	min: 0.001, max: 0.01	min: 49,590, max: 498,928	min: 9.9, max: 99.8	min: 24, max: 139	max: 7, min: 3

*This ER class represents three networks with similar properties and similar contagion dynamics, and so the other two are omitted.

IV. NETWORKS AND RESULTS

In this section, we present the networks used for this study, the rationale for their selection, and the degree distributions of selected networks. We also describe special network structure considerations for the Chwe model; they provide insights into the contagion dynamics results of Section V for this model.

Table I summarizes the networks of this study, and their properties. We use 11 ER graphs and three mined social networks. The networks in the table can be grouped in three sections. The first group includes: (i) a Facebook network (FB) [22], (ii) a high school student network from the New River Valley area of Virginia (NRV), which has an ER-like degree distribution, and (iii) a scale-free network ($SF1$) generated using the preferential attachment method in [6]. The second set contains one graph: an ER graph where the number of vertices and average degree match those of the FB network. This is to study the effects of network structure, since the FB network has an exponential decay (ED) degree distribution [6].

The last group of networks in Table I represents ten networks ($ER1$ through $ER10$) with constant n. They are generated by systematically varying the edge placement probability, p_e, from 0.001 to 0.01 in increments of 0.001. These networks are used for a parametric study to compare dynamics across models in Section V.

Figure 1 provides degree distributions for several networks of this study. For example, the degree distribution for the ER_{FB} network displays the characteristic shape of ER networks indicating that most nodes have the average degree of the network; relatively few nodes have greater and lesser degrees than the average. The NRV network has an ER-like degree distribution. The FB graph shows the classic ED degree distribution. An SF network ($SF1$) has the characteristic straight-line relation in log-log plots [6].

If we write $d_{max} = \alpha d_{ave}$, then $\alpha \approx 2$ for ER networks. In contrast, for ED and SF networks, $\alpha = 10$, 100, or higher. (See Table I and note that in [7], the Epinions network with an SF degree distribution has $\alpha = 223.6$.) This is what we mean by the existence or non-existence of key players (here, with respect to high-degree agents): d_{max} in ER networks is not distinguished with respect to d_{ave}, relative to other types of networks that are more typical of social networks (e.g., ED and SF). See [13] for comparable betweenness data across network classes (not the particular networks studied here); we have analogous k-core data for these networks (not shown for space limitations).

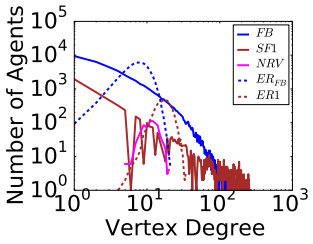


Fig. 1. Degree distributions for selected graphs of this study.

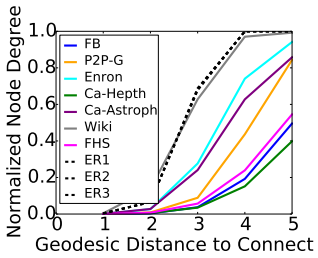


Fig. 2. Geodesic distance over which direct edges are formed between vertices in the Chwe model.

Figure 2 describes how the network structure changes over time following the dynamics in the Chwe model described in Section III. The figure is generated with the following procedure. Given a graph G, each vertex is connected directly to its neighbors of geodesic distance 1. We call this graph $G^1 (= G)$. For each vertex v, we identify the set S^k of vertices within geodesic distance k of v in the original graph (i.e., those vertices that are within k hops of v in the original graph). An edge is placed between v and each vertex in S^k, and redundant edges are eliminated to maintain a simple graph. This generates graph G^k. Graphs G^k, for $2 \leq k \leq 5$, are constructed for each network in this work. For each G^k, we compute the average fraction of vertices to which each vertex is connected; i.e., the normalized average degree in G^k. In Figure 2, we plot data for $ER1$ through $ER3$ and FB, shown in Table I, as well as several mined social networks from [7]. An ordinate value of 1.0 means that each vertex of the graph is connected to every other vertex of the graph; i.e., the original graph has transformed into a clique. Given a geodesic distance, the ER networks have greater connectivity than almost all of the other networks. This is because edges are placed randomly in ER networks, meaning that there are long-range edges in ER graphs; e.g., compare the diameters for FB and ER_{FB}. These long-range edges aid the placement of new edges in the Chwe model.

