Majority Vote and Monopolies in Social Networks

Thesis submitted in partial fulfillment of the requirement for the degree of master of science in the Faculty of Engineering

Submitted by: **Mr. Assaf Mizrachi**

Advisor: **Dr. Chen Avin**

Department of Communication Systems Engineering

Faculty of Engineering

Ben-Gurion University of the Negev

July 30, 2013

Abstract

On various occasions, a society has to reach a decision regarding a question affecting the life of its members. For this, it may use a *voting* mechanism, i.e., collect the votes of the group members and output a decision which best expresses the group's will. In order to make up their minds, individuals often discuss the issue with friends, family and colleagues before they take their vote, and by that may mutually affect each other's vote. Individuals are also, to some extent, influenced by the opinions of key figures in their culture, such as politicians, officers, publicists, writers and so on, who are commonly considered as the "elite" of the society. This work studies the "power of the elite": to what extent can the elite of a social network influence the rest of society to accept its opinion, and thus become a *monopoly*. We present an empirical study of local majority voting in social networks, where the elite forms a coalition against all other (common) nodes. The results, obtained on several social networks, show that an elite of size $\sqrt{m}$ (where m is the number of connections) has disproportionate power, relative to its size, with respect to the rest of society: it wins the majority voting and remains stable over time, in contrast to the predictions of a preferential attachment model.

Keywords: social networks, elite, power of elite, majority voting, coalitions, monopolies, diffusion.

Acknowledgements

I want to thank my academic instructor, Dr. Chen Avin, who thought me how research should go step by step, to not despair from “bad” results and suspect those which are too good to be true. Also, thank you for making my research fun and for your commitment and dedication who’s without, this study would not have been successful.

I want to express my deep gratitude to the faculty of the CSE department at Ben-Gurion University and to the supervisory committee, Dr. Zvi Lotker and Prof. David Peleg, whose great knowledge inspired my work.

Last but not least I want to thank my family. First I want to thank my grandmother who once told me to stay cool if I failed while trying my best. I want to thank my mom and dad for their care and dedication to our education from the day we were born up until now. I am grateful to my lovely wife, Shimrit, whose support and encouragement made this work possible. Thank you, my kids, for just being.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since ancient times, there were numerous occasions in which tribes, nations and other groups of people had to reach a decision regarding a question affecting the life and future of their members. Nowadays, democracies use election systems in order to choose government officials and parliament members who are to set the nation's agenda for the future. In ancient Greece, people used to gather at the city center to take part in polls on issues affecting the affairs of the city (similar to a modern time referendum). In all cases, a group of people who wish to "push" their agenda has to use "politics" – the art of influencing people on a civic or individual level. In order to make up their minds on an issue which is about to be put to vote, people usually discuss the issue with our social environment, including friends, family and colleagues, before taking the vote. This process allows people to mutually affect each other's vote. In addition, individuals are, to some level, influenced by the opinions of key figures in their culture, such as politicians, publicists, writers, former top military officers, celebrities and others. These figures are commonly considered as the *elite* of the society (we discuss several definitions of the elite in section 1.2), and due to their influence on the masses it is natural to think that these "opinion leaders" have a major impact on the final result. This raises the interesting question whether one could measure how large this influence really is.

Clearly, there are also other sources affecting the final vote of individuals, such as their backgrounds, interests and personal tastes. The goal of the present work is to isolate the element of social influence from all other sources of influence, and study only the influence that members of society have on each other's opinions, and particularly, the influence of the elite's opinions on the rest of society.

In order to do so, we examine a hypothetical social voting system on a social network represented by a weighted graph, which runs as follows. In the morning of the decision day (e.g., the elections), each individual (represented by a node in the network) sets its own initial (subjective) vote. We refer to this as the *initialization* phase. Then, in the *update* phase, each individual starts an iterative process, in which it gradually modifies its vote. Specifically, in each round it gathers votes from all sources known to it directly in the network (i.e., its neighbors in the graph), representing friends and family, colleagues, mass media etc., and then changes its vote by adopting the majority of all the votes it has seen. Since one may also be influenced indirectly (e.g., by the friend of a friend), it may take more than one such update round until the system reaches "convergence"[1]. We define the convergence state of the system to be the first round when there were no vote changes from the previous round. Once convergence is reached, the voting process stops and the update phase ends. At the final phase, we count the final votes of all individuals, and the opinion that won the most votes wins the poll. The simplest case is that of a poll involving "Yes" or "No" questions and a democratic voting, i.e., in the *final* phase, each vote counts the same. However, during the *update* phase (only), we study both *uniform* and *non-uniform* influences of nodes. In particular, we may grant a higher influence to certain subgroups of nodes (e.g., the elite), indicating their higher significance. We denote this by *influence factor*. The process of updating the vote of each node is then based on the sum of its neighbors votes weighted by their influence factor. Note that the higher influence is granted *per node* in the initialization phase, and not *per vote* ("Yes" or "No"), i.e. unlike the the vote, it cannot "propagate" from one granted node to its friends.

