Cascades and Myopic Routing in Nonhomogeneous Kleinberg's Small World Model

Jie Gao[1], Grant Schoenebeck[2], and Fang-Yi Yu[2](✉)

[1] Department of Computer Science, Stony Brook University,
Stony Brook, NY 11794, USA
`jgao@cs.stonybrook.edu`
[2] Department of EECS, University of Michigan, Ann Arbor, MI 48109, USA
{`schoeneb,fayu`}`@umich.edu`

Abstract. Kleinberg's small world model [20] simulates social networks with both strong and weak ties. In his original paper, Kleinberg showed how the distribution of weak-ties, parameterized by γ, influences the efficacy of myopic routing on the network. Recent work on social influence by k-complex contagion models discovered that the distribution of weak-ties also impacts the spreading rate in a crucial manner on Kleinberg's small world model [15]. In both cases the parameter of $\gamma = 2$ proves special: when γ is anything but 2 the properties no longer hold.

In this work, we propose a natural generalization of Kleinberg's small world model to allow node heterogeneity: instead of a single global parameter γ, each node has a personalized parameter γ chosen independently from a distribution $\mathcal{D}$. In contrast to the original model, we show that this model enables myopic routing and k-complex contagions on a large range of the parameter space, improving the robustness of the model. Moreover, we show that our generalization is supported by real-world data. Analysis of four different social networks shows that the nodes do not show homogeneity in terms of the variance of the lengths of edges incident to the same node.

1 Introduction

In Milgram's "Small World" experiments [23,26], he gave envelops to random residents of Wichita, Kansas and Omaha, Nebraska, and asked them to forward the envelopes to a personal contact so that they might eventually reach a specific banker in Massachusetts. The success of this experiment (which has since been observed in numerous other contexts – see related work) motivated Kleinberg's small work model which studies why such local decisions work [20]. This ingenious model shows not only that short paths between arbitrary nodes exist (this

---

J. Gao would like to acknowledge support through NSF DMS-1418255, CCF-1535900, CNS-1618391, DMS-1737812 and AFOSR FA9550-14-1-0193. G. Schoenebeck and F. Yu gratefully acknowledge the support of the National Science Foundation under Career Award 1452915 and AitF Award 1535912.

N. R. Devanur and P. Lu (Eds.): WINE 2017, LNCS 10674, pp. 383–394, 2017.
https://doi.org/10.1007/978-3-319-71924-5_27

so-called "small world" phenomena was already embedded into several fundamental models [6,24,27]), but also that these short paths can be easily discovered by myopic routing (i.e., using purely local knowledge).

Kleinberg's small world model considers an underlying metric space capturing the diversity of the population in various social attributes. Social ties are classified into two categories: strong ties that connect an individual to those similar in the social attribute space, and weak ties that may connect individuals far away. Kleinberg's model considers one parameter γ in determining how the weak ties are placed. Each node p takes a weak tie edge to a node q with probability proportional to $1/|pq|^\gamma$ where $|pq|$ denotes the distance between p and q in the social space. Thus at $\gamma = 0$ the weak ties are uniformly randomly distributed, and as γ increases shorter connections are increasingly favored.

However, in this model when the nodes are placed in a 2-dimensional grid the navigability only holds for a particular parameter choice: $\gamma = 2$. At this "sweetspot," a message can be delivered to the destination in $O(\log^2 n)$ hops, by hopping to the neighbor closest to the destination in the Euclidean metric. For *any* constant $\gamma \neq 2$, myopic routing, or, in general, any deterministic routing algorithm using only local information, provably fails to quickly deliver the message. Intuitively why $\gamma = 2$ is crucial, because at this sweetspot each weak tie edge uniformly at random lands in one of the annuli around the destination with inner radius 2^i and outer radius 2^{i+1}, for all i. Therefore, no matter where the destination is, with probability roughly $1/\log n$ there is a neighbor such that taking this neighbor reduces the Euclidean distance to the destination by half. If $\gamma < 2$, it turns out that the weak tie edges are too random and myopic routing loses its sense of direction. If $\gamma > 2$, the weak ties are simply too short and any path to the destination discoverable from local information necessarily takes many hops.

