AN AGENT-BASED MODEL OF COMMON KNOWLEDGE AND COLLECTIVE ACTION DYNAMICS ON SOCIAL NETWORKS

Chris J. Kuhlman
S. S. Ravi

Biocomplexity Institute
University of Virginia
Charlottesville VA 22904, USA

Gizem Korkmaz

Biocomplexity Institute
University of Virginia
Arlington VA 22209, USA

Fernando Vega-Redondo

Department of Decision Sciences
Bocconi University
Milan, ITALY

ABSTRACT

Protest is a collective action problem and can be modeled as a coordination game in which people take an action with the potential to achieve shared mutual benefits. In game-theoretic contexts, successful coordination requires that people know each others' willingness to participate, and that this information is common knowledge among a sufficient number of people. We develop an agent-based model of collective action that was the first to combine social structure and individual incentives. Another novel aspect of the model is that a social network increases in density (i.e., new graph edges are formed) over time. The model studies the formation of common knowledge through local interactions and the characterizing social network structures. We use four real-world, data-mined social networks (Facebook, Wikipedia, email, and peer-to-peer networks) and one scale-free network, and conduct computational experiments to study contagion dynamics under different conditions.

1 Introduction

1.1 Background and Motivation

The significant impacts of civil unrest events (e.g., Arab Spring, Gezi protests, Occupy Wall Street) have motivated the construction of models to understand the mechanisms by which information spreads through social networks, and their consequences. Protest is a collective action problem where an individual wants to protest only if joined by "enough" others in order to mitigate the risk of prosecution. Game theorists often model the collective action problem as a *coordination game* in which two or more people make decisions to participate, with the potential to achieve mutual benefits only if their decisions are consistent. In game-theoretic contexts, coordination requires that agents know about each other's willingness to participate (defined as "threshold") and that this information is *common knowledge* among a sufficient number of people. Common knowledge (CK) is defined as an infinite string of embedded levels of knowledge (i.e., they know each other, they know that they know, and so on.) (Lewis 1969; Schelling 1960).

To choose among multiple options, an individual takes into account others' strategies, which is in turn dependent upon their expectations of what she will do, leading to interdependent expectations that generate an infinite recursion of embedded beliefs (Thomas et al. 2014). In interactive situations, knowledge about

others' knowledge is important. If I want to participate, but I do not know whether you know my intentions, then I do not participate. This is because I do not expect you to participate without sufficient information (that I want to participate if you do).

Social networks that represent local interactions (e.g., Facebook, Twitter, YouTube) may facilitate information sharing in a way that generates CK within groups. CK is important to study because it "has a strong presence in human life" and appears in everyday language that describes concerns such as publicity, privacy, confidentiality, conventional wisdom, hypocrisy, mock outrage, and political correctness (Thomas et al. 2014). So too, mass cultural practices such as ceremonies and rituals, and widespread advertising, can be understood in terms of CK (Chwe 1998).

Although we use protest as the exemplar for CK, there are many other situations in which groups make decisions, i.e., group members make coordinated decisions. These include hiring decisions by managers (Wittenbaum et al. 1998), work task decisions by project teams (Gladstein 1984), community decisions by homeowners (Bond and Keys 1993), and evacuation decisions during natural disasters (e.g., hurricanes) (Nguyen et al. 2017). The considerations and results in this work also apply to these other contexts.

1.2 Chwe Model Overview

Chwe's CK model (Chwe 1999; Chwe 2000) is the first to combine social structure and individual incentives together in a game-theoretic model of collective action to study and provide a rigorous formalization of CK, and the characterizing network structures. In this model, social structure is modeled as a communication network through which every person tells her direct (distance-1) neighbors her willingness to participate, represented by her *threshold*. The communication network helps coordination by creating CK at each discrete time, and informing each stage about earlier stages. Chwe's CK model has the following features: *(i) cliques* in a network are found as the characterizing network structures that are required for the formation of CK, and hence the emergence and the spread of collective action, i.e., groups of agents *simultaneously* transitioning to the active state (similar to the infected state in a contagion model); *(ii)* within these cliques, nodes can change from state 0 (the non-participating state) to state 1 (active state), if the node thresholds are sufficiently small (see details of the Chwe CK model in Section 3); and *(iii)* an original graph's density increases over time as edges between distance-2 neighbors are added (while the set of nodes remains unchanged), resulting in progressively more numerous and larger cliques, the CK-characterizing substructure (Section 3.1). The assumption of successive increases in graph density is introduced based on a friend-of-friend argument. That is, at any time, a node is connected to a set of neighbors, and a node can then come to know its distance-2 neighbors through common friends.

Figures 1 and 2 provide toy examples to illustrate these novel features of the Chwe CK model, in contrast to the widely-used classic threshold models, e.g., Granovetter (1978), to understand the spread of collective action. The formal Chwe model is provided in Section 3. Figure 1 demonstrates the state transition mechanism providing two examples with four agents that form a clique of size 4. Each agent i has a threshold, θ_i, which represents the number of additional people needed to transition to the active state for agent i to benefit from doing so. Figure 1(a) illustrates the case where all four agents simultaneously transition to state 1 since (i) they form a CK set (clique) and hence they know each other's thresholds (and they know that they know, and so on), and (ii) their thresholds are low enough for them to collectively participate and achieve mutual benefits. Figure 1(b) illustrates a case in which only a subset of agents that form a CK set (clique) can transition to state 1. Details of the mechanism are explained in the caption.

Figure 2 illustrates the graph densification process in the Chwe model using an example with four agents that form a path graph (a connected graph, where two end nodes have degree 1 and all internal nodes have degree 2). At each time step, edges between distance-2 neighbors are added which results in two cliques of size 3 at $t = 2$ and a clique of size 4 at $t = 3$. As a result of graph densification, a subset of agents (those with threshold 2) can simultaneously transition to state 1 at $t = 2$, while the agent with threshold 4 stays in state 0 due to its high threshold.