Our focus in this work is on the power of "cooperating" groups in social networks.

[1]Theoretically, the system might never converge, cf. [47]; for now, we assume it does.

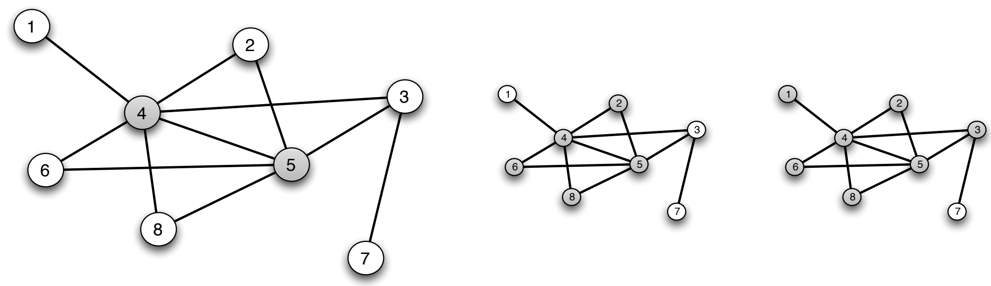


Figure 1.1: Example of majority voting in a social network: (A) An initial state with a coalition of nodes 4 and 5. (B) The final votes in the irreversible model with coalition influence factor 1. (C) The final votes in the reversible model with coalition influence factor 2.

In this context, a central notion is that of a *coalition*, i.e., a group of individuals that vote unanimously in order to influence the resulting decision. For the purpose of this study, a coalition is called a *monopoly* if at the end of the voting process, it convinces more than 50% of the population to vote according to the coalition's initial vote. As mentioned earlier, a coalition may or may not be a monopoly, and this depends to some extent on whether it is granted higher influence or not.

We also consider two types of annexation behaviors a coalition can take. In the *reversible* mode, the coalition plays *fair*, in the sense that each of its members can change its vote in each round according to the above rules (i.e., its vote is reversible). In contrast, in the *irreversible* mode the coalition plays *unfair* and its members (and only them) stick to their initial vote and ignore the majority vote rules during the update phase.

Consider for example the network in Fig 1.1, where nodes colors denote a "Yes"/"No" vote. In (A), nodes 4 and 5 form a coalition with initial vote "Yes". (B) shows the result of the voting process after one round, in the uniform, irreversible mode. Note the this is also the stable state of the voting and nodes 4 and 5 are also a monopoly, since they have succeeded in convincing the majority of the nodes in the network. In the uniform *reversible* mode (not shown in the figure), the coalition loses and upon convergence all nodes vote

"No". Figure 1.1(C) still considers the *reversible* mode, but assumes the above coalition is granted with a influence factor of 2, and shows the final (stable) state of the voting, in which the coalition is also a monopoly. Note that further increasing the coalition influence factor will not change the stable state. To sum up the example, nodes 4 and 5 form a monopoly in the uniform, irreversible case, and also in the reversible case provided they are granted a influence factor of 2.

In this setting of on-line social networks, we study the following interesting questions:

1. Can a relatively small coalition, say of size sub-linear in the population size, be a monopoly? Does providing the coalition with higher influence factor have a significant impact? If so, should this factor be a function of the population size or would a constant increase suffice?

2. How does the monopoly size change over time (as a function of the population size) as society grows?

3. How significant is the effect of "playing fair"? Is there a significant difference in the required monopoly size between the reversible and irreversible modes?

In particular, we study the minimum k such that the *k-rich club* [54] (namely, the k highest degree nodes) in the network, acting as a coalition, can become a monopoly. Answering the questions above can help us understand the *power of the elite* in social networks, where by "power" we mean the level of control a unanimously voting elite has over the final results.

We focus our attention on elite size of $\sqrt{m}$ (where m is the number of edges in the network). Our previous and on going work ([47], [7]) indicate that $\sqrt{m}$ is an interesting scale to for monopolies. Linial et. al. [40] and Peleg [47] showed that $\Omega(\sqrt{m})$ is a lower bound for monopolies in a single round majority voting. In [7]) we showed, using axiomatic approach, that the elite must be of size $\Theta(\sqrt{m})$.

1.1 Overview of the Results

We studied the process of majority voting on several real social networks and models. We consider the basic case of *binary votes* and two coalitions. The first coalition consists of the members of the k-rich-club [54] of the network. The k-rich-club is defined as the sub-group of the nodes with the highest degree in the network. We used this group as it can simply approximate the elite of the network while we aware that these two groups are not necessarily the same. All other, common nodes or periphery nodes, form the second coalition. We study two modes of reversibility: In the first version, all nodes are reversible and we denote this version as *all-reversible*. In the second version, only the elite coalition is irreversible and we denote this mode *irreversible*. The outcome of majority voting were examined for different sizes of elites (i.e., k-rich-club) and different influence factors that were given to the elite. In particular, for a given size of a rich club and reversibility we found the minimum influence factor for it to become a monopoly and vise versa for a given elite influence factor and reversibility we found the minimum k for the elite to become a monopoly. Our main four empirical findings are:

1. The size of monopoly elites: A relatively small elite, particularly an elite of size $\sqrt{m}$, has a disproportional power toward the rest of the network and it can dramatically affect the results of majority voting. In other words, the elite has a monopoly over the final result of majority voting. Moreover, granting the elite with small influence factor is leading to a sharp decrease in the minimum elite size required to win the poll. In particular, for an elite of size of $\sqrt{m}$, the *maximal* observed value of all networks, of the minimum elite influence factor required to win the poll is 8 (Youtube). For the irreversible mode, the value is even smaller: 2. Note this result is obtained for a collection of networks of several sizes and densities, hence it seems that the factor is independent of the network size (in terms of number of nodes and edges). However, the power of elite is bounded, meaning that from some point, granting the elite with higher influence factor does not reduce the minimum elite size required to win the poll.

2. Reversible vs. irreversible: *Quantitatively* speaking, an elite operating in irreversible mode performs better compared to when operating in reversible mode. For example, a smaller elite is required in order to win the poll for the same elite influence factor, when operating in irreversible mode compared to when operating in reversible mode. However, asymptotically there is no *qualitative* difference between the behavior of the elite in the two modes.

3. The power of elites over time: An elite of size $\sqrt{m_t}$ (where m_t is the number of edges in the network at time t), keeps its monopoly power on the society along the life-time of the social network. We observed that, the small influence factor that an elite of size $\sqrt{m_t}$ needs to win the poll, is a constant independent of time and as more and more nodes are joining the network .

4. The power of elite of analytical models: Real networks elites are more powerful than those implied by the analytical models. For example, provided with some influence factor, $phi > 1$, a smaller elite can win the poll in real network compared to an elite of same size for matching[2] *Erdős–Rényi* and Preferential Attachment models[3]. Moreover, in these models, the influence factor an elite size of $\sqrt{m_t}$ needs be granted with to become a monopoly, grows asymptotically along the lifetime of a social network; Therefore there is qualitative difference between these models to real networks when it concern the power of elites in majority voting.

The rest of the thesis is structured as follows. The next section discuss the definition of the elite in sociological aspects. Chapter 2 discuss the background and related work including majority voting. Chapter 3 then describe our formal model and Chapter 4 present our empirical results. In Appendix A we discuss the diffusion process of local majority voting.

[2] Similar number of nodes and average degree.

[3] As expected, we also observe that Erdős-Rényi model generates weaker elite than Preferential attachment.

1.2 Who are The Elite?

As mentioned earlier our empirical study is done on the k-rich-club, but our goal is to understand the power of the elites. We find it important to discuss the differences and similarities between these concepts. The Cambridge dictionary defines the elite as "the richest, most powerful, best educated or best trained group in a society". Wikipedia adds that "an elite, in political and sociological theory, is a small group of people who control a *disproportionate* amount of wealth or political power". In his 1957 book, "The Power Elite", the American sociologist C. Wright Mills writes: "those political, economic, and military *circles*, which as an intricate set of *overlapping small but dominant groups* share decisions having at least national consequences...the power elite are those who decide them". The Italian sociologist Vilfredo Pareto puts it succinctly [45]: "Every people is governed by an elite, by a chosen element of the population". The political scientist Robert A. Dahl [16] distinguishes between "the potential" for control and "actual" control. Potential for control is when a set of individuals has the following property: there is a high probability that if they agree on an alternative, and if they all act in some specified way, then that alternative will be chosen. "Actual" control is the active participation of this set in decision making.

Who are the people constructing the elite? What is the size of the elite and how is it formed and structured in terms of (internal and outgoing) connections? In his book "Who's Running America?" [21], Thomas R. Dye defines the elite as those individuals who control or occupy formal positions of authority in top institutions over ten sectors: industrial (non-financial) corporations, banking, insurance, investments, mass media, law, education, foundations, civic and cultural organizations, and government. He identifies 7,314 institutional positions of power encompassing 5,778 individuals yielding an elite size of 0.0002%, or $n^{0.44}$ where n is the size of the population in the U.S. ($\sim$300 millions). Doob [19] found that most holders of top positions in the power elite possess exclusive membership in one or more social clubs. About a third belong to a small number of especially prestigious clubs in major cities like New York, Chicago, Boston, and Washington D.C. This leads to the conclusion that elite members are well connected to each other.

1.3 Additional Applications

The hypothetical social voting system, described in the beginning of this chapter, can also be considered as a voting mechanism for individual decision making on private matters, such as which mobile phone to buy or which university to go. When coming to take such decisions we often consult with people we trust. Trust networks can be thought of as a derivative of social networks. One usually does not trust all connections, some may be more trustworthy than others, so the opinions of our friends are weighted in some manner (and some may not be considered at all). Similarly, the endorsements of public figures for certain products may positively influence us, just as well-known scientists or politicians who graduated from some university may positively affect our decision to apply to it. However, the questions on which we consult our trust network can have more than two possible options and the recommendations we get from our friends might not be clear cut ("buy" / "don't buy") but include some level of uncertainty. Hence in order to use the above model for studying decision making among consumers, we need to expand its definition to include weighted votes and multiple (i.e. more than two) choices.