Other good properties also hold at special ranges of the parameter γ. In recent work on understanding complex social influence, it was shown how the distribution of weak-ties impacts the spreading behavior of k-complex contagions, in which a node becomes infected if at least k neighbors are infected [13,15]. Again it was shown that when $\gamma = 2$, for any constant k, the k-complex contagion spreads in a polylogarithmic number of rounds to the entire network while when $\gamma \neq 2$ complex contagions necessarily require a polynomial number of rounds. The analysis here connects to the intuition presented earlier for myopic routing. The sweetspot $\gamma = 2$ substantially speeds up the spreading of the contagions.

While the existence of the sweetspot is both insightful and elegant, it has raised new questions for modeling practical networks. The model feels fragile if the good properties only hold at a single parameter value and stop holding even with slight deviation. As put by Jackson [17]: "It is unlikely that societies just happen to hit the right balance. More likely there is something missing from the models, and it is clear the network-formation process underlying many social networks is much more complex than in these models." If Jackson is correct, then a theoretical model that more robustly justifies the empirical observations of Milgram and those who followed is needed.

Our Results. In this work, we generalize Kleinberg's small world model by considering a personalized, possibly heterogeneous γ_u for each node u in the network. In particular, each node u chooses its parameter $\gamma_u \in [0, \infty)$ i.i.d from a distribution $\mathcal{D}$. The weak tie edges issued by u will be placed on node v with probability proportional to $1/|uv|^{\gamma_u}$, where $|uv|$ denotes the distance between u and v in some underlying metric.

This model is motivated by both intuition and observations in real world data sets. It is natural to believe that some people have weak ties that are more/less dispersed (geographically or otherwise) that others. We also provide empirical evidence for node heterogeneity using real world social network data. Given a network, we can embed it in Euclidean space using spectral methods and examine the length of the edges attached to each node. We find that the empirical variance of the lengths of edges incident on the same vertex is substantially less than when the edge lengths are randomly permuted—suggesting that lengths of edges incident on the same vertex are indeed more correlated.

In this paper the main technical results we report is that both myopic routing and k-complex contagions operate quickly in the new model as long as the distribution $\mathcal{D}$ for the personalized γ has *non-negligible mass* around 2. Thus our model provides a robust justification for the observed properties of both myopic routing and k-complex contagions. Moreover it does this by only slightly tweaking Kleinberg's original model.

In particular, we can show that even if there is just $\Omega(\epsilon^\alpha)$ mass in the interval $[2-\epsilon, 2+\epsilon]$ of the distribution $\mathcal{D}$, where $\alpha > 0$ is *any* constant, then myopic routing and k-complex contagions (for any k) still only take polylogarithmic time! For example, it is enough that $\mathcal{D}$ be uniform on the interval $[a, b]$ for any $0 \leq a \leq 2 \leq b$. Note that in such a case, no particular γ_u will be exactly 2 (with probability 1). However, it turns out that enough of the γ_u are close enough to 2, which still enables these social processes.

We also show lower bounds. For myopic routing we show that if for some ϵ, there is no mass in $[2-\epsilon, 2+\epsilon]$, then the typical myopic routing time is polynomial. This is not obvious, as there can be a distribution $\mathcal{D}$ that allows weak ties that are short — connecting nodes nearby, and weak tie that are long – connecting nodes far away. Recall that in the original Kleinberg proof it was shown that *short ties only*, or *long ties only*, are not enough to enable myopic routing but it did not exclude the possibility when *both long and short ties* exist simultaneously. We show that in fact the combination of these weak tie edges are still not enough for enabling efficient myopic routing. In particular, there is a range of distances when none of the two types of ties are helpful, which forces the greedy routing to take a long time.

For complex contagions, our first lower bound shows that if for some $\epsilon > 0$, there is no mass in $[2-\epsilon, 2+\epsilon]$, then there is some k such that k-complex contagions require a polynomial time to spread. Again we must show that the synergy between short and long weak ties cannot enable complex contagions to quickly spread.