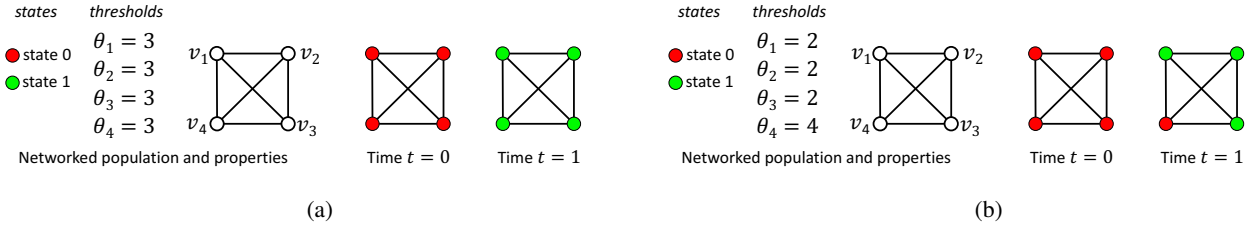(a)   (b)

Figure 1: Toy examples to illustrate the novel state transition mechanism of Chwe CK model. In both examples there are four agents that form a clique of size 4. Green represents the active state (state 1) and red represents state 0. All agents are in state 0 at $t = 0$. For illustrative purposes, the state transitions in this example are deterministic. **(a)** The thresholds are given as $\theta_{1-4} = 3$, hence each agent needs three other agents to transition to the active state (i.e., state 1, shown in green). They all activate. **(b)** Agents 1, 2, and 3 form a sub-clique (as part of the larger clique) and their thresholds are are low enough ($\theta_{1-3} = 2$) for each of them to transition to state 1 at $t = 1$. In this case, agent 4 needs four other individuals to participate with him ($\theta_4 = 4$), and hence he cannot activate.
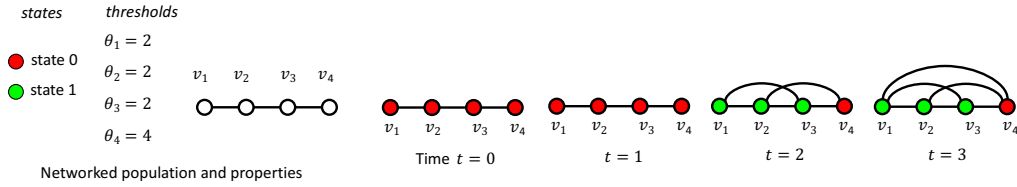

Figure 2: Toy example of the graph densification in Chwe CK model. There are four agents that form a path graph of size 4. At each time step, edges between distance-2 neighbors are added. All agents are in state 0 at $t = 0$. For illustrative purposes, the state transitions in this example are deterministic. No transitions occur at $t = 1$. At $t = 2$, as a result of graph densification, two cliques of size 3 are formed and agents with threshold 2 transition to state 1. At $t = 3$, a larger clique of size 4 is formed. In spite of this, the agent with threshold 4 stays in state 0 at $t = 2$ since he needs four other agents to transition.

1.3 Contributions and Novelty of This Work

The Chwe CK model is the only contagion model where an original graph's density increases over time as agents connect to their friends of friends (its distance-2 neighbors). The model has been applied to path graphs and other small, simplified graphs (Chwe 1999; Chwe 2000). It has not been used to study coordination-based contagion dynamics on realistic social networks, and the growing graph feature mentioned above—a unique feature of contagion dynamics models, to our knowledge—has not been systematically studied to determine its effect on those dynamics. These are the aims of this paper.

We develop an agent-based model based on the CK model proposed by Chwe (1999, 2000), and study the contagion dynamics (using size and speed of the collective action as the measures) under different conditions, and realistic network structures. To our knowledge, this is the first paper to study the effects of this densification mechanism on real networks. Based on our stated goals, our main contributions follow.

Realistic networks. We study contagion dynamics on mined social networks that prevail in the real world, such as Facebook, Wikipedia, and email networks, and a fifth scale-free network, a common network structure. We characterize these graphs using parameters that are useful in understanding the dynamics on these networks, including some time-dependent properties. These networks cover wide ranges of values of network measures, such as number of vertices ($6\times$), number of edges ($4\times$), average degree ($4\times$), maximum degree ($13\times$), and average clustering coefficient ($80\times$), and hence represent a broad sampling of web-based mined network features. Some of these are very large networks when considering the requirement to find cliques in networks, as described in Section 5. For two of the networks, this process required over 100 hours of compute time on a Dell PowerEdge C6420, 2666 MHz, cluster where each node is an Intel Xeon Gold 6148, 2.40 GHz processor, with three cache levels.

Effects of network growth mechanisms on the contagion dynamics. We present the Chwe CK model and then the concomitant agent model used within an agent-based modeling (ABM) platform to run

simulations. Main findings from the simulations follow. *(i)* The speed of contagion spread can be very high as a result of graph densification. When a Facebook graph grows over the first five time steps of a simulation, the fraction of agents in state 1 (active) can go from 0 (initially) to 1.0. When the graph grows for only three time steps, the spreading is very similar. *(ii)* In all five networks, contagion growth is very little in the original graphs. It is only when we consider the square of the original graph (defined below) that contagion starts to spread. *(iii)* When node thresholds are equal to the average graph degree, a contagion can spread through the graph in the first two time steps. When node thresholds are increased to the maximum graph degree, complete spreading (i.e., a cascade) can require three time steps, where the graph evolves to the cube of the original graph. *(iv)* This last result is robust to a range of participation rates (specified by probabilities of nodes, ranging from 0.1 to 1.0). *(v)* As the ratio of number of nodes to maximum degree decreases in a graph, the more difficult it is to spread high-threshold contagion at lower participation probabilities. This is because it is difficult to have enough nodes participating, at any one time, to overcome the node thresholds.

Chwe CK model's assumption that a node knows progressively more neighbors in their network, results, over time, in connections to nodes that are originally more distant in the original social network. These results demonstrate that this assumption is powerful in initiating contagion where none existed and in subsequently driving this contagion through a social network, often at high rates of transmission.

