Bull. Aust. Math. Soc. **90** (2014), 363–375
doi:10.1017/S0004972714000604

ON DYNAMIC MONOPOLIES OF GRAPHS WITH PROBABILISTIC THRESHOLDS

HOSSEIN SOLTANI and MANOUCHEHR ZAKER✉

(Received 24 March 2014; accepted 25 June 2014; first published online 9 September 2014)

Abstract

Let G be a graph and τ be an assignment of nonnegative thresholds to the vertices of G. A subset of vertices, D, is an irreversible dynamic monopoly of (G, τ) if the vertices of G can be partitioned into subsets $D_0, D_1, \ldots, D_k$ such that $D_0 = D$ and, for all i with $0 \le i \le k-1$, each vertex v in D_{i+1} has at least $\tau(v)$ neighbours in the union of $D_0, D_1, \ldots, D_i$. Dynamic monopolies model the spread of influence or propagation of opinion in social networks, where the graph G represents the underlying network. The smallest cardinality of any dynamic monopoly of (G, τ) is denoted by $\text{dyn}_\tau(G)$. In this paper we assume that the threshold of each vertex v of the network is a random variable X_v such that $0 \le X_v \le \deg_G(v) + 1$. We obtain sharp bounds on the expectation and the concentration of $\text{dyn}_\tau(G)$ around its mean value. We also obtain some lower bounds for the size of dynamic monopolies in terms of the order of graph and expectation of the thresholds.

2010 *Mathematics subject classification*: primary 05C69; secondary 91D30.

Keywords and phrases: spread of influence in graphs, irreversible dynamic monopolies, probabilistic threshold assignment.

1. Introduction and motivation

The irreversible (progressive) spread of influence, such as diseases or opinions in a population, viruses in virtual networks, and innovation and viral marketing in social networks, has attracted considerable interest in recent years. These phenomena can be formulated and analysed as discrete dynamical processes using the concept of (progressive) dynamic monopolies. Assume that the (simple and undirected) graph G on the vertex set $V(G)$ and the edge set $E(G)$ represents the underlying network. Assign thresholds $\tau : V(G) \to \mathbb{N} \cup \{0\}$ to the vertices of G where the value $\tau(v)$ measures the level of susceptibility of the vertex v. The discrete time dynamic process corresponding to the threshold assignment τ is defined as follows.

The process starts with a subset D of vertices which consists of the vertices having the state + at time 0. We denote the set of vertices with state + at time i by M_i. So at the beginning, at time 0, we have $M_0 = D$. Then at each time $i + 1 \ge 1$, the state of each vertex v changes to the state + provided that at least $\tau(v)$ neighbours of v belong to M_i.

If the state of v is already + at time i (that is v is in M_i), then its state remains +. If at a certain time i of the process a vertex v has state + then v is said to be an active vertex at time i. Note that the process defined above is progressive or irreversible; that is, when the state of a vertex becomes + at some step of the process then its state remains unchanged until the end of the process. By a dynamic monopoly (dynamo) we mean any subset D of the vertices of G such that by starting from D, all the vertices of G reach state + at the end of the process. Throughout this paper, by (G, τ) we mean a graph G together with a threshold assignment τ to the vertices of G. By the size of a dynamic monopoly D we mean the cardinality of D. It is easy to see that a subset, D, of vertices in a graph (G, τ) is a dynamic monopoly if and only if there exists a partition of $V(G)$ into subsets $D_0, D_1, \ldots, D_k$ such that $D_0 = D$ and, for all $i = 1, \ldots, k - 1$, each vertex v in D_{i+1} has at least $\tau(v)$ neighbours in $D_0 \cup \cdots \cup D_i$. The smallest size of any dynamic monopoly of (G, τ) is denoted by $\text{dyn}_\tau(G)$. Throughout this paper we deal with undirected graphs without multiple edges or loops except in Section 3, where we generalise our concepts and results for directed graphs.

Irreversible dynamic monopolies have been widely studied in the literature in [5–7, 12, 14, 16, 19, 20], and also under the equivalent terms 'conversion sets' [2, 4, 11] and 'target set selection' [1, 8, 9, 18]. Dynamic monopolies have applications in viral marketing [10]. A concept similar to dynamic monopolies, the so-called bootstrap percolation, was widely studied in the area of percolation theory (see, for example, [3]). In such applications, the threshold assignment is constant for all vertices of the graph. Different threshold assignments such as constant assignments, simple majority assignment (where, for each vertex v, $\tau(v) = \lceil \deg(v)/2 \rceil$) and strict majority assignment (where, for each vertex v, $\tau(v) = \lceil (\deg(v) + 1)/2 \rceil$) have been studied. Dynamic monopolies in terms of the average threshold were studied in [14]. Dynamic monopolies of random graphs were studied in [5, 7, 16]. Dynamic monopolies with probabilistic thresholds was first considered by the second author in [19], where the concept of a homogeneous society was defined as follows. Assume that in a social network each person v chooses a threshold t with the probability $p_v(t)$. If for any fixed t, $p_v(t)$ is the same for all vertices v, then the network with these probabilistic thresholds is called homogeneous. In other words, if the random variable X_v denotes the threshold of each vertex v in a network G, then G is homogenous if $\{X_v\}_{v \in G}$ is an identically distributed set of random variables. The motivation to study probabilistic thresholds is that in practice it is difficult or impossible to have knowledge about all individual thresholds of the society. Information on the proportion of the vertices having a certain threshold is more accessible using statistical approaches. Indeed, assigning the threshold of an individual in a social network as a deterministic value seems unreasonably idealistic. These comments motivate the study of graphs with probabilistic thresholds. In the following probabilistic model we assume that it is probable that a vertex v has threshold at most $\deg(v) + 1$. This represents the fact that in practice a vertex may remain inactive even if its whole neighbourhood is active. As an example of such a phenomenon, consider the spread of yes/no votes in an election.