Another aspect of social networks that can be modeled by our voting system concerns diffusion processes. In his book, *Diffusion of Innovations*, Everett Rogers seeks to explain how, why, and at what rate new ideas and technology spread through cultures. He suggests that two of the four main elements that influence the spread of a new idea are the communication channels and the social system (the other two are the innovation and time). The innovation must be widely adopted in order to survive. During the adoption process, there is a point at which an innovation reaches "critical mass" – the point in time when the number of individual adopters ensures the continuation of the adoption of the innovation. An example for this process given in Rogers' book is that of the fax machine. It is only in 1987, 150 years after it was invented, that the fax machine reached its critical mass. It was the decreasing prices of the machines and cost per fax, combined with the usage of already in place phone lines for transmission, that brought the Americans “to assume that ’everybody else’ had a fax machine”. And naturally, when everybody else has a fax machine, one had better got himself one as well. Cast in our monopoly framework, it is likely that the same elite group that can form a monopoly over society can also affect the early creation of a critical mass, if not create it by its sole decision.

Chapter 2

Background and Related Work

This thesis studies the power of the elite coalition in local majority voting over social networks, both for static networks and over time. Let us next briefly introduce each of the main components of the system and their basic properties.

2.1 Social Networks

A social network is a structure composed of agents and ties between them, hence it can be represented by a graph consisting of a node for each agent and an edge for each tie. The ties may represent connections in the real world, such as friendships, relatives, colleagues, etc., or in a virtual space (i.e., the www). The network is represented by graph $G(V, E)$ where V and E are the sets of nodes and edges respectively. Typically each agent is a person, but it can also represent a more abstract entity, e.g., article, websites, and so on. The edges can be directed (e.g., with $A \rightarrow B$ representing the fact that A trusts B), or undirected (e.g., with $A \leftrightarrow B$ representing the fact that A and B are co-authors). The edge weights may represent level of trust, frequency of communication, influence, and so on. This broad definition encompasses many types of social networks. In several known online networks such as Facebook and Twitter, the edges are defined explicitly by the agents. Other network types, such as co-authorship, co-purchasing, and citation networks, are "alive" without the initiative of their members. In the past couple of decades, the extensive study of complex systems and social networks has yielded an impressive body of knowledge on such networks and the *universal* properties they share, namely, properties and behaviors that is repeatedly seen across different networks. We review the main characteristics of social networks in the following subsection.

Social Networks Properties

1. **Small world:** In the famous "small world experiment", Milgram [43] showed that a social network is very well connected, in the sense that there is always a short path between any two individuals belonging to it. The average distance found by Milgram, 5.2, is supported by recent experiments [37, 8] performed on contemporary social networks. For example, in 2008, researchers at Microsoft [37] studied records of 30 billion electronic conversations among 180 million people in various countries. They found that the average length of the communication path was 6.6 hops, and that about 80 percents of the pairs could be connected in seven steps or fewer. In 2011, a similar experiment examined 721 million Facebook users and revealed that the average number of links from one arbitrarily selected person to another was 4.74 [8].

2. **Giant component:** A necessary but insufficient condition for a small world network is that it must have a connected component which containing a constant and large fraction of the entire graph's nodes. For example, following Pareto law, the giant component will contain at least 80% of the nodes in the graph.

3. **Clustering:** A universal behavior is shared among all societies – people tend to cluster. The probability that two randomly selected individuals know each other is much higher if they happen to have a common friend (in graph theoretic terms, a social network graph is likely to have many triangles). In their joint paper [53], Watts and Strogatz proposed a model to capture this property and quantified it by a parameter they called *clustering coefficient*.

4. **Power law degree distribution:** By nature, there exist a "preferential attachment". New players in the network prefer to link to strong (i.e. with many connections) exiting exiting players, so over time we get few nodes with sky high degree and a long tail of nodes with small degree, and by that shape the degree distribution of the network to follow a *power law*: the probability of a node to have degree k is proportional to $p(k) = k^{-\alpha}$, $\alpha > 1$ [9]. For example, scientific collaboration networks are growing with the introduction of new scientists which are more likely to publish their first works in cooperation with *senior* scientists. Also, as prestige matters, notable scientists are preferred collaborators over their colleagues.

5. **Navigability:** The small world experiment revealed the existence of relatively short paths between every two people in a social network. However, it had also implied that people are able to find these short paths while each individual is operating with purely local information (as the structure of the whole network is unknown to the individual). This is due to the characteristic feature of small-world networks where their diameter is exponentially smaller than their size, being bounded by a polynomial in $\log n$. Klienberg [35] showed that it is only under unique properties there exist a decentralized algorithm that can achieve a delivery time bounded by any polynomial in $\log n$.