2   Related Work

Common knowledge (CK) has come up in many different scholarly contexts, from the philosophy of language to game theory to sociology. Aumann (1976) originally used his definition of CK to prove a result that says that agents cannot "agree to disagree" about their beliefs, formalized as probability distributions, if they start with common prior beliefs. Rubinstein (1989) developed a model and showed that under a specific and restrictive set of assumptions, common knowledge is necessary for coordination. Subsequent work has suggested that shared knowledge can also facilitate coordination (Binmore and Samuelson 2001; Monderer and Samet 1989; Dalkiran et al. 2012). The problem of coordination and common knowledge has been examined in many disciplines, including political science (Ostrom 1990), philosophy (Lewis 1969; Hume 2003; Skyrms 2004), economics (Chwe 2013; Geanakoplos 1994), sociology (Willer et al. 2009), and linguistics (Clark 1992; Clark 1996). Purely experimental studies of CK have been undertaken from a social science perspective; e.g., (Thomas et al. 2014; Korkmaz et al. 2018).

Game-theoretic models of CK and collective action (Chwe 1999; Chwe 2000; Korkmaz et al. 2014) allow small and large groups of agents to *coordinate* their actions, thereby enabling them to transition state simultaneously if it is mutually beneficial to do so. CK emerges within a group when each member knows the states and attributes (e.g., preference, type) of the other members, and critically, each member knows that everyone else knows her attributes.

3   The Model

This section provides a formal description of the CK model proposed by Chwe (1999, 2000), referred to as the Chwe CK model, that we study in this paper. The first four subsections describe model components. Subsection 3.5 provides an algorithm for the full Chwe model. Examples of the Chwe model dynamics are provided in Subsection 3.6. Our presentation is both agent-centric and clique-centric because this aids in understanding the model. However, the entire model can be described solely in terms of individual agent behaviors (and hence is an ABM), and we address this issue at the end of Section 3.5.

3.1 Graph Densification

The initial network involves a population of individuals represented by $G(V,E)$ with a node set $V = \{1, 2, \ldots, n\}$ of n nodes (people) and edge set E, where an undirected edge $\{i, j\} \in E$ means that nodes $i, j \in V$ can communicate with each other. The Chwe CK model specifies a process whereby the initial social network $G(V,E)$ grows in time by adding edges at successively greater geodesic distances. Since the node set V is fixed, the graph density increases, and hence this process is also referred to as *graph densification* or *growing graph (GG)* mechanism. The instantaneous network $G^t(V,E^t)$ at time t is defined

over the fixed node set V, where for each $i, j \in V$, with $i \neq j$, the edge $\{i, j\} \in E^t$ iff in G (the original graph), $d(i, j) \leq t$. Here, $d(i, j)$ is the geodesic distance between nodes i and j in G. Thus, at $t = 1$, $G^1 = G$, the initial graph. For a given edge $\{i, j\} \in E$, consider a node $k \in G$, where the shortest path from i to k is through j, such that $d(i, k) = q \geq 2$. Then in G^q, a new edge $\{i, k\}$ will be added.

3.2 State Transition

This model formalizes the process of a node transitioning from the inactivated state 0 to the activated state 1. The Chwe state transition model is a *progressive* model (Kempe et al. 2003) in that the only permissible node state transition is from state 0 to state 1. Once a node is in state 1, it remains in that state; it cannot transition back to state 0 at a later time.

Chwe (1999, 2000) models collective action as a *coordination game* in which two or more people each make a decision to activate, with the potential to achieve shared *mutual* benefits only if their decisions are consistent. Individuals can be in one of two states at each time t: if i activates at t, then $a_{it} = 1$ for all times $\geq t$; otherwise, $a_{it} = 0$. Each person i has a threshold θ_i that represents his willingness to activate: a person wants to activate only if the number of people (other than him) who do so is at least θ_i. Person i's utility U_{it} at $t \in \{0, 1, \ldots, t_{max}\}$ can be specified as

$$U_{it} = \begin{cases} 0 & \text{if } a_{it} = 0 \\ 1 & \text{if } a_{it} = 1 \text{ and } |\{j \in V \backslash i : a_{jt} = 1\}| \geq \theta_i \\ -z & \text{if } a_{it} = 1 \text{ and } |\{j \in V \backslash i : a_{jt} = 1\}| < \theta_i \end{cases} \tag{1}$$

where $-z < 0$ is the penalty i gets if she activates and not enough people (i.e., less than θ_i) join her, a_{jt} denotes the states of nodes j at the end of time t. Thus, a person will activate as long as she is sure that there is a sufficient number of people in state 1 at the end of t. A person gets utility 0 by staying in state 0 regardless of what others do. Individuals communicate their tuples (θ, a) with their neighbors on the network. CK *among a group of people* implies that they *know* each others' tuples (θ, a) at t and they know that everyone in the group knows these values. Based on the model, for person i to activate it must be that he belongs to a minimal sufficient network, which is characterized by a *clique* (Chwe 2000). That is, each clique subgraph $G'(V', E')$ embedded in a graph G^t identifies a set V' of nodes that form a common knowledge set. A clique subgraph with r nodes is also denoted as K_r.

3.3 Agent Participation Probability

We introduce stochastic state transition as an extension to the Chwe CK model. *Crucially, however, the extension preserves the original Chwe model by an appropriate choice of a parameter value.* The original model is deterministic: all nodes participate in the social network interactions at all times. However, it is straight-forward to envision that as an event unfolds, not all people are attentive to the event: they may be occupied with other commitments. This is consistent with other social contagion transmission models, e.g., (Kempe et al. 2003; Myers and Leskovec 2012), that introduce stochasticity to recognize that contagion is not transmitted at every step of the dynamical process. We use *participation probability* $p_{p,i}$ for $i \in V$ as the probability that a node is attentive (i.e., is paying attention) to the event (e.g., protest) at each time t. Hence, at each time t, a Bernoulli trial is performed for each node $i \in V$, and with probability $p_{p,i}$, i participates in the contagion process at t. *Note that the original (deterministic) Chwe model is easily recovered by setting $p_{p,i} = 1.0$ for all $i \in V$.* We study both cases ($p_{p,i} = 1.0$ and $p_{p,i} < 1.0$ for all $i \in V$) in this work.