We now give some formal definitions. Let G be a graph, where the threshold of each vertex v is a random variable X_v such that $0 \leq X_v \leq \deg(v) + 1$, where $\deg(v)$ is the degree of v in G. By the interval $[0, \deg(v) + 1]$ we mean $\{0, \ldots, \deg(v) + 1\}$. The random variable X_v chooses each threshold $t \in [0, \deg(v) + 1]$ with a certain probability. Let $X = (X_v)_{v \in G}$. For any $x = (x_v)_{v \in G} \in \prod_{v \in G}[0, \deg(v) + 1]$, $\mathrm{dyn}_x(G)$ is defined as the smallest size of a dynamic monopoly of (G, x). In fact when the threshold assignment τ is a member, say x, of the probability space $\prod_{v \in G}[0, \deg(v) + 1]$, then instead of $\mathrm{dyn}_\tau(G)$ we may write $\mathrm{dyn}_x(G)$. Each element $x = (x_v)_v$ occurs with the probability $\prod_{v \in G} \Pr(X_v = x_v)$. We define $\mathrm{dyn}_X(G)$ as a random variable on the probability space $\prod_{v \in G}[0, \deg(v) + 1]$ whose value on each element $x = (x_v)_{v \in G}$ is $\mathrm{dyn}_x(G)$. For any $x = (x_v)_v$, let $M(x)$ be any fixed dynamic monopoly with the smallest cardinality in (G, x). A random version of $M(x)$ is defined as $M(X) : \prod_{v \in G}[0, \deg(v) + 1] \to 2^{V(G)}$ whose value on each $x = (x_v)_v$ is $M(x)$. Note that the two random variables $\mathrm{dyn}_X(G)$ and $|M(X)|$ are the same. Let $\mathcal{Q}$ be any property for random vectors X. As an example of $\mathcal{Q}$ consider the set of threshold assignments $X = (X_v)_{v \in G}$ for arbitrary graphs G such that for some constant λ and each vertex v, $X_v \le \lambda \mathbb{E}(X_v)$. Consider now the space $\mathcal{S}_n = \{(G, X) : |G| = n, X \in \mathcal{Q}\}$. Assume that to any $(G, X) \in \mathcal{S}_n$, there corresponds an interval $I(G, X) \subseteq \mathbb{R}$ such that its endpoints can be either open or closed and are determined in terms of the graph parameters of G (e.g. the order) and values of X (the endpoints can also be $\pm\infty$). We say that (G, X) has property $\mathcal{I}$ if $\mathrm{dyn}_X(G) \in I(G, X)$. We also say that 'with high probability (or simply whp) $\mathrm{dyn}_X(G) \in I(G, X)$' if

$$\lim_{n \to \infty} \Pr((G, X) \in \mathcal{S}_n : (G, X) \text{ has property } \mathcal{I}) = 1.$$

Throughout this paper, by the edge density $\epsilon(G)$ of a graph G we mean $|E(G)|/|V(G)|$. Also we write $g(n) = o(f(n))$, where $f(n)$ and $g(n)$ are any functions of n defined on $\mathbb{N}$, whenever $g(n)/f(n) \to 0$ as $n \to \infty$. As we mentioned earlier, the concept of homogeneous society was defined in [19] as a graph (G, X) such that, for any t and any two vertices u and v of G, one has $\Pr(X_u = t) = \Pr(X_v = t)$. Let (G, X) be any homogeneous society with edge density ϵ such that $\epsilon < \mathbb{E}(X_v)$. Then it was proved in [19] that any dynamic monopoly of (G, X) has with high probability at least $n^{1-\delta}$ elements, where δ is any arbitrary fixed positive real number and $n = |G|$. In other words, in such a network we have $\mathrm{dyn}_X(G) \ge n^{1-\delta}$ with high probability. The following question was posed in [19]. Let (G, X) be homogeneous with $\epsilon(G) < \mathbb{E}(X_v)$. Is it true that with high probability $\mathrm{dyn}_X(G) \ge n(1 - \epsilon(G)/\mathbb{E}(X_v))$? This question is discussed with details in Section 3, where a complete answer is presented.

The organisation of the paper is as follows. In Section 2 we obtain a concentration result for $\mathrm{dyn}_X(G)$ around $\mathbb{E}(\mathrm{dyn}_X(G))$ (Theorem 2.3). Next, an upper bound for $\mathbb{E}(\mathrm{dyn}_X(G))$ (Theorem 2.5) and, using the previous concentration result, another upper bound for $\mathrm{dyn}_X(G)$ (Theorem 2.6) are obtained. We show that the resulting bounds are tight (Remark 2.8). Section 3 is devoted to lower bounds for the size of dynamic monopolies in general networks. We first discuss the question raised by the second author in [19]. Two useful examples (Examples 3.1 and 3.2) lead us to investigate a suitable lower bound for $\mathrm{dyn}_X(G)$, where G satisfies some conditions. Specifically, we prove the following result (Theorem 3.5). Let $\{(G_n, X_n)\}_{n=1}^{\infty}$ be a family of graphs with probabilistic thresholds such that $|G_n| = n$. Denote the random variable corresponding

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

to any vertex v of G_n by $X_{n,v}$. Set $\overline{\alpha}_n = \sum_{v \in G_n} \mathbb{E}(X_{n,v})/n$. Suppose there is a function $f(n) = o(n)$ such that $X_{n,v} \leq f(n)\overline{\alpha}_n$, for all $v \in G_n$ and a constant $k > 1$ such that $\overline{\alpha}_n \geq k\epsilon(G_n)$ for all n. Then for any positive constant δ, with high probability, $\mathrm{dyn}_{X_n}(G_n) \geq n(1 - (\epsilon(G_n))/\overline{\alpha}_n)(1 - \delta)/f(n)$. After this result (Theorem 3.5), we obtain the lower bound $\mathrm{dyn}_{X_n}(G_n) \geq \lambda|G_n|$, where λ is a constant number, without the condition $X_{n,v} \leq f(n)\overline{\alpha}_n$ for G_n but with some extra conditions on (G_n, X_n) (Theorem 3.7). We discuss the necessity of our conditions in Theorems 3.5 and 3.7. We observe that the results in this paper are also valid for directed graphs with probabilistic thresholds.