This construction is significant because it is part of the *dynamics* of the Chwe model. That is, at each successive time step, a vertex increases its knowledge of the states and thresholds of other vertices at an increased geodesic distance of 1. Thus, because initial conditions correspond to $t = 0$ and G^1, and the dynamics at $t = 1$ are computed on G^1. At time t, the graph on which dynamics are computed is G^t. From Figure 2, we see that by time $t = 5$, most graphs—and particularly ER graphs—are very highly connected and many cliques are formed. This affects contagion dynamics.

V. Simulation Results: Contagion Dynamics

This section provides results of simulations that are conducted to capture the contagion dynamics based on the three models described in Section III. In the first subsection, we describe simulation parameters. Subsequent subsections present results that justify the contributions listed in Section I-B.

A. Test Conditions and Simulation Procedure

We introduce stochastic behavior to the models represented by *participation probability*, p_p, which is the probability that an agent is participating (not absent) in the contagion process at each time step. The behavior of a participating agent follows the models as described. If an agent is not participating (absent), then in the CD model and Chwe models, it is (temporarily) removed from the social network at that time (but returns for the next time if it is now participating). In the CKF model, since a non-participating person's Facebook wall still exists, the structure of the bicliques in which he participates remains intact, but the agent himself is not included among the nodes in the CK set for that discrete time step.

The thresholds θ and participation probabilities p_p used for simulations are listed in Table II. For each pair of (θ, p_p) conditions, one simulation is executed. One simulation is comprised of 50 separate diffusion instances (i.e., runs). Diffusion instances are run for 30 time steps each to capture the early stages of diffusion. For each instance at each discrete time t, Equation (1) or (2), depending on the model being run, is computed for all nodes in state 0 to determine which nodes transition to state 1. We use a high and a low threshold in the simulations. The low threshold θ_l is equal to the average degree in the network. This means that in order for vertices in a CK set to transition from state 0 to 1, the size of the CK set must be at least $d_{ave} + 1$. The high threshold θ_h is chosen as the theoretical maximum threshold that will produce diffusion in the CKF model (even when running contagion diffusion in the other models), which is d_{max}.

TABLE II
Simulation Matrix for Each Network. Each Vertex is Assigned the Same θ and p_p in a Simulation.

Thresholds θ	**Participation Probabilities p_p**
$\theta_l = d_{ave}$, $\theta_h = d_{max}$	0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 1.0

The CD model requires some nodes to be in state 1 ("seeds") for the initiation of contagion. We use the seeding method described in [5]: for each diffusion instance we pick one vertex and assign it and all of its distance 1 neighbors the initial state of 1, with all other vertices in state 0. In contrast, the common knowledge models (Chwe and CKF) require no seed vertices because individuals can coordinate to change state from 0 to 1 jointly, and hence all vertices are initially in state 0 at $t = 0$.

B. Comparison of Models on Facebook and ER Networks

Figure 3 contains contagion histories for ER_{FB} and FB, which nominally match in n and d_{ave}. Curves are the time-wise averages over the 50 runs. The first two plots are for $\theta = d_{ave} = 9$, for the Chwe model and CKF model, respectively. The latter two are plotted for $\theta = d_{max} = 21$, for Chwe and CKF, respectively. There is no contagion spreading in CD model for these networks.

These plots contain the fraction of agents in state 1 as a function of time. Each curve represents diffusion for different participation probabilities p_p. The colors corresponding to these p_p are: 1.0 (magenta), 0.4 (gray), 0.3 (purple), 0.2 (green), 0.1 (cyan), 0.05 (orange), and 0.01 (blue).