3.4 Effect of Participation Probability on the Clique Graph Structure of the Chwe Model

In Section 3.2, the clique is identified as the network substructure that gives rise to CK among a group of nodes V'. It is important to note that any subset $V'' \subset V'$ of nodes still forms a clique $G''(V'', E'')$. With regard to participation probability, given a clique $K_{n_1} = G'(V', E')$ on $n_1 = |V'|$ nodes, assume that $0 \leq n_2 < n_1$ of these nodes are participating at time t. Then, the participating nodes form a *participating clique* $K_{n_2} \equiv G''(V'', E'') \subset K_{n_1}$ on these n_2 nodes. When $n_2 = 0$, then there is no participating clique, and no contagion can spread.

3.5 Dynamics of Contagion Spread in the Chwe Model

We now describe the contagion dynamics of the Chwe CK model using the foregoing model descriptions, and present an algorithm for these computations. Individuals can be in one of two states at each time t: if i possesses contagion (i.e., activates) at t, then $a_{it} = 1$; if it does not, then $a_{it} = 0$. The system state at t is $a_t = (a_{1t}, a_{2t}, \ldots, a_{nt})$. The goal is to compute the next system state $a_{(t+1)}$ at time $t+1$. As this model allows for mutual activation, contagion can emerge where none previously existed, i.e., "seeds" (nodes assigned state 1 at $t = 0$) are not required for the initiation of the contagion as they are in classic diffusion models (Granovetter 1978).

To compute $a_{(t+1)}$, the participation of each node i is computed using $p_{p,i}$. From a participating clique $K_{|M^{clique}|}$, we define M^{clique} as the set of participating nodes that share CK. The contagion dynamics computations at each t consists of evaluating Equation (1) for each $i \in V$ in each participating CK set M^{clique}. That is, V in Equation (1) is replaced with M^{clique}.

The Chwe CK contagion dynamics evolve over time in discrete time steps. A set of initial conditions at $t = 0$ is specified, which are: an initial graph $G(V, E)$ (with $G^1(V, E) = G(V, E)$), a threshold vector $\Theta = (\theta_1, \theta_2, \ldots, \theta_n)$ of node thresholds θ_i for all n nodes $i \in V$, a participation probability vector $P = (p_{p,1}, p_{p,2}, \ldots, p_{p,n})$ for all nodes, and an initial system state vector $a_0 = (a_{1,0}, a_{2,0}, \ldots, a_{n,0})$ for the n nodes. The dynamics are then computed at each time step in a sequence $\langle 1, 2, \ldots, t_{max} \rangle$ of times, up through a specified maximum time t_{max}. At each time, Algorithm 1 is invoked. Inputs to the algorithm at time t include $G^t(V, E)$ and a_t, which are computed with the call to the algorithm at time $(t-1)$.

Algorithm 1: Steps in the Chwe Common Knowledge Model at Time t.

1 **Input:** Time t. Instantaneous social network $G^t(V, E)$. Collection Θ of node thresholds $\Theta = (\theta_1, \theta_2, \ldots, \theta_n)$, where θ_i is the threshold for all $i \in V$. The vector $P = (p_{p,1}, p_{p,2}, \ldots, p_{p,n})$ of participation probabilities $p_{p,i}$ for all $i \in V$. System state $a_t = (a_{1,t}, a_{2,t}, \ldots, a_{n,t})$ at time t.

2 **Output:** The densified graph $G^{(t+1)}(V, E)$ and the updated system state $a_{(t+1)}$, both at time $t+1$.

3 **Steps:**

    A.   Form the densified graph $G^{(t+1)}(V, E)$ from $G^t(V, E)$ by adding edges to $G^t(V, E)$ based on the model described in Section 3.1.

    B.   Compute all of the clique subgraphs $G'(V', E') \subseteq G^{(t+1)}(V, E)$.

    C.   Using the participation probabilities $p_{p,i}$ for all nodes i that are in at least one node set V' in some clique $G'(V', E')$, perform one Bernoulli trial for each i per Section 3.3. This determines which nodes are participating in each clique at t.

    D.   Form all of the participating cliques $G''(V'', E'') \subseteq G'(V', E')$, per Section 3.4. Form all of the node sets $M^{clique} = V''$.

    E.   For each M^{clique}, if there is at least one participating node $i \in M^{clique}$ such that $a_{it} = 0$, then perform the computations of the state transition model on M^{clique} in Section 3.2 to determine whether any nodes in M^{clique} transition from state 0 to state 1. Replace V in Equation (1) with M^{clique}. If a node i in state 0 transitions to state 1 in any clique, then $a_{i,t+1} = 1$; otherwise $a_{i,t+1} = 0$. Record the nodes that transition to state 1. (The order of evaluation of all M^{clique} is arbitrary.)

    F.   Form the new system state vector $a_{(t+1)}$ by updating the states of participating nodes $i \in M^{clique}$ that have transitioned to state 1, i.e., to $a_{i,t+1} = 1$.

    G.   Return $G^{(t+1)}(V, E)$ and $a_{(t+1)}$.

Recasting the model solely in terms of an individual agent model. We now specify how the Chwe model can be described in terms of individual agent behaviors. We use the **Steps** in Algorithm 1. Steps A and B can be performed by individual agents because they know their friends-of-friends (Section 1). Hence, each agent i can determine the cliques G' that they are members of at the current time step and with whom they will form direct edges at the next time step. Step C is an individual agent action. For Step D, because i knows the thresholds and states of its neighbors, and whether its neighbors are participating, i can determine the cliques G'' and node sets M^{clique} of which it is a member. Step E is already an individual agent operation: each agent i executes Equation 1. Step F merely assembles the updated node states.

3.6 Examples of the Contagion Dynamics

These examples illustrate some of the effects of the graph densification mechanism and thresholds on the contagion dynamics of the Chwe CK model. They also provide intuition for reasoning about contagion simulation results on mined networks in Section 6.