6. **Social networks over time: Densification and shrinking diameter:** In [38] Leskovec, Kleinberg and Faloutsos examined the evolution of real-world networks over long periods. They observed snapshots of some networks taken at regularly spaced points in time (e.g. days, months). They've found two major properties:

   (a) **Densification:** As the number of edges growing much faster (super linearly) than the number of nodes, the average degree of the graphs is increasing over time and we can say that the graphs are getting denser over time.

   (b) **Shrinking diameter:** The diameter of a graph is defined as the maximum length of undirected shortest path over all connected pairs of nodes. It appears that the diameter of real-world networks gets smaller as the network gets older. This phenomenon was observed also in a mature graph, when the giant component already contains nearly 90% of the nodes, so it is not that the decrease is due to many small disconnected components got connected to each other.

7. **Rich club:** The k-rich-club [54] is defined as the k nodes with the highest degree in the network. By the power law property we get that in social networks, this group is small and yet "holding" high potion of the links in the graph. This group is also very well connected: there is a very high probability that if two rich nodes are not directly connected, there exist another rich node connecting them, i.e. the diameter of the induced sub graph of the rich club is 2.

8. **Elite properties: Influence, stability and Density:** Concerning the universal properties of the elite, in a recent, yet unpublished report [7], Avin et. al. suggest an axiomatic approach to finding the size of the elite of social networks. Bases on three axioms for the elite: *Influence*, *Stability* and *density* they proved that the elite must be of size of $\sqrt{m}$ (where m is the number of edges in the network). Empirically, they find that the $\sqrt{m}$-rich-club has the following properties:

   (a) **Influence:** Thinking of an edge as representing a source of influence (a natural interpretation in social networks), they find that the elite has "disproportionate" power over the rest of the network: a significant constant fraction of the edges in the network are controlled by the elite.

   (b) **Stability:** The elite is robust against influence from the rest of the network, as the number of links between the elite members is also a significant fraction of the number of links connecting the elite to the outside.

   (c) **Density:** The elite, of size $\sqrt{m}$, is a dense subgraph.

Random Models

The common assumption in modeling social networks is that the network links are essentially "random" (in a sense to be made more precise next).

One common model is the *Erdős–Rényi* model. In this model, the network is constructed as a randomly chosen n-node graph from the class $G_{n,p}$, where for every two nodes x, y, a link xy exists in the graph with probability p, independently of the other links. For $n \cdot p = c > 1$ the model generates a graph with unique giant component with very high probability (no other component will be of size larger than $\log(n)$). We later use this model as a base case to compare our results to.

An alternative approach is to model the network by an evolutionary process describing its growth. A well-known representative of this approach is the *preferential attachment* model (PA) [4]. In this model, nodes join the network one by one, and each new node randomly selects some existing nodes according to their degrees and attaches itself to those nodes. The higher the degree of a node, the more likely it is to attract new nodes to connect to it. The network starts as an initial network of n_0 nodes. New nodes are added to the network one at a time. Each new node is connected to $d \leq m_0$ existing nodes with a probability that is proportional to the number of neighbors that the existing nodes already have. Formally, the probability p_i that the new node is connected to node i is [4] $p_i = deg(i)/\sum_j deg(j)$, where $deg(i)$ is the degree of node i. In this work we adopt the convention $m_0 = d$ and start with an initial network forming a complete graph (clique).

2.2 Majority Voting

There are many occasions where a group needs to express its will using one and simple parameter. For this it is using some *Voting system.* The voting system is using a preset of rules and mechanisms to synthesize the final decision. One of the ways of doing it, considering the subject is numeric, is averaging. For example, the number of sits a party gets in the parliament, and as a result the composition of the parliament, is expressing the average will of democratic societies. However, in some applications, averaging is unlikely or impossible. For example, consider a group of friends trying to choose the destination for their next summer trip abroad, or, a prize comity having to choose the best of all nominees. In this case, the group is usually using voting and the choice with most votes is chosen. In cases where there are more than two alternative to choose between, the system can be used to rank the alternative rather than choose a single one. A clear advantage of these three voting systems is that there are using clear and simple rules. But apparently, they also hide some pitfalls and pathologies. Choosing a single alternative, for example, can lead a frustration and disapproval of the results from the loosing side in case the numbers were razor tight. Producing a ranked list of alternative might lead to Condorcet paradox [22] in which the list can be cyclic (i.e. not transitive), even if the preference of each individual voter is not.

2.2.1 Local Majority Voting

With no intention to fix the pathologies of the above systems, we would now like to suggest a new voting system based on majority voting. We use the following allegory, suggested by Peleg [47], for the informal definition of the system: Tomorrow morning the members of a society is to vote – Yes or No – upon an important question. The decision will be taken by counting the votes and the party with most votes wins. We refer to the last, rather common, procedure as *global majority voting*. Now, let the society be modeled by a graph. Each node of the graph represent a citizen and there is an edges from each citizen to his acquaintances - friends, colleagues, neighbours, etc. Prior to the global majority voting, each of the citizens is polling his acquaintances (i.e. his neighbors in the graph) for their vote, just to get a sense of the outcome. In case he is "seeing" a majority of his acquaintances voting the opposite from him, he backs down from his original vote and decide to vote with the majority. Otherwise, he stays with his original vote. We refer to this poll as *local majority voting*.