2. Upper bounds

In this section we first obtain a concentration result for $\mathrm{dyn}_X(G)$ around its mean value. Next, an upper bound for $\mathbb{E}(\mathrm{dyn}_X(G))$ and, using the previous concentration result, another upper bound for $\mathrm{dyn}_X(G)$ are obtained. We show the sharpness of these bounds. In order to begin, we need the following result due to McDiarmid [17].

Theorem 2.1 [17]. *Let $X = (X_1, X_2, \ldots, X_n)$ be a family of independent random variables with X_k taking values in a set A_k for each k. Suppose that the real-valued function f defined on $\prod A_k$ satisfies $|f(x) - f(x')| \leq c_k$ whenever the vectors x and x' differ only in the kth coordinate. Let μ be the expected value of the random variable $f(X)$. Then for any $t \geq 0$,*

$$\Pr(|f(X) - \mu| \geq t) \leq 2\exp\Big(\frac{-2t^2}{\sum c_k^2}\Big).$$

In order to apply Theorem 2.1 for $\mathrm{dyn}_X(G)$ we need to prove the following lemma.

Lemma 2.2. *Let G be a graph and τ and τ' be two threshold assignments to the vertices of G such that $\tau(u) = \tau'(u)$ holds for all vertices u of G except for one. Then*

$$|\mathrm{dyn}_\tau(G) - \mathrm{dyn}_{\tau'}(G)| \leq 1.$$

Proof. Let v be the only vertex such that $\tau(v) \neq \tau'(v)$. Without loss of generality we may assume that $\tau'(v) > \tau(v)$. Let M be a dynamo of the minimum size for (G, τ). Because the threshold of all vertices other than v, are the same in τ and τ', it is clear that $\mathrm{dyn}_{\tau'}(G) \geq \mathrm{dyn}_\tau(G)$. On the other hand, $M' = M \cup \{v\}$ is a dynamic monopoly for (G, τ'). Then $\mathrm{dyn}_{\tau'}(G) \leq |M| + 1 = \mathrm{dyn}_\tau(G) + 1$, and hence

$$|\mathrm{dyn}_\tau(G) - \mathrm{dyn}_{\tau'}(G)| \leq 1. \qquad \square$$

The following theorem shows the concentration of $\mathrm{dyn}_X(G)$ around $\mathbb{E}(\mathrm{dyn}_X(G))$.

Theorem 2.3. *Let (G, X) be given and $|G| = n$. Let $\omega(n)$ be any arbitrary slowly increasing function with $\omega(n) \to \infty$ as $n \to \infty$. Then, with high probability,*

$$|\mathrm{dyn}_X(G) - \mathbb{E}(\mathrm{dyn}_X(G))| \leq \omega(n)\sqrt{n}.$$

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

PROOF. Let the vertices of G be $v_1, v_2, \ldots, v_n$ and X_i the threshold chosen by v_i. Set $A_i = [0, \deg(v_i) + 1]$ and, for any $x \in \prod A_i$, define $f(x) = \mathrm{dyn}_x(G)$. Set also $t = \omega(n)\sqrt{n}$. Now by applying Theorem 2.1 with $c_i = 1$ for all i and using Lemma 2.2, we have

$$\Pr(|f(X) - \mu| \geq \omega(n)\sqrt{n}) \leq 2\exp(-2\omega^2(n)).$$

Note that $2\exp(-2\omega^2(n)) = o(1)$. The proof is complete. □

We also need the following result due to Ackerman *et al.* [1].

THEOREM 2.4 [1]. *Let G be a graph and τ be a threshold assignment for the vertices of G. Then*

$$\mathrm{dyn}_\tau(G) \leq \sum_{v \in V(G)} \frac{\tau(v)}{\deg(v) + 1}.$$

The promised upper bound for $\mathbb{E}(\mathrm{dyn}_X(G))$ is as follows.

THEOREM 2.5. *Let (G, X) be given with $X = (X_i)_{i=1}^n$ where, for each i, $1 \leq i \leq n$, X_i corresponds to the vertex v_i. Then*

$$\mathbb{E}(\mathrm{dyn}_X(G)) \leq \sum_{i=1}^{n} \frac{\mathbb{E}(X_i)}{\deg(v_i) + 1}.$$

PROOF. It is enough to take the expectation of both sides of the inequality given in Theorem 2.4 and then use the linearity property of the expectation. □

The following theorem gives an upper bound for $\mathrm{dyn}_X(G)$. Its proof is immediately obtained using Theorems 2.5 and 2.3.

THEOREM 2.6. *Let G be a graph with $V(G) = \{v_1, v_2, \ldots, v_n\}$ and $X = (X_1, X_2, \ldots, X_n)$ be a random threshold assignment for G. Also let $\omega(n)$ be any arbitrary slowly increasing function with $\omega(n) \to \infty$ as $n \to \infty$. Then, with high probability,*

$$\mathrm{dyn}_X(G) \leq \sum_{i=1}^{n} \frac{\mathbb{E}(X_i)}{\deg(v_i) + 1} + \omega(n)\sqrt{n}.$$

In the following we show that the upper bounds of Theorems 2.3 and 2.6 are the best possible in the sense that if we replace $\omega(n)\sqrt{n}$ in both bounds by $\lambda\sqrt{n}$ where λ is any fixed positive number, then the assertions are not valid. For this purpose we need the central limit theorem (see, for example, [13]) as follows.

THEOREM 2.7. *Let $Y_1, Y_2, \ldots$ be a sequence of independent identically distributed random variables with finite means μ and finite nonzero variances σ^2, and let $S_n = Y_1 + Y_2 + \cdots + Y_n$. Then, as $n \to \infty$,*

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{D} N(0, 1).$$

We make the following remark concerning Theorems 2.3 and 2.6.