First, each of the four plots shows appreciable contagion spread for at least some conditions (and in some cases, for all conditions) in the ER_{FB} network (solid curves). These curves demonstrate that **common knowledge models can propagate contagion in hubless, unstructured networks**, which is Contribution 1. By inspection, since the same sets of conditions are used with both the Chwe and CKF models, we see immediately that **the Chwe model propagates contagion over a broader range of conditions, compared to the CKF model**, which is part of Contribution 2. For example, for $p_p = 0.05$ and $\theta = 9$, the Chwe model (results in Figure 3a) generates significant contagion, while the CKF model (results in Figure 3b) produces less contagion.

In Figures 3a and 3c, the Chwe model produces about the same dynamics on FB and ER_{FB} networks; the solid and corresponding dashed lines are close. That is, **the Chwe results are relatively insensitive to network class**, which is also part of Contribution 2. This is due to the growing network that is assumed in the Chwe model, which generates more cliques and more CK over time.

By contrast, the CKF model dynamics depend on local structure, and hence Figure 3d shows a large difference between ER and ED graph classes (solid versus dashed curves). The reason that contagion initiation is difficult in ER_{FB} for the CKF model is that there is one star subgraph with 22 agents: the one whose hub or center node has degree 21. Thus, for contagion to initiate, all 22 agents must be participating at the same time. This is a stringent requirement for most p_p values investigated.

C. Comparison of Models Across SF, ED, and ER Classes

The box plots in Figure 4 depict the cumulative fraction of nodes in state 1 at $t = 30$ for each of the 50 instances for Chwe and CKF models. The three graph classes are also provided.

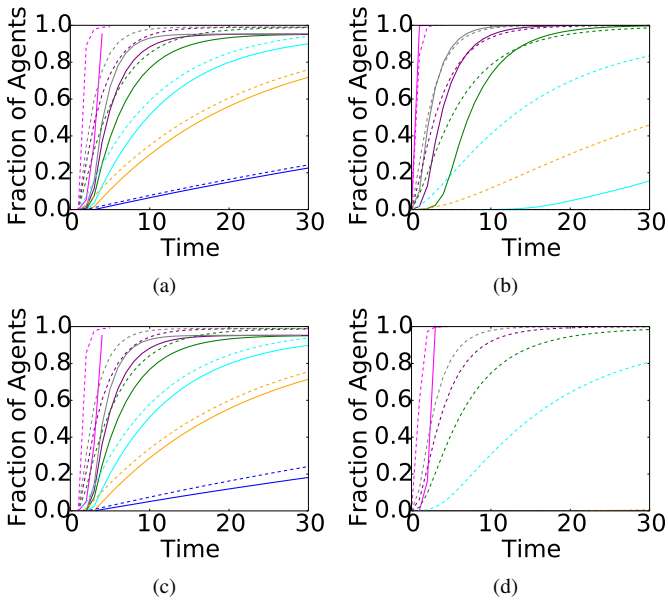


Fig. 3. Comparison of diffusion dynamics using the Chwe model in (a) and (c), and the CKF model in (b) and (d), In each figure, results for ER_{FB} (solid curves) and FB (dashed curves) are given. Both networks have approximately the same n and d_{ave}, but d_{max} in FB network is almost an order of magnitude greater. In (a) and (b), the uniform threshold is $\theta = d_{ave} = 9$. Each curve corresponds to a different p_p; see text. In (c) and (d), $\theta = 21 = d_{max}$. There is no contagion with the CD model.

The SF network is $SF1$, the ED network is FB, and the ER network is NRV. Here, we use $p_p = 0.05$, with uniform thresholds $\theta = d_{ave}$, where d_{ave} is based on the respective network. We observe that **the Chwe model is essentially unaffected by network structure, whereas the CKF model is far more dependent on the network structure**, which is part of Contribution 2. Also part of that contribution is that **results between the two models are in closest agreement for SF networks**. This is because SF networks have more hub nodes of greater degree, which assist the CKF model in initiating contagion at multiple locations in an SF network, when thresholds are less than d_{max}. For greater p_p, the CKF model generates contagion in the NRV (right-most) network. We do not obtain any diffusion with the CD model.