The above results for complex contagion apply for *any* k. We also study what happens for a particular k. Here we show that for each k there is an interval $[2, \beta_k)$ where $\beta_k = \frac{2(k+1)}{k}$ such that when $\mathcal{D}$ has constant support on $[2, \beta_k)$, k-complex contagions spread in polynomial time, but when, for any $\epsilon > 0$, $\mathcal{D}$ has no support on $[2-\epsilon, \beta_k + \epsilon]$, then k-complex contagions requires polynomial time to spread with high probability.

2 Related Work

Small World Graphs. The small world property—that there exists short paths between two random members of a network–appears in many real world complex networks in vastly different contexts ranging from film collaboration networks and neural networks [11] to email networks [10], food webs [28] and protein interaction networks [19].

It has been discovered in a number of settings that random edges introduced to a graph can dramatically reduce the network diameter, creating a small world graph. This observation was made in the Watts-Strogatz model [27] (when edges are rewired to a random destination) as well as for regular random graphs [6] (a graph in which all nodes have the same constant degree and edges are uniformly randomly placed). Kleinberg's small world model can be considered as an extension to such models. In particular, the Newmann-Watts model [24] (a variant of the Watts-Strogatz model in which random edges are added in *addition* to existing edges) is a special case of Kleinberg's model for choosing $\gamma = 0$ — i.e., the weak ties are uniformly randomly added.

Navigability. Milgram's "Small World" experiments [23,26] illustrated not only the small world property—that short paths exist—but, in fact, showed a stronger property—that such paths can be efficiently found using only local information—called navigability. A short path was discovered through a *local* algorithm with the participants forwarding to a friend who *they believed* to be more likely to know the target. Although forwarding decision-making was not systematically recorded, geographical proximity was found to be an important forwarding criterion in some cases. Other criteria such as profession and popularity may have been used as well. A later study using email-chains [10] confirms this as well, finding that at least half of the choices were due to either geographical proximity of the acquaintance to the target or occupational similarity.

Besides the Kleinberg's small world model, several other models also considered using metric distances in modeling social ties. For example, Kumar *et al.* [22] extended the Kleinberg's model to include the underlying metrics with low-doubling dimension. This model also requires a specific distribution of the weak ties.

Another line of work diverges from distance function defined over some low-dimensional space, but instead defines a distance function based on some hierarchical structure. For example, Watts *et al.* considered a hierarchical professional organization of individuals and a homophilous network with ties added between two nodes closer in the hierarchy with a higher probability. If each node has a fixed probability of dropping the message, they show a greedy routing algorithm sending packages to the neighbor most similar to the target (called homophily-based routing) successfully delivers a fraction of the messages before they are dropped. Kleinberg also confirmed similar results on a hierarchical network, in which the nodes are represented as leaf nodes of a hierarchical organization structure and random edges are added to the leaves with probability dependent on their tree distance. When each node has polylogarithmic out-degree, greedy routing based on the tree distance arrives at the destination in $O(\log n)$ hops. While the aforementioned models also successfully create a more robust network model for myopic routing, in doing so they abandoned the spatial structure of Kleinberg's small world model. While certain structures can be modeled well as a hierarchy, others are much more natural as a continuum, as in Kleinberg's model—e.g. distances, wealth, political ideology, and education.

Boguñá et al. [5] proposed a model that assumes a social metric space and the power law degree distribution. They considered nodes on a ring and assigned target degrees from a power law distribution. An edge is then placed between two nodes with a probability positively dependent on their distance on the ring and negatively dependent on their degrees. They investigated greedy routing with the distances on the ring as a means of navigating in the network. Krioukov et al. [21] considered using a hyperbolic plane as the hidden social space. Nodes are uniformly distributed in a radius R disk in a hyperbolic plane with edges placed in pairs with distance smaller than r. They show that such a graph has power law degree distribution and that greedy routing with hyperbolic distance has a high success rate.