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

REMARK 2.8. The upper bound of Theorem 2.6 is the best possible.

PROOF. Consider the complete graph K_n on n vertices, where each vertex chooses the threshold 1 with probability half and the threshold $n-1$ with probability half. Denote the threshold of the vertex i by X_i. Let $Y_i = (X_i - 1)/(n-2)$ and $Y = \sum Y_i$. We have $\mathbb{E}(Y_i) = \frac{1}{2}$ and $\mathrm{Var}(Y_i) = \frac{1}{4}$. The random variable Y counts the number of vertices with threshold $n-1$. We make the following claim.

Claim.

$$\mathrm{dyn}_X(K_n) = \begin{cases} Y-1 & \text{if } Y \geq 2, \\ 1 & \text{if } Y < 2. \end{cases}$$

To prove the claim, assume that U is the set of vertices with threshold $n-1$ in K_n. Note that $Y = |U|$. Clearly any dynamic monopoly of K_n needs at least $Y-1$ vertices from U. Also when $Y \geq 2$, each set consisting of $Y-1$ vertices of U is a dynamic monopoly. This proves the claim in case $Y \geq 2$. If $Y < 2$, each vertex of the graph is a dynamic monopoly and any such set needs at least one vertex. This completes the proof of the claim.

We now apply Theorem 2.7 for $Y_1, \ldots, Y_n$ and obtain the following inequalities. Choose η such that $0 < \eta < (1/\sqrt{2\pi}) \exp(-(2\lambda+2)^2/2)$. Then

$$\begin{aligned}
\Pr\Big(\mathrm{dyn}_X(K_n) - \frac{n}{2} > \lambda\sqrt{n}\Big) &= \Pr\Big(Y - 1 - \frac{n}{2} > \lambda\sqrt{n}\Big) \\
&= \Pr\Big(\frac{Y - n/2}{\sqrt{n/4}} > 2\lambda + \frac{1}{\sqrt{n/4}}\Big) \\
&\geq \Pr\Big(\frac{Y - n/2}{\sqrt{n/4}} > 2\lambda + 1\Big) \\
&\geq \int_{2\lambda+1}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\, dx - \eta \\
&> \frac{1}{\sqrt{2\pi}} \exp(-(2\lambda+2)^2/2) - \eta,
\end{aligned}$$

which implies that for any positive constant λ, the statement that $|\mathrm{dyn}_X(G) - \mathbb{E}(\mathrm{dyn}_X(G))| \leq \lambda\sqrt{n}$ holds whp for all graphs is not valid. □

3. Lower bounds and related discussions

The concept of homogeneous societies was introduced in [19] as follows. Let (G, X) be given and $X = (X_v)_{v \in G}$. Assume that for any value t, $\Pr(X_v = t)$ is the same for all vertices v of G. Then (G, X) is called a homogeneous society. Let G be any graph on n vertices with the edge density $\epsilon = \epsilon(G)$. In [19, Theorem 4], it was proved that if G is homogeneous and $\alpha = \mathbb{E}(X_v) > \epsilon(G)$ then for any positive δ, with high probability any dynamic monopoly of G has at least $n^{1-\delta}$ vertices. But in the proof of that theorem we need to have $\alpha > \epsilon(G)$ even when the order of G tends to infinity. Since α and ϵ are both functions of $|G|$, in order to fulfil this condition we should

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

replace $\alpha > \epsilon(G)$ by $\alpha \geq k\epsilon(G)$ in [19, Theorem 4], where $k > 1$ can be any positive fixed number. In order to observe why we need $\alpha \geq k\epsilon(G)$, it is enough to consider $G = K_{1,n-1}$ where each vertex chooses the threshold 1 with probability 1. We have $\mathrm{dyn}(G) = 1$ and hence the corresponding assertion does not hold for this example. Also in [19], the following question concerning homogeneous networks was raised. Is it true that $\mathrm{dyn}(G) \geq n(1 - \epsilon(G)/\alpha)$? The necessary comment concerning this question is that the lower bound $n(1 - \epsilon(G)/\alpha)$ for $\mathrm{dyn}(G)$ should be replaced by $n(1 - \epsilon(G)/\alpha)(1 - \delta)$, where δ is any arbitrarily small positive and fixed number, otherwise the lower bound does not hold. The following guiding example leads us to obtain an appropriate lower bound for $\mathrm{dyn}_X(G)$. Then in Theorem 3.5 we show the validity of this lower bound.

Example 3.1. Let k be any positive integer divisible by 4 and set $n = 2k$. Let $G_n = K_{k,k}$ whose bipartite sets are A and B. We consider the probabilistic threshold assignment X for G_n as follows. For each vertex $v \in A$, set $\Pr(X_v = k) = \frac{3}{4}$ and $\Pr(X_v = 0) = \frac{1}{4}$. And for each vertex $v \in B$, set $\Pr(X_v = 3k/4) = 1$ (that is each vertex in B has deterministic threshold $3k/4$). It is clear that for each vertex v of G, $\mathbb{E}(X_v) = 3k/4$. In the following we show that for any positive constant δ, whp $(1 - \delta)n/4 \leq \mathrm{dyn}_X(G_n) \leq (1 + \delta)n/4$.

For this purpose, for each vertex $v \in A$, define the random variable

$$Y_v = \begin{cases} 1 & \text{if } X_v = 0, \\ 0 & \text{if } X_v = k. \end{cases}$$

It is clear that $\mu = \mathbb{E}(Y_v) = \frac{1}{4}$. Set $S_k = \sum_{v \in A} Y_v$. In fact the random variable S_k counts the number of vertices in the bipartite set A with threshold 0. It is easy to see that $\mathrm{dyn}_X(G_n) = 3k/4 - S_k$.