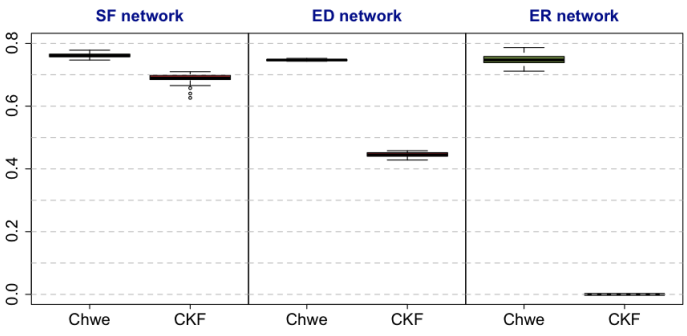


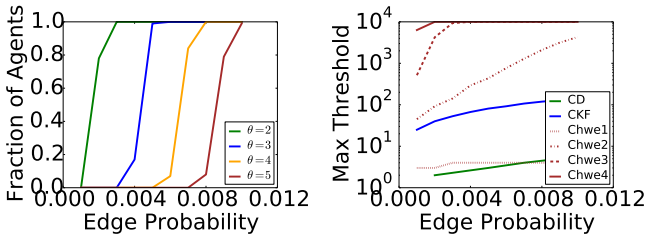
Fig. 4. Box plots illustrating the cumulative fraction of nodes in state 1 at $t = 30$ for each of 50 instances for Chwe and CKF models. The participation probability is $p_p = 0.05$. $SF1$ (left), FB (center), and NRV (right) networks are shown. The threshold is d_{ave} of the respective networks. There is no CD diffusion for these conditions.

The qualitative results in Figure 4 hold over a broad range of θ and p_p. Differences between Chwe and CKF models increase as p_p decreases and threshold increases.

D. Parametric Study of Results For Different ER Networks

Here, we study the series of ten ER networks ($ER1$ to $ER10$). For each of these networks, simulations are run where threshold is systematically varied to identify the maximum threshold that will propagate contagion. In all simulations, we use $p_p = 1.0$. Figure 5a shows the results for CD. Each curve corresponds to results for a particular threshold. For example, when $p_e = 0.004$, about 15% of vertices, on average, change to state 1 for $\theta = 3$. For $p_e = 0.005$, 99% of nodes, on average, change to state 1 when $\theta = 3$. Hence, we take $(0.005, 3)$ as the (p_e, θ) pair that generates widespread diffusion (i.e, a cascade) in these ER networks. Similar results for other (p_e, θ) pairs can be gleaned from this plot. This process is repeated for the Chwe and CKF models, and results are plotted in Figure 5b. This latter plot shows the maximum allowable threshold θ_{all} that will produce ≥ 0.9 fraction of nodes in state 1 at $t = 30$.

Four curves for the Chwe model are provided, one for each of G^1 through G^4, in Figure 5b, to account for the increasing number of graph edges in time, and one curve for each of CD and CKF. The maximum allowable threshold in the Chwe model increases as additional edges are added to the base (original) graph G^1. **For G^1, if we ignore the fact that agents learn their neighbors' thresholds and states at incrementally greater geodesic distance as a simulation progresses, then the thresholds that the Chwe model can sustain are comparable to those for CD** (Contribution 3). This plot clearly shows that the common knowledge models can propagate contagions in ER (unstructured, hubless) networks at much greater thresholds than can CD across a wide range of ER graphs (by varying p_e). Furthermore, as stated in Contribution 3, **for CD and CKF models, θ_{all} changes relatively little with increasing p_e, but for the Chwe model, θ_{all} can increase significantly with p_e.** See, for example, the Chwe2 and Chwe3; i.e., G^2 and G^3, curves, respectively.



(a) (b)

Fig. 5. The effect of edge probability p_e in generating full cascades in 10^4-vertex ER graphs. Here, p_e ranges from 0.001 to 0.01 in 0.001 increments. (a) Results for CD only. Across a set of networks, when all vertices are assigned the same threshold (see legend), the fraction of agents that reach state 1 are shown. (b) Results for all three models: CD, CKF, Chwe. Data for CD are derived from (a); data for CKF and Chwe are derived theoretically from the maximum sizes of CK sets.