The Chwe CK model uses an initial graph whose connectivity changes over time. Figure 3 shows growing graph (GG) mechanism between $t = 0$ to $t = 5$. The original graphs at $t = 0$ and $t = 1$ are denoted as G or G^1. At $t = 2$, each node of G^1 is joined to all distance-1 neighbors of its immediate (i.e., distance-1) neighbors; the result is the graph G^2. At $t = 3$, each node of G^2 is joined to all distance-1 neighbors of its immediate neighbors; this results in G^3. And so on, until a clique K_6 on $n = 6$ nodes is formed at $t = 4$. At $t = 5$, no more edges can be added, and the graph structure no longer changes. In general, by time $t = n - 2$, a clique K_n on n nodes will be formed in a connected graph. Based on the Chwe CK model, each node will communicate with every node in the population when K_n is formed on a connected graph.
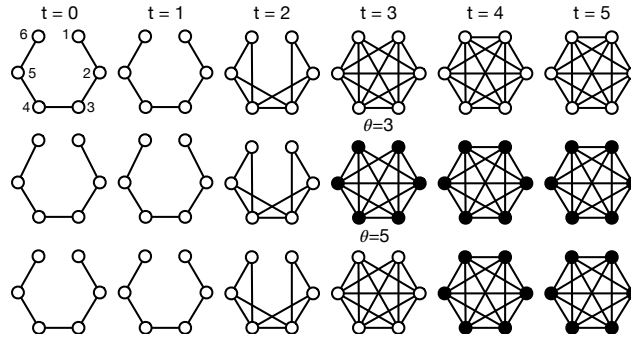

Figure 3: Illustration of how a Path$_6$ graph (upper left) adds edges over time $0 \leq t \leq 5$ in the Chwe CK model (first row). The edge additions are made according to the mechanism described in Section 3.1. Node IDs are given at the upper left (at $t = 0$) and are the same IDs in all other graphs (not shown for clarity). Rows 2 and 3 show the contagion dynamics for $\theta = 3$ and 5, respectively, for all $i \in V$; open circles are nodes in state 0 and filled circles denote state 1.

The time through which a graph grows is denoted as t_{GG}. In the simulations of Section 6, we halt network growth after a particular time t_{GG} to evaluate the effects of the GG mechanism. For example, if $t_{GG} = 3$, then this means that G^1 is used to compute contagion at $t = 1$, G^2 is used at $t = 2$, and G^3 is used at $t \geq 3$; the graph does not grow for $t > t_{GG}$. However, contagion may still spread at $t > t_{GG}$ if there exists at least one node i such that $0 < p_{p,i} < 1.0$ ($p_{p,i} = 0$ is the case with no participation and $p_{p,i} = 1.0$ is the deterministic case).

In the second row of graphs, labeled with $\theta = 3$, we show contagion dynamics when all nodes have a threshold of 3. First, let us assume that $p_p = 1$, so that the dynamics are deterministic (as in the original Chwe CK model). All nodes are in state 0 at $t = 0$. At $t = 3$, all nodes change to state 1, because for $\theta = 3$, each node i has at least θ_i neighbors in an M^{clique} that contains i. (There are two cliques of size 5, i.e., two K_5s, with node sets $M^{clique} = \{1, 2, 3, 4, 5\}$ and $\{2, 3, 4, 5, 6\}$. Note that cliques of this size do not form until $t = 3$.) All nodes in each of these sets transition to state 1 per Equation (1). Note that nodes 3, 4, and 5 are each part of both cliques and will change state as a result of their inclusion in either of the K_5.

The dynamics for $p_p = 1$ when all nodes have $\theta = 5$, are provided in the last row of Figure 3. These dynamics thus conform to the original Chwe CK model. The reasoning for these dynamics follows the same ideas as for $\theta = 3$. In this case, nodes do not transition until $t = 4$. We will see for real networks and simulations in Section 6, in contrast, that $t_{GG} \leq 3$ is often sufficient to drive the contagion.

4   The Agent-Based Model (ABM)

An agent-based model (ABM), for running simulations in an agent-based simulation (ABS) framework, was developed based on the model described in Section 3. The ABM design and software implementation

follow the procedures for system state update in Algorithm 1. The only exception is the computation of cliques in the growing graphs G^t.

For the ABM simulations, all node-maximal cliques K in each G^t must be computed because the clique is the CK-inducing substructure in a graph (Chwe 1999). Computing all node-maximal cliques in a graph is an NP-hard problem (Lawler et al. 1980), so we devised an approximate approach described in (Korkmaz et al. 2014) that is a *pre-processing step* and not part of simulations per se (the pre-computed cliques are inputs to simulations, enabling reuse of them). In essence, the approach computes cliques in G and then grows these cliques in size with the additional edges that are introduced at each successive t in G^t. Cliques are then compared because a new clique may contain as proper subgraphs other cliques, and in these cases, we keep only the largest clique. This process enables us to reuse the computed cliques across many simulations. We compute cliques K in graphs G^1 through G^5 and we will demonstrate in the results below that this is an ample number of graphs (this is a finding of our work). From each clique, we extract the CK sets. Computing these cliques is computationally intensive: computing cliques in G^1 through G^5 takes more than 100 hours of compute time on dedicated Intel Sandybridge nodes (that are used for high-performance computing) for many graphs with roughly 10,000 to 30,000 nodes.

5   Social Networks and Network Structure Results

To study the dynamics of the Chwe model on social networks, we adopt the approach taken by other researchers, e.g., (Tong et al. 2012; Zhang et al. 2016). Specifically, studies are carried out on a group of networks with wide ranges in structural property values. We use five different networks presented in Table 1. The first four networks are real (i.e., mined) web-based networks: FB is a Facebook user network (Viswanath et al. 2009), P2PG is a peer-to-peer network, Enron is an Enron email network, and Wiki is a Wikipedia network of voting for administrators (Leskovec and Krevl 2014). The SF1 is a stylized scale free (SF) network, that was generated by a standard preferential attachment method (Barabasi and Albert 1999) for comparison; they are widely found in practice. These networks, combined, cover very wide ranges in values of network measures, such as number of vertices (6×), number of edges (4×), average degree (4×), maximum degree (13×), and average clustering coefficient (80×), and hence represent a broad sampling of web-based mined network features. The purpose of our experiments is to demonstrate the behavior of the CK model on a set of networks whose structural properties vary over a wide range. Since the CK model does not use any information regarding the semantics of the underlying network, the behavior of the CK model cannot be directly related to the phenomenon that resulted in the formation of the network. Selected properties of the networks are provided in Figure 4 that are particularly relevant to CK model dynamics (of Section 6). Both demonstrate rapid increases in node degrees.