Now we want to prove that with high probability $(1 - \delta)n/4 \leq \mathrm{dyn}_X(G_n) \leq (1 + \delta)n/4$. Set for simplicity $\clubsuit = \Pr((1 - \delta)n/4 \leq \mathrm{dyn}_X(G_n) \leq (1 + \delta)n/4)$. Then

$$\begin{aligned} \clubsuit &= \Pr\Big(\frac{k}{2}(1 - \delta) \leq \frac{3k}{4} - S_k \leq \frac{k}{2}(1 + \delta)\Big) \\ &= \Pr\Big(\frac{k}{4} - \frac{\delta k}{2} \leq S_k \leq \frac{k}{4} + \frac{\delta k}{4}\Big) \\ &= \Pr\Big(-\frac{\delta k/2}{\sqrt{3k/16}} \leq \frac{S_k - k\mu}{\sqrt{k\sigma^2}} \leq \frac{\delta k/2}{\sqrt{3k/16}}\Big). \end{aligned}$$

Since $\int_{-\infty}^{+\infty} (1/\sqrt{2\pi}) \exp(-x^2/2)\, dx = 1$, for any positive constant ε there is constant N such that

$$\int_{-N}^{+N} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\, dx \geq 1 - \frac{\varepsilon}{2}. \tag{3.1}$$

We may assume that k is large enough such that $(\delta k/2)/\sqrt{3k/16} \geq N$ and

$$\Pr\Big(-N \leq \frac{S_k - k\mu}{\sqrt{k\sigma^2}} \leq +N\Big) \geq \int_{-N}^{+N} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\, dx - \frac{\varepsilon}{2}.$$

These inequalities together with (3.1) imply

$$\Pr\Big(-\frac{\delta k/2}{\sqrt{3k/16}} \le \frac{S_k - k\mu}{\sqrt{k\sigma^2}} \le +\frac{\delta k/2}{\sqrt{3k/16}}\Big) \ge 1-\varepsilon.$$

Then, since the left-hand side is equal to ♣,

$$\Pr((1-\delta)n/4 \le \mathrm{dyn}_X(G_n) \le (1+\delta)n/4) \ge 1-\varepsilon.$$

But since $\varepsilon > 0$ is arbitrary, with high probability $(1-\delta)n/4 \le \mathrm{dyn}_X(G_n) \le (1+\delta)n/4)$.

For each graph G_n in Example 3.1 consider the combination $|G_n|(1-\epsilon(G_n)/\alpha(G_n))$. We have $|G_n|(1-\epsilon(G_n)/\alpha(G_n)) = n/3$. Our result above in Example 3.1 concerning $\mathrm{dyn}_X(G_n)$ shows that the inequality $\mathrm{dyn}_X(G_n) \ge |G_n|(1-\epsilon(G_n)/\alpha(G_n))(1-\delta)$ (whp) does not hold. But we note that the inequality $\mathrm{dyn}_X(G_n) \ge |G_n|(1-\epsilon(G_n)/\alpha(G_n))$ $(1-\delta)(3/4)$ holds for the family $\{G_n\}_n$ with high probability. This suggests that a reasonable lower bound for $\mathrm{dyn}_X(G_n)$ is $|G_n|(1-\epsilon(G_n)/\alpha(G_n))(1-\delta)/\kappa$ where κ is any constant depending only on the order of graph. Fortunately, Theorem 3.5 shows that this lower bound is valid. Before we present Theorem 3.5, let us show by an example that even for homogeneous networks the inequality $\mathrm{dyn}_X(G) \ge |G|(1-\epsilon(G)/\alpha(G))$ does not hold.

Example 3.2. Let G_n consist of $k = n/2$ vertex disjoint copies of K_2. Consider the probabilistic threshold assignment X_v for each vertex v of G_n as follows: $\Pr(X_v = 1) = \frac{3}{4}$ and $\Pr(X_v = 0) = \frac{1}{4}$. It is clear that G_n is a homogeneous network with $\alpha = \frac{3}{4}$ and $\epsilon(G_n) = \frac{1}{2}$. The inequality $\mathrm{dyn}_X(G_n) \ge n(1-\epsilon(G_n)/\alpha(G_n))$ is equivalent to $\mathrm{dyn}_X(G_n) \ge n/3$. We make the following claim.

Claim. As $n \to \infty$, $\Pr(\mathrm{dyn}_{X_n}(G_n) \le 9n/32) \nrightarrow 0$.

For each edge $e = uv \in E(G_n)$, define the random variable $Y_e = X_u X_v$. It is clear that

$$\Pr(Y_e = 1) = \Pr(X_u X_v = 1) = 3/4 \times 3/4 = 9/16.$$

Set $S_k = \sum_{e\in E(G)} Y_e$. It is clear that $\mathrm{dyn}_X(G_n)$ is the number of edges with both of its vertices having threshold 1. In other words, $\mathrm{dyn}_X(G_n) = S_k$. We use Theorem 2.7 for S_k with the assumption that k is large enough and obtain the inequalities

$$\begin{aligned}\Pr\Big(\mathrm{dyn}_{X_n}(G_n) \le \frac{9k}{16}\Big) &= \Pr\Big(S_k \le \frac{9k}{16}\Big)\\ &= \Pr\Big(\frac{S_k - k\mu}{\sqrt{k\sigma^2}} \le 0\Big)\\ &\ge \Big(\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\,dx\Big) - \delta\\ &= \frac{1}{2} - \delta.\end{aligned}$$

It follows that $\Pr(\mathrm{dyn}_{X_n}(G_n) \le 9n/32) \nrightarrow 0$. This contradicts the validity of $\mathrm{dyn}_X(G_n) \ge n/3$ whp.

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

We now prove lower bounds for $\mathrm{dyn}_X(G)$. The first result is Theorem 3.5. In the proof of Theorem 3.5 we shall make use of the following concentration result obtained by McDiarmid [17].