Complex Contagions. The model of k-complex contagions belongs to the general family of *threshold models*, in which each node has a threshold on the number of infected edges/neighbors needed to become infected [16]. The threshold model is motivated by certain coordination games studied in the economics literature in which a user maximizes its payoff when adopting the behavior as the majority of its neighbors.

k-complex contagions have been previously studied in the Kleinberg small world model [15] and their spreading behaviour was almost completely classified [13]. Ghasemiesfeh *et al.* [15] showed that for any k, if $\gamma = 2$ then complex contagions spread quickly, in a polylogarithmic number of rounds. Further, Ebrahimi *et al.* [13] showed that for each $k \geq 2$, there exists an interval of values, $[2, \alpha_k]$, such that when $\gamma \in (2, \alpha_k)$, a k-complex contagion spreads quickly on the corresponding graph, in a polylogarithmic number of rounds. However, if γ is outside this range, then a k-complex contagion requires a polynomial number of rounds to spread to the entire network. They also showed similar results for a variant of the Kleinberg model where edges are added without replacement (thus multi-edges are allowed).

k-complex contagions have also been studied in other social network models, for examples, networks that have a time-evolving nature (e.g. the Preferential

Attachment model) [12,14], and configuration model networks with power-law degree distribution [25].

k-complex contagions are referred to as *bootstrap percolation* [1,9] in the literature, especially when initial seeds are chosen randomly at random. Here, the focus is often to examine the threshold of the number of initial seeds with which the infection eventually 'percolates', i.e. diffuses to the entire network. Studies have been done on the random Erdős-Rényi graph [18], random regular graphs [4], and the configuration model [2], etc [3]. All of these results show that for a complex contagion to percolate, the number of initial seeds is a growing function of the network size and in many cases a constant fraction of the entire network. In contrast, we always start with a constant number of seeds and we would like to examine whether a fast spreading is possible.

3 Preliminaries

Recall that in the Kleinberg's small world model [20], nodes are defined on a $n \times n$ grid[1]. Each node u connects to nodes within grid Manhattan distance $\lceil q \rceil$, where q is a constant. These edges are referred to as ***strong ties***. In addition, each node generates p random outgoing edges (without replacement), termed ***weak ties***. The probability that node u connects to node v via a random edge is $1/\lambda_\gamma d(u,v)^\gamma$, in which $d(u,v)$ is the Manhattan distance of u, v and $\lambda_\gamma = \sum_v d(u,v)^{-\gamma}$ is a normalization factor. Further, we remark that the graph is directed — the weak ties issued by a node u have u as the tail and the strong ties are bidirectional.

For ***Heterogeneous Kleinberg's small world*** $HetK_{p,q,\mathcal{D}}(n)$, we define p, q, n as in the original model, but, instead of one global γ, each node u independently chooses its personalized parameter γ_u from the distribution $\mathcal{D}$ on $[0, \infty)$ with probability density function[2] $f_\mathcal{D}$ and cumulative distribution function $F_\mathcal{D}$. Let $M_\mathcal{D}(\epsilon) = F_\mathcal{D}(2+\epsilon) - F_\mathcal{D}(2-\epsilon)$ measure the "mass" of $\mathcal{D}$ around 2.

We study two dynamics on this heterogeneous Kleinberg's small world model: *decentralized routing*, and *k-complex contagion.*

In the ***decentralized routing algorithm***, a message is passed to one of its (local or long-range) contacts using only local information. Given the source s and destination t in the graph, we denote the routing process/algorithm $\mathcal{A}$: a sequence of nodes on the graph $(x_i)_{i \geq 0}$ where $x_0 = s$. The *delivery time* from s to t of algorithm $\mathcal{A}$ is defined as $\min\{i \geq 0 : x_i = t\}$ which is a random variable with σ-space generated by $HetK_{p,q,\mathcal{D}}(n)$ and the myopic routing algorithm. The *expected delivery time* of a decentralized algorithm $\mathcal{A}$ is the expected delivery time for uniformly chosen sources s and destinations t. The *myopic greedy algorithm* routes the message from the current location to be as close as possible to

---

[1] In order to eliminate the boundary effect, we wrap up the grid into a torus – i.e., the top boundary is identified with the bottom boundary and the left boundary is identified with the right boundary.

[2] For discrete distribution, the probability density function exists if we allow using Dirac delta function.

the destination vertex (according to the grid distance) using only one hop from the current node.