Table 1: Characteristics of web-based social networks. If there are multiple connected components in a graph, we use only the giant component. Here, n and m are numbers of vertices and edges, respectively, in the giant component; d_{ave} and d_{max} are average and maximum degrees; c_{ave} is average clustering coefficient; and Δ is graph diameter.

Network	Type/Domain	n	m	d_{ave}	d_{max}	c_{ave}	Δ
FB	Facebook	43,953	182,384	8.30	223	0.115	18
Enron	Email	33,696	180,811	10.7	1,383	0.509	17
P2PG	Peer Comms.	10,876	39,940	7.34	103	0.00622	10
Wiki	Online Voting	7,115	100762	28.3	1065	0.141	7
SF1	Stylized	4,956	45,031	18.2	270	0.0780	8

6   ABM Simulations and Results

6.1 Parameter Specifications for Simulations

We conduct discrete time agent-based simulations based on the Chwe CK model described in Section 3. A *simulation* consists of a set of runs, where a *run* covers the spread of contagion from an initial configuration with all nodes in state 0 at time $t = 0$, to a specified maximum time t_{max}. Values for all input parameters
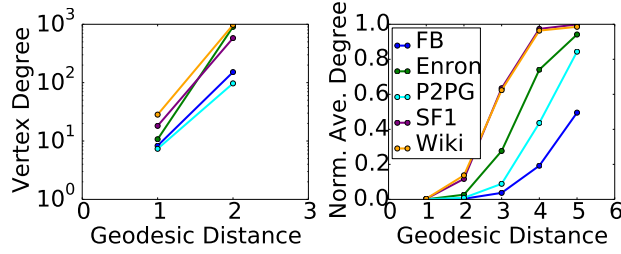
Figure 4: Network structures for the five graphs: (left) average vertex degree for geodesic distances 1 and 2 (i.e., for G^1 and G^2), and (right) normalized average degree d_{ave}/n as edges are formed between vertices that originally are separated by the specified geodesic distance. These data result from the densification mechanism of the Chwe model.

are provided in Table 2 and are associated with each simulation. Differences among runs is due to the stochasticity in models. We provide the average results from 50 runs in Section 6.2.

Table 2: Summary of simulation parameters.

Parameter	Description
Networks.	Presented in Table 1.
Node Thresholds θ_i.	Uniform values in $\{d_{ave}, d_{max}\}$ for a simulation for all nodes i. "Uniform" means all nodes in a network have the same value.
Participation Probabilities $p_{p,i}$.	Uniform value of 0.05, 0.1, 0.2, 0.4, or 1.0 assigned to all nodes in a simulation.
Time that graph grows t_{GG}.	Values of 1, 2, 3, 4, and 5.
Seed Vertices.	No specified seed vertices; all vertices initially in state 0. CK initiates contagion.
Simulation Duration t_{max}.	30, 90, 300 time steps.

6.2 Simulation Results

Simulations focus on the contagion spread size (as a fraction of $n = |V|$) as a function of time. We investigate the effects of the time through which the graph grows (t_{GG}), participation probability, and thresholds.

Contagion dynamics as a function of network growth. Figure 5 provides time histories of cumulative fraction of agents in state 1 for the FB network. In all simulations, all nodes have $\theta = 9 \approx d_{ave}$. Different colors correspond to different t_{GG} values. The rate at which agents (nodes) change from state 0 to 1 slows considerably as p_p decreases from 0.4 to 0.1.
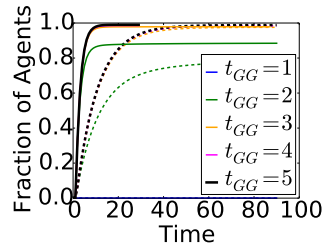

Figure 5: Chwe CK model results. Cumulative fraction of agents in state 1 as a function of time and growing graph mechanism (t_{GG}) in the FB network and with all nodes having threshold $\theta = 9 \approx d_{ave}$. Different color curves correspond to $t_{GG} \in \{1, 2, 3, 4, 5\}$. The dashed and solid curves correspond to $p_p = 0.1$ and $p_p = 0.4$.

Focusing on the solid curves in Figure 5 for $p_p = 0.4$, we observe that the results for $t_{GG} = 5$, 4, and 3 are virtually identical. All agents are in state 1 by $t = 10$. The green curve for $t_{GG} = 2$ indicates that about 85% of nodes reach state 1, and do so by $t = 10$. For $t_{GG} = 1$, there is almost zero contagion initiation and spreading; the curve lays near the x-axis. Thus, the dynamics change significantly in growing the graph from G^1 to G^2 to G^3 (i.e., in using $t_{GG} = 1$, 2, and 3), but not thereafter.

Hence, the GG mechanism for $t_{GG} = 3$ is sufficient to capture the contagion dynamics in the FB network. This is somewhat surprising since the diameter Δ of FB is $\Delta = 18$ which is much larger than 5, and the average clustering coefficient of 0.115 is not large. This result is useful because as part of the dynamics, all cliques in a graph must be identified, and this is an NP-hard problem (Lawler et al. 1980); it is widely believed that there is no efficient algorithm for any NP-hard problem.

Different networks and thresholds. We now turn to the behavior of all graphs in this study to determine whether the previous results hold for all of the networks, and for different thresholds. Figure 6 contains two plots of cumulative *final* fraction of affected agents as a function of geodesic distance for forming edges between pairs of vertices, for each of the five networks. The x-axis label is synonymous with t_{GG}. Hence, the final cumulative fraction of agents in state 1, for an x-axis value of ℓ, are generated on graphs through G^ℓ and correspond to $t_{GG} = \ell$. For all data, $p_p = 1$. In Figure 6(a), $\theta = d_{ave}$ for each node. Increasing the geodesic distance by 1 (meaning forming edges with friends-of-friends)—so that G^1 is used at $t = 1$ and G^2 is used for $t \geq 2$—increases the sizes of cliques to such a large extent that between 70% (for P2PG) and 100% (for SF1) of nodes change to state 1. Note that across all networks, the results for $t_{GG} = 3$ are approaching those for $t_{GG} = 5$ due the smaller thresholds and G^3.