THEOREM 3.3 [17]. *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables such that $0 \le X_k \le 1$, for each $k \in \{1, \ldots, n\}$. Let $S_n = \sum_k X_k$ and $\mu = \mathbb{E}(S_n)$. Then, for any positive number δ,*

$$\Pr(|S_n - \mu| \ge \delta\mu) \le 2e^{-(1/3)\delta^2\mu}.$$

We also need the following result concerning dynamic monopolies from [14].

THEOREM 3.4 [14]. *Let G be a graph and τ be a threshold assignment to the vertices of G. Let $\bar{t}$ and t_M denote the average and maximum threshold of τ, respectively. For any dynamic monopoly M of (G, τ),*

$$|M| \ge |G|\Big(1 - \frac{\epsilon(G)}{\bar{t}}\Big)\frac{\bar{t}}{t_M}.$$

Let $\{(G_n, X_n)\}_{n=1}^{\infty}$ be any family of graphs with probabilistic thresholds. In the following theorem we denote the random variable corresponding to any vertex v of G_n by $X_{n,v}$.

THEOREM 3.5. *Let $\{(G_n, X_n)\}_{n=1}^{\infty}$ be a family of graphs with probabilistic thresholds, where $|G_n| = n$. Set $\overline{\alpha}_n = \sum_{v\in G_n} \mathbb{E}(X_{n,v})/n$. Suppose that there is a function $f(n) = o(n)$ such that $X_{n,v} \le f(n)\overline{\alpha}_n$, for all $v \in G_n$, and that there is a constant $k > 1$ such that $\overline{\alpha}_n \ge k\epsilon(G_n)$ for all n. Then for any positive constant δ, with high probability,*

$$\mathrm{dyn}_{X_n}(G_n) \ge n\Big(1 - \frac{\epsilon(G_n)}{\overline{\alpha}_n}\Big)\frac{1-\delta}{f(n)}.$$

PROOF. Let the threshold assignments in G_n be $X_n^1, X_n^2, \ldots, X_n^n$. Also let $\mathbb{E}(X_n^i) = \alpha_n^i$ for each $1 \le i \le n$. Set $Y_n^i = X_n^i/(f(n)\overline{\alpha}_n)$. We have $0 \le Y_n^i \le 1$ because by the assumption $0 \le X_n^i \le f(n)\overline{\alpha}_n$. Set $Y_n = \sum_{1\le i\le n} Y_n^i$. Then, by the linearity of expectation,

$$\mathbb{E}(Y_n) = \sum_{i=1}^{n} \mathbb{E}(Y_n^i) = \sum_{i=1}^{n} \frac{\alpha_n^i}{f(n)\overline{\alpha}_n} = \frac{n\overline{\alpha}_n}{f(n)\overline{\alpha}_n} = \frac{n}{f(n)}.$$

By applying Theorem 3.3 for $Y_n^1, \ldots, Y_n^n$ and Y_n we obtain

$$\Pr\Big(\Big|Y_n - \frac{n}{f(n)}\Big| \ge \frac{(k-1)\delta}{k}\frac{n}{f(n)}\Big) \le 2\exp\Big(-\frac{1}{3}\Big(\frac{(k-1)\delta}{k}\Big)^2\frac{n}{f(n)}\Big).$$

Since $f(n) = o(n)$, the right-hand side of the inequality tends to 0. It follows that with high probability $|Y_n - n/f(n)| \le ((k-1)\delta/k)(n/f(n))$ or equivalently $|(f(n)\overline{\alpha}_n/n)Y_n - \overline{\alpha}_n| \le ((k-1)\delta/k)\overline{\alpha}_n$, and then $|(\sum_{i=1}^{n} X_n^i/n) - \overline{\alpha}_n| \le ((k-1)\delta/k)\overline{\alpha}_n$. Denote the average threshold $\sum_i X_n^i/n$ by $\bar{t}_n$. It follows that, with high probability,

$$\overline{\alpha}_n\Big(1 - \frac{k-1}{k}\delta\Big) \le \bar{t}_n \le \overline{\alpha}_n\Big(1 + \frac{k-1}{k}\delta\Big).$$

Now, by using the last inequality and Theorem 3.4, with high probability,

$$\begin{aligned}\mathrm{dyn}_{X_n}(G_n) &\geq n\Big(1 - \frac{\epsilon(G_n)}{(1-(k-1/k)\delta)\overline{\alpha}_n}\Big)\frac{(1-(k-1/k)\delta)\overline{\alpha}_n}{f(n)\overline{\alpha}_n} \\ &= n\Big(1 - \frac{k-1}{k}\delta - \frac{\epsilon(G_n)}{\overline{\alpha}_n}\Big)\frac{1}{f(n)} \\ &\geq n\Big(1 - \frac{\epsilon(G_n)}{\overline{\alpha}_n}\Big)(1-\delta)\frac{1}{f(n)}.\end{aligned}$$

which completes the proof. □

In [19], a family of graphs $\{(G_n, \tau_n)\}_n$, where τ_n is a deterministic threshold assignment for the vertices of G_n, is said to be a dynamo-unbounded family if there exists a function $g(n)$ with $g(n) \to \infty$ as $n \to \infty$ such that $\mathrm{dyn}_{\tau_n}(G_n) \geq g(|G_n|)$. Similarly, we define $\{(G_n, X_n)\}_n$ to be a dynamo-unbounded family, if there exists a function $g(n)$ with $g(n) \to \infty$ as $n \to \infty$ such that with high probability $\mathrm{dyn}_{X_n}(G_n) \geq g(|G_n|)$. Homogenous networks with $\alpha \geq k\epsilon$ and networks satisfying the conditions of Theorem 3.5 are two examples of dynamo-unbounded families. Even if in Theorem 3.5, $f(n) \to \infty$, the related family is dynamo-unbounded, since $n/f(n) \to \infty$. Note also that if we remove the condition $f(n) = o(n)$ in Theorem 3.5 then the family is not necessarily dynamo-unbounded. For this claim, consider $G_n = K_{1,n-1}$, where the threshold of each vertex $v \in G_n$ is $X_{n,v} = \deg_{G_n}(v)$ (that is deterministic threshold). The conditions of Theorem 3.5 except $f(n) = o(n)$ hold for G_n, but we have $\mathrm{dyn}_{X_n}(G_n) = 1$. Therefore $\{G_n\}_n$ is not a dynamo-unbounded family.