We define a k-complex contagion process in a directed graph following the definition in Ghasemiesfeh et al. [15]. We assume k is a small constant. A ***k-complex contagion*** $\mathrm{CC}(G, k, \mathcal{I})$ is a contagion that initially infects vertices of $\mathcal{I}$ and spreads over graph G. The contagion proceeds in rounds. At each round, each vertex with at least k infected neighbors becomes infected. The vertices of $\mathcal{I}$ are called the initial seeds. We say that k nodes $(u_1, \cdots, u_k)$ are a ***k-seed cluster*** if they form a connected subgraph via *only* the grid structure. A k-complex contagion spreads in the inverse direction of an edge: a node becomes infected if it follows at least k infected neighbors. In this work, we define the ***speed of a k-complex contagion*** as the number of rounds it takes to infect the whole graph which is always finite if we take $q \geq k$ and $\mathcal{I}$ is a k-seed cluster.

4 Myopic Routing Upper Bounds

In this section, we prove the following theorem:

Theorem 1 (Myopic Routing Upper Bounds). *Given a* $HetK_{p,q,\mathcal{D}}(n)$ *with constant* $p, q \geq 1$ *and distribution* $\mathcal{D}$*. If there exists some constants* $\epsilon_0 > 0$*,* $\alpha \geq 1$ *and* $K > 0$ *such that* $\forall \epsilon < \epsilon_0$*,* $M_{\mathcal{D}}(\epsilon) \geq K\epsilon^{\alpha}$*, the expected delivery time of the myopic greedy algorithm is at most* $O(\log^{2+\alpha} n)$*.*

The above theorem proves fast myopic routing over a large class of Heterogeneous Kleinberg's Small world models. The only distributions that this theorem fails to apply to are distributions with negligible mass near 2. In particular, if $\mathcal{D}$ is uniform over *any* finite interval containing 2, then myopic routing will take time at most $O(\log^3 n)$, and as long as the mass near 2 is non-trivial (i.e., lower bounded by the inverse of some fixed polynomial), then delivery only takes poly-log time.

Remark 1. Note that if the random variable associated with $\mathcal{D}$ is a constant random variable that takes a constant value 2, the $HetK_{p,q,\mathcal{D}}(n)$ degenerates to the original Kleinberg's model with $\gamma = 2$, and the Theorem 1 is tight which yields the same $O(\log^2 n)$ upper bound on delivery time on myopic greedy routing algorithm.

The proof of Theorem 1 follows the general outline of the proof in Kleinberg's original paper: measure the progress of process $\mathcal{A} = (x_i)_{i \geq 0}$ in terms of phases which will be defined later and show the following: (1) monotone property of the process, (2) upper bound the total number of phase, (3) lower bound the probability of finishing each phase. The formal proof will be in the full version.

5 Myopic Routing Lower Bounds

In this section we prove a lower bound for any decentralized algorithms on the Heterogeneous Kleinberg Small World $HetK_{p,q,\mathcal{D}}(n)$ in the following theorem:

Theorem 2. *Given a Heterogeneous Kleinberg's Small World network* $HetK_{p,q,\mathcal{D}}(n)$ *with constant parameters* p, q *and probabilistic density function* $f_{\mathcal{D}}$ *for the distribution* $\mathcal{D}$ *on the personalized* γ_u *for each node* u*, if there exists a constant* $\epsilon_0 > 0$ *such that* $F(2+\epsilon_0) - F(2-\epsilon_0) = 0$*, where* F *is the cumulative density function of* $\mathcal{D}$*, then the expected routing time for all decentralized algorithms is* $\Omega(n^{\xi})$ *where* $\xi = \frac{\epsilon_0}{3(3+\epsilon_0)}$*.*

In the original Kleinberg's model [20], all nodes use the same γ parameter. When γ is greater than 2 the weak ties are too short in expectation such that it would need a polynomial number of hops to reach a far away destination. When γ is smaller than 2 the edges are too random to be useful for nearby destinations. But in a heterogeneous model, the nodes may have different γ values. The nodes with $\gamma_u > 2$ have *concentrated edges* while those with $\gamma_u < 2$ have *diffuse edges.* A network with only concentrated edges or only diffuse edges cannot support polylogarithmic myopic routing. But it is unclear whether the combination of them, as in the heterogeneous model, can lead to polylogarithmic delivery time. Theorem 2 states that this is not true. We show this by considering a scope where neither type of edges is helpful. The formal proof will be in the full version.