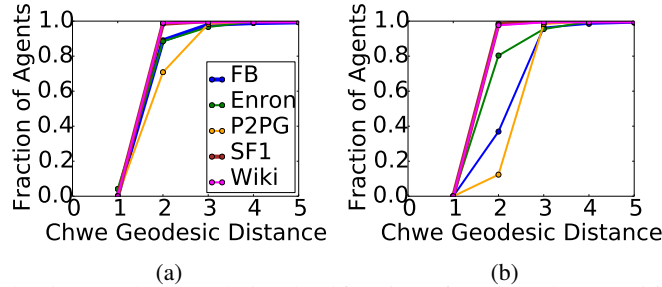(a)   (b)

Figure 6: Effect of GG mechanism on the cumulative *final* fraction of vertices that transition to state 1 in five networks (legend same in both plots). Simulations were run for 30 time steps. The x-axis corresponds to the value of t_{GG}. These are deterministic computations where $p_p = 1$. (a) results for $\theta = d_{ave}$ and (b) for $\theta = d_{max}$. The amount of contagion initiation at $t = 1$ on the original graph is zero for P2P, Enron, and Wiki in (a), and for all graphs in (b). Hence, the GG mechanism is critical to initiate contagion. Widespread contagion diffusion occurs at t_{GG}/Δ of roughly 1/6 to 1/2.

Another important factor is that contagion can initiate with CK models at multiple locations simultaneously: wherever an M^{clique} of nodes exists that satisfies Equation (1). Based on Figure 4, many nodes in FB have degree $d > \theta = 9$ in G^3, so contagion initiates at many locations in G.

Figure 6(b) displays similar data for all networks, except that now the thresholds are large: $\theta = d_{max}$ rather than d_{ave}. There are differences from the previous results for $t_{GG} = 1$ and 2; however, for $t_{GG} = 3$, again almost all nodes have transitioned to state 1. Note that these results, generated for $p_p = 1$, also hold for lesser p_p, as long as contagion will initiate. That is, the times to achieve the cumulative final fractions of nodes in state 1 will be greater for lesser p_p, but the trends with t_{GG} are still valid.

The GG mechanism is not only powerful, but operates over very small geodesic distances compared to the graph diameters Δ when p_p is large. (Δ is an upper bound on the time to spread a simple contagion (Centola and Macy 2007) through a network, where $\theta = 1$ for all nodes in a deterministic classic threshold model, and is achieved when one of the end nodes on the longest shortest-path is seeded.) These data and Table 1 show that full cascades occur for t_{GG}/Δ values that range from (1/6) to (1/2) when $p_p = 1$. These results also justify the use of graphs G^1 through G^3 in computing cliques (the substructure for CK in the Chwe model), rather than going to G^Δ, at least for large p_p, for the wide range in thresholds examined here, which is a significant computational savings in obviating the need to find cliques in G^t for $t > 3$.

Participation probability and network structure. We reduce $p_{p,i}$ by 10×, from 1.0 to 0.1, to investigate the effect of participation probability. Figure 7 contains data analogous to those in Figure 6, except that now $p_{p,i} = 0.1$ for all $i \in V$ in all analyses. The y-axis values are the cumulative fractions of agents in state 1 at $t = 30$; results are becoming asymptotic at this time (e.g., see Figure 5). Interestingly,

the GG mechanism overcomes the lower participation rate, so that the results for lesser thresholds in Figures 6(a) and 7(a) are quite similar.

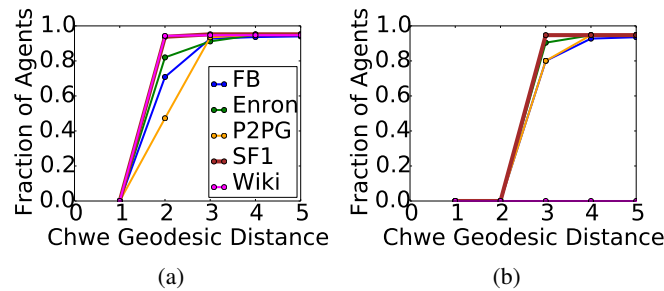(a)   (b)

Figure 7: Effect of participation probability in the Chwe CK model by comparison with the data of Figure 6. The y-axis is the cumulative fraction of agents in state 1 at $t = 30$. Here, $p_p = 0.1$, one-tenth the value in the previous figure, but the results do not change much in the final fraction of agents activating. This is the power of the [GG] mechanism of the Chwe CK model. Thresholds are (a) d_{ave} and (b) d_{max}. However, there are appreciable changes at time $t = 2$ in both plots (i.e., at intermediate times), compared to the results at this time in Figure 6. In (b), there is no contagion initiation for Wiki through $t = 5$. Legend is the same for both plots.

At greater thresholds $\theta = d_{max}$, the results in Figures 6(b) and 7(b) differ the most for $t_{GG} = 2$. There is, however, one exception: the Wiki network in Figure 7(b), for which no contagion initiates, even for $t_{GG} = 5$. All other networks have a ratio of $n/d_{max} \approx 20$ to 200, but for Wiki, the ratio is only about 7. Hence, even when the Wiki network evolves to K_{7115}, with $p_{p,i} = 0.1$, this participation probability is too low to generate enough participating nodes at any one time to initiate contagion. Initiation requires 1066 ($= d_{max} + 1$) nodes, which is $1066/7115 = 0.150 > p_{p,i} = 0.1$. We see the result that although greater d_{max} means that contagion can initiate at greater thresholds for greater $p_{p,i}$, if d_{max} is a large fraction of n, lesser $p_{p,i}$ will stymie contagion when $\theta = d_{max}$. This is an interesting example of the interplay between graph structure and contagion dynamics parameters. We note that if $p_{p,i}$ increases to 0.4, then contagion spreads rapidly (plots not shown) in Wiki and the other networks.

7   Conclusion and Future Work

In this paper, we study a game-theoretic model of collective action that is achieved through CK. This model has a unique feature that graph density increases as a function of time. The node sets of cliques are CK groups and coordination and cooperation can take place within these groups. Results are summarized in Section 1.3. Future work includes studying and comparing the mechanisms of other CK models.

REFERENCES

Aumann, R. J. 1976. "Agreeing to Disagree". *The Annals of Statistics* 4(6):1236–1239.