One can see that local majority voting is in fact a localized algorithm where information is exchanged over the edges. Thus, controlling the information exchanged over the edges will probably lead to control over the results. Recalling the properties of the elite, where there is a high portion of edges are going in, out and between elite members, it looks like the elite has the potential to control the result of majority voting. An interesting majority vote game will be to match a coalition of the "elite" vs. coalition of all other nodes. In the following section we discuss definitions of the elite as small, well connected and ruling group of the society from more sociological aspects.

Peleg [47], is overviewing the process of local majority voting, its variants, and properties. The variants are actually rules. Each rule holds for all vertices, there is no "free will" for a vertex whether to follow a rule or not. Apparently, each variant can affect the results dramatically. Following two subsections we discuss the variants of and properties of local majority voting.

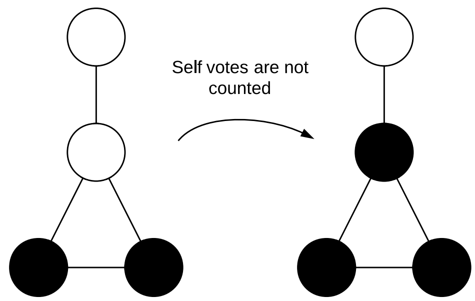
2.2.2 Local Majority Voting Variants

Self Votes Counting Mode

Polling for neighbors' votes, the nodes can take their own vote in account or not. These are denoted by SI (Self-Included) and SN (Self-Not-included). Figure 2.1 demonstrate the affect of self votes counting on the outcome of the polling game. When self votes are not counted, the vertex in the middle is seeing a majority of two "No"s vs. one "Yes" leading him to flip his vote. However, if self votes are counted, he is seeing a tie. The last is bringing us to the next variant – how ties are broken.

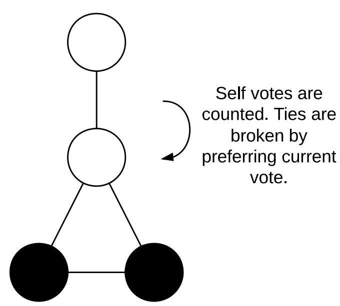
Tie Breaking Mode

As we saw, counting the votes may end in a tie. We have three options to break the tie: we can prefer one of the votes over another (w.l.o.g. "Yes"), we can prefer to flip our current vote or to stay with it. These are denoted by PY (Prefer-Yes), PC (Prefer-Current) and PF (Prefer-Flip). Figure 2.2 demonstrates the result for the same setup as in figure 2.1 for SI and PC modes. The result is that the middle vertex do not change his mind.



.

Figure 2.1: Local majority self vote example. Not counting his self vote, the middle voter will change his vote to "No".

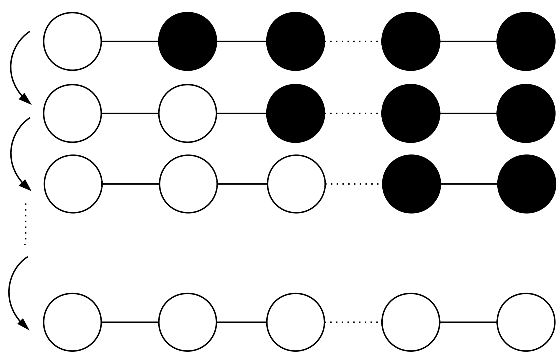


.

Figure 2.2: Local majority tie breaking example. Self votes are counted. The middle voter breaks the tie by preferring his current vote.

Repetitive Voting Game

Let us define a voting system which goes as follow: In the Initialization phase each nodes chooses his own, subjective vote. In the Update phase, the vertices are conducting a local majority voting game. In the last phase, the vertices are conducting a global majority voting and the party with most votes wins. Repeating the game in the update phase for more than single round can dramatically affect the results of the final phase. Figure 2.3 shows the result of the game on the Line for PF and SN modes. The second node from the left breaks the tie by flipping his vote. One can see, that, given the game continue for another round, another node will flip his vote, and so on. This brings us to the next variant of the game. Given the update phase contains single round the "No" voters wins. But, as one vertex is changing his vote to "Yes" in each round, it is easy to see that from some point forward, the "Yes" voters will win. The repetitive voting variant can represent models of social contagion – flow of information, diffusion of innovation and spreading of diseases. Note we only consider the case in which the polling is done in discrete time steps. The state of the system in time t is represented by a vector $X_t = (x_t(v_1), x_t(v_2), ..., x_t(v_n))$ where $x_t(v_i)$ represents the value of node v_i after time (i.e. round) t.