In general if a family $\mathcal{F}$ does not satisfy the condition $X_{n,v} \leq f(n)\alpha_n$, Theorem 3.5 asserts nothing about the dynamo-unboundedness of $\mathcal{F}$. For such networks Theorem 3.7 can be applied. In the proof of Theorem 3.7, we make use of the following lemma from [19].

Lemma 3.6 [19]. *Let (G, τ) be a graph and ϵ be the edge density of G. For any dynamic monopoly M of G,*

$$\sum_{v \notin M} \tau(v) \leq |E(G)|.$$

The following theorem introduces another dynamo-unbounded family of graphs under some conditions.

Theorem 3.7. *Let $\{(G_n, X_n)\}_{n=1}^{\infty}$ be a sequence of graphs with probabilistic thresholds such that $|G_n| = n$. Let $X_n^1, \ldots, X_n^n$ be the threshold assignments in G_n. Assume that for each n there exists a constant α_n such that for each i with $1 \leq i \leq n$, $\mathbb{E}(X_n^i) = \alpha_n$. Assume also that there exist a constant $k > 1$ such that for all n, $\alpha_n \geq k\epsilon(G_n)$ and a positive constant s such that $\epsilon(G_n) \geq s$. Then for some positive constant λ, with high probability, $\mathrm{dyn}_{X_n}(G_n) \geq \lambda n$.*

Proof. Let δ be a sufficiently small positive number such that $1 - \delta > 1/k$. Let $c > 4$ be such that $c - 4/c(1-\delta) > 1/k \geq \epsilon(G_n)/\alpha_n$ and $cs \geq 2$. It follows that for some positive constant ξ, $\xi < (c - 4/c)(1 - \delta) - \epsilon(G_n)/\alpha_n$, for all n. Now partition the vertex set of G_n into subsets U_n and W_n, such that for each vertex $v \in U_n$, $\deg(v) + 1 > c\alpha_n$ and for each vertex $v \in W_n$, $\deg(v) + 1 \leq c\alpha_n$. The following inequalities hold:

$$|U_n|c\epsilon(G_n) \leq |U_n|ck\epsilon(G_n) \leq |U_n|c\alpha_n \leq \sum_{v \in U_n} (\deg(v) + 1)$$
$$\leq \sum_{v \in V(G_n)} \deg(v) + |U_n| \leq 2n\epsilon(G_n) + |U_n|.$$

Then $|U_n|(c\epsilon(G_n)/2) \leq 2n\epsilon(G_n)$ or equivalently $|U_n| \leq 4n/c$ and therefore $|W_n| \geq (c - 4)n/c$. Applying Theorem 3.3 to the set W_n and proceeding as in the proof of Theorem 3.5, we have, with high probability,

$$\alpha_n(1 - \delta) \leq \frac{\sum_{v_i \in W_n} X_n^i}{|W_n|} \leq \alpha_n(1 + \delta). \tag{3.2}$$

Let $M(X_n)$ be a dynamic monopoly of graph (G_n, X_n). Using Lemma 3.6, we obtain

$$\sum_{v_i \in W_n \setminus M(X_n)} X_n^i \leq \sum_{v_i \notin M(X_n)} X_n^i \leq |E(G_n)|$$

or, equivalently,

$$\sum_{v_i \in W_n} X_n^i - \sum_{v_i \in W_n \cap M(X_n)} X_n^i \leq |E(G_n)|.$$

The latter inequality, together with (3.2), implies

$$|W_n|\alpha_n(1 - \delta) - n\epsilon(G_n) \leq \sum_{v_i \in W_n} X_n^i - |E(G_n)| \leq \sum_{v_i \in W_n \cap M(X_n)} X_n^i \leq |W_n \cap M(X_n)|\alpha_n c.$$

Now $|W_n| \geq (c - 4)n/c$, so $((c - 4)n/c)\alpha_n(1 - \delta) - n\epsilon(G_n) \leq |W_n \cap M(X_n)|\alpha_n c$ and

$$\frac{n}{c}\left(\frac{c - 4}{c}(1 - \delta) - \frac{\epsilon(G_n)}{\alpha_n}\right) \leq |W_n \cap M(X_n)| \leq |M(X_n)|.$$

Recall that $\xi < ((c - 4)/c)(1 - \delta) - \epsilon(G_n)/\alpha_n$. It follows that $n\xi/c \leq |M(X_n)|$, as desired. □

In Theorem 3.7 we have the following condition. For each n there exists a constant α_n such that, for all i with $1 \leq i \leq n$, $\mathbb{E}(X_n^i) = \alpha_n$. The following example shows that Theorem 3.7 is not valid without this condition. For each n, let $G_n = K_{1,n-1}$. Assume that each vertex $v \in G_n$ has deterministic threshold $X_{n,v} = \deg_{G_n}(v)$, but all other conditions of Theorem 3.7 hold for G_n. Note that $\mathrm{dyn}_{X_n}(G_n) = 1$. This violates the assertion of Theorem 3.7.

Finally, we show that the condition $\epsilon(G_n) \geq s$ in Theorem 3.7 is necessary. For this purpose, let G_n be the graph on n vertices with only one edge. Let $X = (X_v)_{v \in G_n}$ be a threshold assignment for G_n such that for all $v \in G_n$, $\Pr(X_v = 0) = (n - 2)/2$ and $\Pr(X_v = 1) = 2/n$. We have $\epsilon(G_n) = 1/n$ and $\alpha_n = 2/n$. Note that $\epsilon(G_n) \to 0$. The assertion of Theorem 3.7 does not hold for $\{G_n\}_n$. In fact by Theorem 3.7, with

https://doi.org/10.1017/S0004972714000604 Published online by Cambridge University Press

high probability, $\mathrm{dyn}_{X_n}(G_n) \geq n(1 - (1/n)/(2/n))(1 - \delta) = (n/2)(1 - \delta)$. On the other hand, for large n the number of vertices with threshold 1 in G_n obeys the Poisson distribution. Consequently, the probability that there exist no vertices of threshold 1 is $(2^0/0!)e^{-2}$. This shows that with the same probability $\mathrm{dyn}_X(G_n) = 0$. This clearly violates the assertion of Theorem 3.7 for this family. Hence the condition $\epsilon(G_n) \geq s$ is necessary in Theorem 3.7.