6 Complex Contagion Upper Bounds

The spreading of k-complex contagion on the original Kleinberg's model has been fully characterized in [13,15]. If a k-seed cluster is infected initially, the contagion spreads to the entire network in $O(\text{polylog}(n))$ rounds if $\gamma \in [2, \beta_k)$, where $\beta_k = \frac{2(k+1)}{k}$, and in $\Omega(\text{poly}(n))$ rounds otherwise.

6.1 Non-negligible Mass Near 2

In the heterogeneous Kleinberg model, we first show a result that is analogous to our results for myopic routing: as long as the distribution $\mathcal{D}$ for γ_u has a non-negligible amount of mass near 2, then for any k, k-complex contagions spread in polylog time—but the exponent of $\log n$ depends on k and $\mathcal{D}$.

Theorem 3. *Fix a distribution* $\mathcal{D}$*, an integer* $k > 0$ *and* $\eta > 0$*. If there exist constants* $\epsilon_0 > 0$ *and* $\alpha \geq 0$ *where* $M_{\mathcal{D}}(\epsilon) \geq K\epsilon^{\alpha}$ *for all* $\epsilon \leq \epsilon_0$*, and* $p, q \geq k$*, there exists* $\kappa = k\alpha + \frac{k(k+1)}{2}$*, such that a* k*-complex contagion* $\mathrm{CC}(HetK_{p,q,\mathcal{D}}(n), k, \mathcal{I})$ *starting from a* k*-seed cluster* $\mathcal{I}$ *takes at most* $O(\log^{(3+\kappa)/2} n)$ *rounds*[3] *to spread to the whole network with probability at least* $1 - n^{-\eta}$ *over the randomness of* $HetK_{p,q,\mathcal{D}}(n)$*.*

The theorem is based on the observation that the infected region doubles its size in a polylogarithmic number of steps. In this way the general proof framework is similar to that in [15], and the complete proof will be in the full version.

---

[3] The scalar depends on the constants k, η, α, K.

6.2 Fixed k

For a specific k, we can show that as long as the distribution $\mathcal{D}$ has constant mass in the interval $[2, \beta_k)$ (recall for the beginning of the section that $\beta_k = \frac{2(k+1)}{k}$), then the k-complex contagion will spread to the entire network in a polylogarithmic number of rounds. Recall that the results in Theorem 3 only require non-negligible mass near 2. Here we require constant mass, but the mass need not be asymptotically close to 2 as long as it is in the interval $(2, \beta_k)$.

Theorem 4. *Fix a distribution $\mathcal{D}$, an integer $k > 0$ and $\eta > 0$. If $\Pr_{\gamma\sim\mathcal{D}}[\gamma \in [2, \beta_k)] > 0$ where $\beta_k = \frac{2(k+1)}{k}$, and $p, q \geq k$. There exists $\xi > 0$ depending on $\mathcal{D}$ and k such that, the speed of a k-complex contagion $\mathrm{CC}(HetK_{p,q,\mathcal{D}}(n), k, \mathcal{I})$ starting from a k-seed cluster $\mathcal{I}$ is at most $O\left(\log^{\xi} n\right)$ with probability at least $1 - n^{-\eta}$.*

The proof of Theorem 4 uses the same divide and conquer strategy as in [13], and the proof will be in the full version.

7 Complex Contagion Lower Bounds

In this section, we describe a polynomial time lower bound for the spreading rate of k-complex contagion on the Heterogeneous Kleinberg Small World $HetK_{p,q,\mathcal{D}}(n)$, when the distribution $\mathcal{D}$ on the personal parameter γ_u has zero weight around two. Here we first state the theorem for a fixed k, and the result near two is a natural corollary.