Barabasi, A., and R. Albert. 1999. "Emergence of Scaling in Random Networks". *Nature* 286:509–512.

Binmore, K., and L. Samuelson. 2001. "Coordinated Action in Electronic Mail Game". *Games and Economic Behavior* 35(1):6–30.

Bond, M. A., and C. B. Keys. 1993. "Empowerment, Diversity, and Collaboration: Promoting Synergy on Community Boards". *American Journal of Community Psychology* 21(1):37–57.

Centola, D., and M. Macy. 2007. "Complex Contagions and the Weakness of Long Ties". *American Journal of Sociology* 113(3):702–734.

Chwe, M. S.-Y. 1998. "Culture, Circles, and Commercials Publicity, Common Knowledge, and Social Coordination". *Rationality and Society* 10(1):47–75.

Chwe, M. S.-Y. 1999. "Structure and Strategy in Collective Action". *American Journal of Sociology* 105(1):128–156.

Chwe, M. S.-Y. 2000. "Communication and Coordination in Social Networks". *The Review of Economic Studies* 67(1):1–16.

Chwe, M. S.-Y. 2013. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton, New Jersey: Princeton University Press.

Clark, H. H. 1992. *Arenas of Language Use*. Chicago: University of Chicago Press.

Clark, H. H. 1996. *Using Language*. Cambridge, MA: Cambridge University Press.

Dalkiran, N. A., M. Hoffman, R. Paturi, D. Ricketts, and A. Vattani. 2012. "Common Knowledge and State-dependent Equilibria". In *Algorithmic Game Theory*, 84–95. Berlin, Heidelberg: Springer.

Geanakoplos, J. 1994. "Common knowledge". *Handbook of Game Theory with Economic Applications* 2:1437–1496.

Gladstein, D. L. 1984. "Groups in Context: A Model of Task Group Effectiveness". *Administrative Science Quarterly* 29(4):499–517.

Granovetter, M. 1978. "Threshold Models of Collective Behavior". *American Journal of Sociology* 83(6):1420–1443.

Hume, D. 2003. *A Treatise of Human Nature*. 2nd ed. New York: Oxford University Press.

Kempe, D., J. Kleinberg, and É. Tardos. 2003. "Maximizing the Spread of Influence Through a Social Network". In *Proceedings of the KDD Conference*. August 24-27, Washington, DC, 137–146.

Korkmaz, G., M. Capra, A. Kraig, K. Lakkaraju, C. J. Kuhlman, and F. Vega-Redondo. 2018. "Coordination and Common Knowledge on Communication Networks". In *Proceedings of the AAMAS Conference*. July 10-15, Stockholm, Sweden, 1062–1070.

Korkmaz, G., C. J. Kuhlman, A. Marathe, M. Marathe, and F. Vega-Redondo. 2014. "Collective Action Through Common Knowledge Using a Facebook Model". In *Proceedings of the AAMAS Conference*. May 5-9, Paris, France, 253–260.

Lawler, E. L., J. K. Lenstra, and A. H. G. R. Kan. 1980. "Generating All Maximal Independent Sets: NP-hardness and Polynomial-time Algorithms". *SIAM Journal on Computing* 9:558–565.

Leskovec, J., and A. Krevl. 2014. "SNAP Datasets: Stanford Large Network Dataset Collection".

Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Monderer, D., and D. Samet. 1989. "Approximating Common Knowledge with Common Beliefs". *Games and Economic Behavior* 1(2):170–190.

Myers, S., and J. Leskovec. 2012. "Clash of the Contagions: Cooperation and Competition in Information Diffusion". In *Proceedings of the ICDM Conference*. December 10-13, Brussels, Belgium, 539–548.

Nguyen, C., K. J. Schlesinger, and J. M. Carlson. 2017. "Collective Decision Dynamics in Group Evacuation". In *Proceedings of the Summer Simulation Multi-Conference*. July 9-12, Bellevue, Washington, 1–12.

Ostrom, E. 1990. *Governing the Commons*. Cambridge, MA: Cambridge University Press.

Rubinstein, A. 1989. "The Electronic Mail Game: Strategic Behavior Under "Almost Common Knowledge"". *The American Economic Review* 79(3):385–391.

Schelling, T. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Skyrms, B. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge, MA: Cambridge University Press.

Thomas, K. A., P. DeScioli, O. S. Haque, and S. Pinker. 2014. "The Psychology of Coordination and Common Knowledge". *Journal of Personality and Social Psychology* 107(4):657–676.

Tong, H., B. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos. 2012. "Gelling, and Melting, Large Graphs by Edge Manipulation". In *Proceedings of the CIKM Conference*. October 29 - November 2, Maui, Hawaii, 245–254.

Viswanath, B., A. Mislove, M. Cha, and K. P. Gummadi. 2009. "On the Evolution of User Interaction in Facebook". In *Proceedings of the WOSN Conference*. August 17, Barcelona, Spain, 37–42.

Willer, R., K. Kuwabara, and M. W. Macy. 2009. "The False Enforcement of Unpopular Norms". *American Journal of Sociology* 115(2):451–490.

Wittenbaum, G. M., S. I. Vaughan, and G. Strasser. 1998. "Coordination in Task-Performing Groups". *Theory and Research on Small Groups* 4:177–204.

Zhang, Y., A. Adiga, S. Saha, A. Vullikanti, and B. A. Prakash. 2016. "Near-Optimal Algorithms for Controlling Propagation at Group Scale on Networks". *IEEE Transactions on Knowledge and Data Engineering* 28(12):3339–3352.

AUTHOR BIOGRAPHIES

CHRIS J. KUHLMAN is faculty in the Biocomplexity Institute of the University of Virginia. His email address is cjk8gx@virginia.edu.

GIZEM KORKMAZ is faculty in the Biocomplexity Institute of the University of Virginia. Her email address is gkorkmaz@virginia.edu.

S. S. RAVI is faculty in the Biocomplexity Institute of the University of Virginia. His email address is ssr6nh@virginia.edu.

FERNANDO VEGA-REDONDO is a professor in the Department of Decision Sciences at Bocconi University. His email address is fernando.vega@bocconi.it.