4. Concluding remarks

We conclude by observing that the results of this paper can be extended to directed graphs. Let G be a directed graph whose minimum in-degree is at least one. Also let $\tau : V(G) \to \mathbb{N}$ be an assignment of thresholds to the vertices of G such that $\tau(v) \leq \deg^-(v)$, for each vertex v, where $\deg^-(v)$ stands for the in-degree of v. A subset M of vertices of G is called a dynamic monopoly for (G, τ) if the vertex set of G can be partitioned into $D_0 \cup \cdots \cup D_t$ such that $D_0 = M$ and, for each $i \geq 1$ and each $v \in D_i$, the number of edges from $D_0 \cup \cdots \cup D_{i-1}$ to v is at least $\tau(v)$. Dynamic monopolies of directed graphs were studied in [1, 6, 15]. As proved in [1], Theorem 2.4 is also valid for directed graphs. Lemma 3.4 holds for directed graphs too. Let X be a probabilistic threshold assignment for the vertices of a directed graph G. Similar to the case of undirected graphs, we can define $\mathrm{dyn}_X(G)$. The directed version of the results of the paper are also valid for directed graphs with probabilistic threshold assignments. The proofs are similar and we omit the details.

References

[1] E. Ackerman, O. Ben-Zwi and G. Wolfovitz, 'Combinatorial model and bounds for target set selection', *Theoret. Comput. Sci.* **411** (2010), 4017–4022.

[2] S. S. Adams, P. Booth, D. S. Troxell and S. L. Zinnen, 'Modeling the spread of fault in majority-based network systems: dynamic monopolies in triangular grids', *Discrete Appl. Math.* **160** (2012), 1624–1633.

[3] J. Balogh and B. Bollobás, 'Bootstrap percolation on the hypercube', *Probab. Theory Related Fields* **134** (2006), 624–648.

[4] C. C. Centeno, M. C. Dourado, L. D. Penso, D. Rautenbach and J. L. Szwarcfiter, 'Irreversible conversion of graphs', *Theoret. Comput. Sci.* **412** (2011), 3693–3700.

[5] C.-L. Chang and Y.-D. Lyuu, 'Spreading messages', *Theoret. Comput. Sci.* **410** (2009), 2714–2724.

[6] C.-L. Chang and Y.-D. Lyuu, On irreversible dynamic monopolies in general graphs, www.arXiv.org.

[7] C.-L. Chang and Y.-D. Lyuu, 'Bounding the sizes of dynamic monopolies and convergent sets for threshold-based cascades', *Theoret. Comput. Sci.* **468** (2013), 37–49.

[8] N. Chen, 'On the approximability of influence in social networks', *SIAM J. Discrete Math.* **23** (2009), 1400–1415.

[9] C.-Y. Chiang, L.-H. Huang, B.-J. Li, J. Wu and H.-G. Yeh, 'Some results on the target set selection problem', *J. Comb. Optim.* **25** (2013), 702–715.

[10] P. Domingos and M. Richardson, 'Mining the network value of customers', *Proc. 7th ACM Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, (2001), 57–66.

[11] P. A. Dreyer and F. S. Roberts, 'Irreversible k-threshold processes: graph-theoretical threshold models of the spread of disease and of opinion', *Discrete Appl. Math.* **157** (2009), 1615–1627.

[12] P. Flocchini, R. Kralovic, A. Roncato, P. Ruzicka and N. Santoro, 'On time versus size for monotone dynamic monopolies in regular topologies', *J. Discrete Algorithms* **1** (2003), 129–150.
[13] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 2nd edn (Oxford University Press, Oxford, 1992).
[14] K. Khoshkhah, H. Soltani and M. Zaker, 'On dynamic monopolies of graphs: The average and strict majority thresholds', *Discrete Optim.* **9** (2012), 77–83.
[15] K. Khoshkhah, H. Soltani and M. Zaker, 'Dynamic monopolies in directed graphs: The spread of unilateral influence in social networks', *Discrete Appl. Math.* **171** (2014), 81–89.
[16] T. Kulich, 'Dynamic monopolies with randomized starting configuration', *Theoret. Comput. Sci.* **412** (2011), 6371–6381.
[17] C. McDiarmid, 'Concentration', in: *Probabilistic Methods for Algorithmic Discrete Mathematics* (Springer, New York, 1998).
[18] A. Nichterlein, R. Niedermeier, J. Uhlmann and M. Weller, 'On tractable cases of target set selection', in: *Algorithms and computation. Part I*, Lecture Notes in Computer Science, 6506 (Springer, Berlin, 2010).
[19] M. Zaker, 'On dynamic monopolies of graphs with general thresholds', *Discrete Math.* **312** (2012), 1136–1143.
[20] M. Zaker, 'Generalized degeneracy, dynamic monopolies and maximum degenerate subgraphs', *Discrete Appl. Math.* **161** (2013), 2716–2723.

HOSSEIN SOLTANI, Department of Mathematics,
Institute for Advanced Studies in Basic Sciences,
Zanjan 45137-66731, Iran
e-mail: h_soltani@iasbs.ac.ir

MANOUCHEHR ZAKER, Department of Mathematics,
Institute for Advanced Studies in Basic Sciences,
Zanjan 45137-66731, Iran
e-mail: mzaker@iasbs.ac.ir