Theorem 5 (Lower bound for fixed k). *Given distribution $\mathcal{D}$, constant integers $k, p, q > 0$, and $\epsilon_0 > 0$ such that $F_{\mathcal{D}}(\beta_k + \epsilon_0) - F_{\mathcal{D}}(2 - \epsilon_0) = 0$, then there exist constants $\xi, \eta > 0$ depending on $\mathcal{D}$ and k, such that the time it takes a k-contagion starting at seed-cluster $\mathcal{I}$, $\mathrm{CC}(HetK_{p,q,\mathcal{D}}(n), k, \mathcal{I})$, to infect all nodes is at least $\Omega(n^{\xi})$ with probability at least $1 - O(n^{-\eta})$ over the randomness of $HetK_{p,q,\mathcal{D}}(n)$.*

If $\mathcal{D}$ satisfies the condition in Theorem 5, we can partition the support into two disjoint sets $Supp\{\mathcal{D}\} = D_1 \cup D_2$ such that $\gamma_1 = 2 - \epsilon_1 = \sup\{\gamma \in D_1\} < 2 - \epsilon_0$, and $\gamma_2 = 2 + \epsilon_2 = \inf\{\gamma \in D_2\} > 2(1 + 1/k) + \epsilon_0$.

Ebrahimi et al. [13] proved for the original Kleinberg model if $\gamma > \frac{2(k+1)}{k}$ the weak ties will be too short to create remote k-seeds; on the other hand, if $\gamma < 2$ the weak ties will be too random to form k-seeds at all. Similar to proving the lower bound for myopic routing, the challenge in proving this theorem is the synergy between concentrated and diffuse edges which can possibly be exploited by k-complex contagions in the heterogeneous Kleinberg model. We resolve this by considering a scale where neither type of edges is helpful.

Before proving Theorem 5 we state a corollary concerning a lower bound when there is no mass around 2.

Corollary 1 (Lower bound for no mass around 2**).** *Given distribution* $\mathcal{D}$*, constant integers* $p, q > 0$*, and* $\epsilon_0 > 0$ *such that* $F_{\mathcal{D}}(2+\epsilon_0) - F_{\mathcal{D}}(2-\epsilon_0) = 0$*, there exist a constant integer* $k > 0$ *and* $\xi, \eta > 0$ *such that the time it takes a* k*-contagion starting at seed-cluster* $\mathcal{I}$*,* $\mathrm{CC}(HetK_{p,q,\mathcal{D}}(n), k, \mathcal{I})$*, to infect all nodes is at least* n^{ξ} *with probability at least* $1 - O(n^{-\eta})$ *over the randomness of* $HetK_{p,q,\mathcal{D}}(n)$*.*

The corollary follows directly from Theorem 5 by taking a sufficiently large k.

8 Empirical Results

See full version.

9 Conclusion

We introduced a generalization of the Kleinberg small world model where the parameter which determines how concentrated or diffuse long ties are can be different for each node, and showed empirical results which support our new model. We proved that this model overcomes a weakness of the original model, which is that the parameters needed to be tuned just right to facilitate fast myopic routing, which was the original motivation behind the model's development. For a wide array of parameters, our new model facilitates both fast myopic routing and the fast spread of complex contagions.

One future direction would be try to learn the heterogeneous distribution in real-world network data. Another future direction would be to connect this model to the "structural holes" theory [7,8] which posits that agents gain power by sitting along many shortest paths, by allowing agents to mediate the passing of information. That is, in the hierarchical small world model, which types of individuals are mostly likely to lie on shortest paths, or, in general, are more useful in myopic routing and complex contagions. A final future direction would be to study the Kleinberg small world model where nodes have a non-uniform (e.g. powerlaw) degree distribution of weak ties. This may provide an alternative way to generalize the Kleinberg small world model so that it supports myopic routing and complex contagions over a larger parameter range.

References

1. Adler, J.: Bootstrap percolation. Phys. A: Stat. Theor. Phys. **171**(3), 453–470 (1991)
2. Amini, H.: Bootstrap percolation and diffusion in random graphs with given vertex degrees. Electr. J. Comb. **17**(1), R25 (2010)
3. Amini, H., Fountoulakis, N.: What i tell you three times is true: bootstrap percolation in small worlds. In: Goldberg, P.W. (ed.) WINE 2012. LNCS, vol. 7695, pp. 462–474. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35311-6_34