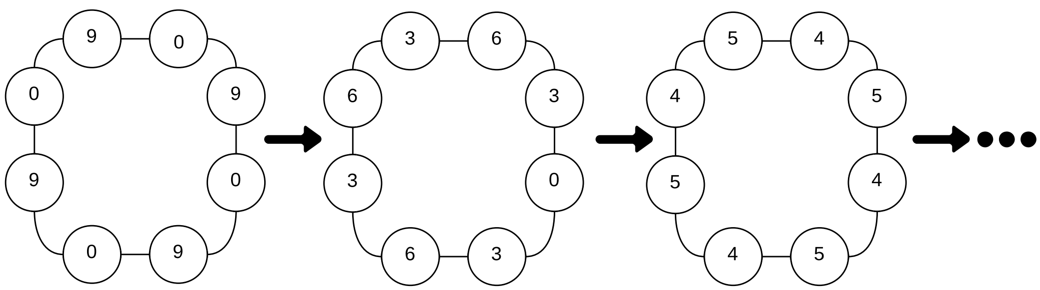


.

Figure 2.3: Repetitive local voting procedure on the Line for the (SN, PF) model. In every round, another "No" voter is changing his vote to "Yes", ending in all "Yes" consensus.

Votes Space, Counting Mode and Edge Weights

We can look on a version of the game where more than two votes are allowed. For example, votes can be on a discrete ladder from 1 to 5 or having some real value between the bounds. In this case, the more common application is that, counting the votes in each round, the vertices are *averaging* the votes rather than taking the majority. More than that, we can decide to assign extra weight to the vote of some nodes and in this case the averaging is done in a weighted manner. Figure 2.4 [47], demonstrating a repetitive averaging process on the Ring.



.

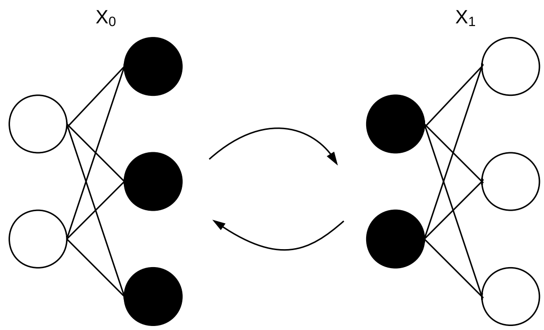
Figure 2.4: Repetitive averaging on the Ring. Picture taken from [47].

2.2.3 Local Majority Voting Properties

Periodicity Behavior

The period of repetitive local voting process is defined as the minimum number of rounds for its state to repeat itself, from some time t^* forward. Formally, $Period(X_t) = \min\{k | X_{t^*+k} = X_{t^*} \text{ for some } t^*\}$. We say that the process reaching a fixed point (i.e. steady state) if $Period(X_t) = 1$. Since for finite graphs the number of states is also finite, it is clearly periodic. Actually, we can say more than that: for a finite graph, the period is 1 or 2

[29][49][50]. Figure 2.3 is demonstrating a system with period of 1 while Figure 2.5 [47] is demonstrating a system with period of 2.



.

Figure 2.5: Repetitive process with period of 2. Picture taken from [47].

Control, Coalitions and Monopolies

The questions we raised at the Introduction are actually questions of control: Can a group of nodes voting unanimously affect the outcome of the local poll carried by one or more vertices? Let N be the set of nodes in the graph G. We define a subset of nodes voting unanimously as coalition:

Definition 1 (Coalition) *Coalition is a group of nodes* $M \subseteq N$ *such that* $X_0(v) = C$ *for all* $v \in M$.

Note that our definition for coalition is weaker and less generic than this defined by in [47]: A set of vertices $M \subseteq N$ which is able to *determine the outcome* of the local poll carried by v. We say that a coalition has a monopoly if, by coordinating their votes, they can "convince" (i.e. force) more than 50% of the nodes to vote the same as them. Note, that it may be a case where a coalition of consist of more than 50% of the nodes in the graph may not be enough for a monopoly.

Definition 2 (Monopoly) *Monopoly is a coalition which, given a local poll carried out by all vertices, can force the result of the poll such that* $\sum_{v \in N | x(v)=C} > \frac{n}{2}$.

Note that our definition of monopoly is also weaker than the one given in [47]. There, a monopoly is defined as a coalition which can determine the votes of *all* the nodes in the graph. In addition, there defined *dynamic monopoly*, abbreviated *dynamo*, a coalition which can reach a monopoly only in repetitive polling process (i.e. it need more than a single round to force his vote). Figure 2.6 [47] is demonstrating a coalition of $2\sqrt{n}$ forcing a 2:1 majority in all polls. The coalition is consisting of a core of nodes forming a clique. All other nodes are forming the periphery. The periphery is consisting of disjoint groups, each of size $\sqrt{n}$, where each node in the group is connected to the same set two nodes in the core. The set of two nodes is unique per group and there are no internal edges between nodes in the periphery. It is interesting that the core in this example, is also the core by means of the k-rich club as defined by Mislove et al. in [44]. An important lesson we can learn from it is that the influence of vertex depends on his connections to other vertices, the more it is connected to other major influencing vertices, the more it is ability to influence the outcome.