VI. Conclusions

This paper analyzes two contagion models that incorporate different common knowledge (CK) mechanisms, enabling agents to coordinate their actions. Since classic diffusion (CD) models demonstrate that complex contagions generally do not propagate in unstructured networks that lack key players, our main goal is to determine whether CK mechanisms will propagate contagions in such networks. We find that CK mechanisms can drive complex contagions in networks that lack key players, but that the two CK models can produce vastly different contagion dynamics on networks void of key players. These data suggest that fostering CK among people may be one way to produce or increase contagion.

Acknowledgments. We thank the anonymous reviewers for their useful comments and suggestions. We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL) and the Social and Decision Analytics Laboratory (SDAL). This work was partially supported by DTRA Grant HDTRA1-11-1-0016, DTRA CNIMS Contract HDTRA1-11-D-0016-0001, and NSF NetSE Grant CNS-1011769.

References

[1] S. Gonzalez-Bailon, J. Borge-Holthoefer, A. Rivero, and Y. Moreno, "The dynamics of protest recruitment through an online network," *Scientific Reports*, pp. 1–7, 2011.
[2] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *WWW*, 2012.
[3] N. O. Hodas and K. Lerman, "The simple rules of social contagion," *Scientific Reports*, vol. 4, 2014.
[4] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the Spread of Influence Through a Social Network," in *KDD*, 2003, pp. 137–146.
[5] D. Centola and M. Macy, "Complex contagions and the weakness of long ties," *Am. J. of Soc.*, vol. 113, no. 3, pp. 702–734, 2007.
[6] A. Barabasi and R. Albert, "Emergence of scaling in random networks," *Nature*, vol. 286, pp. 509–512, 1999.
[7] J. Leskovec, "SNAP Datasets," http://snap.stanford.edu/data, 2016.
[8] R. C. Gunasekara, K. Mehrotra, and C. K. Mohan, "Multi-objective optimization to identify key players in social networks," in *ASONAM*, 2014, pp. 443–450.
[9] D. McAdam, "Recruitment to high-risk activism–the case of freedom summer," *American Journal of Sociology*, vol. 92, pp. 64–90, 1986.
[10] S. P. Borgatti, "Identifying sets of key players in a social network," *Comput Math Organiz Theor*, vol. 12, pp. 21–34, 2006.
[11] M. S.-Y. Chwe, "Structure and strategy in collective action," *Am. J. of Soc.*, vol. 105, pp. 128–156, 1999.
[12] G. Korkmaz, C. J. Kuhlman, A. Marathe, M. V. Marathe, and F. Vega-Redondo, "Collective action through common knowledge using a facebook model," in *AAMAS*, 2014.
[13] D. A. Bader, S. Kintali, K. Madduri, and M. Mihail, "Approximating betweenness centrality," in *WAW*, 2007, pp. 124–137.
[14] C. J. Kuhlman *et al.*, "Inhibiting diffusion of complex contagions in social networks: Theoretical and experimental results," *DMKD*, 2015.
[15] S. Myers and J. Leskovec, "Clash of the contagions: Cooperation and competition in information diffusion," in *ICDM*, 2012.
[16] T. Takaguchi, N. Masuda, and P. Holme, "Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics," *Plos One*, vol. 8, pp. e68 629–1–e68 629–5, 2013.
[17] E. Sun, I. Rosenn, C. A. Marlow, and T. M. Lento, "Gesundheit! modeling contagion through facebook news feed," in *ICWSM*, 2009.
[18] P. Devineni, D. Koutra *et al.*, "If walls could talk: Patterns and anomalies in facebook wallposts," in *ASONAM*, 2015, pp. 367–374.
[19] V. Barash, C. Cameron, and M. Macy, "Critical phenomena in complex contagions," *Social Networks*, vol. 34, 2012.
[20] G. Schoenebeck, "Potential networks, contagious communities, and understanding social network structure," in *WWW*, 2013.
[21] M. S.-Y. Chwe, "Communication and coordination in social networks," *Rev. of Econ. Stud.*, vol. 67, pp. 1–16, 2000.
[22] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, "On the evolution of user interaction in facebook," in *WOSN*, 2009.