4. Balogh, J., Pittel, B.: Bootstrap percolation on the random regular graph. Random Struct. Algorithms **30**, 257–286 (2007)
5. Boguna, M., Krioukov, D., Claffy, K.C.: Navigability of complex networks. Nat. Phys. **5**, 74–80 (2009)
6. Bollobás, B., Chung, F.R.K.: The diameter of a cycle plus a random matching. SIAM J. Discret. Math. **1**(3), 328–333 (1988)
7. Burt, R.S.: Structural Holes: The Social Structure of Competition. Cambridge University Press, Cambridge (1992)
8. Burt, R.S.: Structural Holes: The social structure of competition. Harvard University Press, Cambridge (1995)
9. Chalupa, J., Leath, P.L., Reich, G.R.: Bootstrap percolation on a Bethe lattice. J. Phys. C: Solid State Phys. **12**(1), L31 (1979)
10. Dodds, P.S., Muhamad, R., Watts, D.J.: An experimental study of search in global social networks. Science **301**, 827 (2003)
11. Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 409–410 (1998)
12. Ebrahimi, R., Gao, J., Ghasemiesfeh, G., Schoenebeck, G.: How complex contagions in preferential attachment models and other time-evolving networks. IEEE Trans. Netw. Sci. Eng. **PP**(99), 1 (2017). https://doi.org/10.1109/TNSE.2017.2718024. ISSN 2327–4697
13. Ebrahimi, R., Gao, J., Ghasemiesfeh, G., Schoenebeck, G.: Complex contagions in Kleinberg's small world model. In: Proceedings of the 6th Innovations in Theoretical Computer Science (ITCS 2015), pp. 63–72. January 2015
14. Gao, J., Ghasemiesfeh, G., Schoenebeck, G., Yu, F.-Y.: General threshold model for social cascades: analysis and simulations. In: Proceedings of the 2016 ACM Conference on Economics and Computation, pp. 617–634. ACM (2016)
15. Ghasemiesfeh, G., Ebrahimi, R., Gao, J.: Complex contagion and the weakness of long ties in social networks: revisited. In: Proceedings of the fourteenth ACM conference on Electronic Commerce, pp. 507–524. ACM (2013)
16. Granovetter, M.: Threshold models of collective behavior. Am. J. Sociol. **83**(6), 1420–1443 (1978)
17. Jackson, M.O.: Social and Economic Networks. Princeton University Press, Princeton (2008). ISBN 0691134405, 9780691134406
18. Janson, S., Luczak, T., Turova, T., Vallier, T.: Bootstrap percolation on the random graph $G_{n,p}$. Ann. Appl. Probab. **22**(5), 1989–2047 (2012)
19. Jeong, H., Mason, S.P., Barabasi, A.-L., Oltvai, Z.N.: Lethality and centrality in protein networks. Nature **411**, 41–42 (2001)
20. Kleinberg, J., The small-world phenomenon: an algorithm perspective. In: Proceedings of the 32-nd Annual ACM Symposium on Theory of Computing, pp. 163–170 (2000)
21. Krioukov, D., Papadopoulos, F., Boguna, M., Vahdat, A.: Greedy forwarding in scale-free networks embedded in hyperbolic metric spaces. In: ACM SIGMETRICS Workshop on Mathematical Performance Modeling and Analysis (MAMA) June 2009
22. Kumar, R., Liben-Nowell, D., Tomkins, A.: Navigating low-dimensional and hierarchical population networks. In: Azar, Y., Erlebach, T. (eds.) ESA 2006. LNCS, vol. 4168, pp. 480–491. Springer, Heidelberg (2006). https://doi.org/10.1007/11841036_44. ISBN 3-540-38875-3
23. Milgram, S.: The small world problem. Phychol. Today **1**, 61–67 (1967)
24. Newman, M.E.J., Moore, C., Watts, D.J.: Mean-field solution of the small-world network model. Phys. Rev. Lett. **84**, 3201–3204 (2000)

25. Schoenebeck, G., Yu, F.-Y.: Complex contagions on configuration model graphs with a power-law degree distribution. In: Cai, Y., Vetta, A. (eds.) WINE 2016. LNCS, vol. 10123, pp. 459–472. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-54110-4_32
26. Travers, J., Milgram, S.: An experimental study of the small world problem. Sociometry **32**, 425 (1969)
27. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)
28. Williams, R.J., Berlow, E.L., Dunne, J.A., Barabasi, A.L., Martinez, N.D.: Two degrees of separation in complex food webs. Proc. Nat. Acad. Sci. **99**(20), 12913–12916 (2002)