Reversibility

The repetitive model has also been studied for its *irreversible* variant. In this variant, w.l.o.g. the "Yes" voters are "not allowed" to reverse their vote back to "No". In the context of social voting, it can model a situations where some voters are considered as "prejudiced". Given the system in figure 2.5 worked in irreversible mode, it would have had a period of 1 instead of 2 thus avoiding the toggling behavior.

In the following section we continue with the formal definition of the voting system and its variants. However, at the introduction of this section, we suggested that the elite of the society, when defined as the k-rich-club has major impact on the outcome of a poll carried by local majority voting process. In the following subsection we discuss this notion.

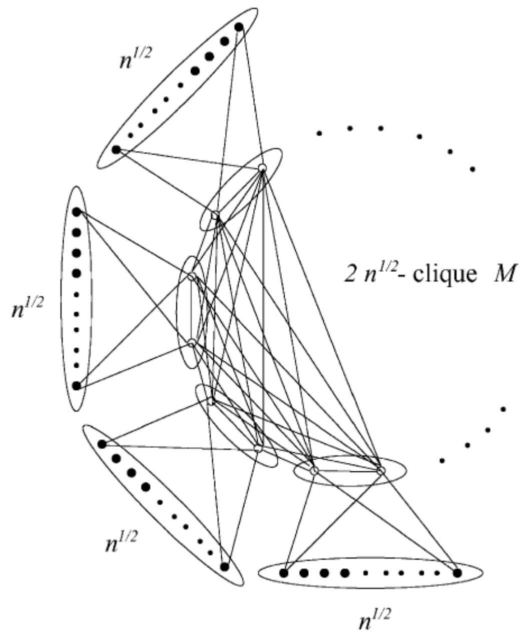


Figure 2.6: $2\sqrt{n}$ coalition forcing a 2:1 majority. Picture taken from [47].

2.2.4 Related Work

Discrete influence systems, including ones based on voting processes, occur naturally in diverse settings, and were studied, for example, in the context of social influence [27, 17] and neural networks [28, 48]. Our voting process is also reminiscent of *epidemic* models (cf. [6] and the references therein).

A dynamic, synchronous, deterministic majority-based version of the voting process, similar to the one considered here, was studied in [48, 46, 28]. In that context, the term *monopoly* is reserved to a particularly powerful set of nodes that, if acting as a coalition, can force the process to end in consensus on its opinion after a *single* voting step; the repetitive variant considered here is referred to therein as a *dynamic monopoly.* Reversible and irreversible dynamic monopolies were studied in [46, 24, 25, 26, 13, 20]. Related problems were studied in different guises and under different names, in [33, 23, 11, 3, 51, 12, 31, 15].

Probabilistic variants of the voting process were extensively studied as well. A repetitive (synchronous) *probabilistic* voting process on weighted graphs was studied in [32]. The nodes start with their initial opinion as their vote. In each round, each node recomputes its vote by choosing at random one neighbor (with probability dependent on the edge weights) and adopting its vote. It is shown that the probability of the process ending with every node holding a vote of "yes" is proportional to the sum of the probabilities of the nodes whose initial opinion is "yes" in the stationary distribution of the process. This theorem is generalized also to processes on graphs with multiple (i.e., non-binary) opinions.

Related results concern a probabilistic *asynchronous* model called the *voter model*, introduced by [34]. This is a continuous time Markov process with a state space consisting of all possible vote distributions. This process was extensively studied in infinite grid graphs [34] and finite connected graphs [18], analyzing the convergence time and convergence probability.

Chapter 3

The Model

The network is modeled by a simple[1] directed weighted graph $G(V, E, \Phi)$ where V is the set of nodes, E is the set of edges and Φ a weight function. Let n and m denote the number of node and edges respectively. For every two nodes, either both or none of the two directed edges appear in the graph. (The advantage of this formalism over using undirected edges is that it enables assigning different edge weights, modeling different influences of the two neighbors on each other.) Let $N(v)$ denote the set of neighbors of v in G, *including* v (i.e., $N(v) = \{u \mid uv \in E\} \cup \{v\}$).

In the context of a voting system, the nodes of the graph represent voters and the edges represent influence. A directed edge uv with weight $w_{uv} > 0$ from node u to node v represents the fact that voter u influences voter v proportionally to w_{uv}. Different votes are modeled as colors. At any time, each voter is colored with its current vote. The simplest case concerns binary votes, so each voter is colored with either white (i.e., "yes") or black ("no"). Formally, the value of a "Yes" (respectively, "no") vote is set to $+1$ (resp., -1). Neutral votes are disallowed. A *coalition* is a set of nodes that coordinates the same initial vote. A coalition is called a *monopoly* if at the end of the voting process it manages to convince more than 50% of the population to vote according to its own initial vote.

In the *initialization* phase, each node sets its vote to its initial (subjective) opinion.

[1] with no parallel or self edges

Next starts the *update* phase, a synchronous process consisting of several *rounds*. In each round, each node updates its current vote to the majority of the votes in its neighborhood, including its own vote, and considering the weight of its neighbors. In case of a tie, the node keeps its current vote. A bit more formally, let $X_t = \{x_t(v_1), x_t(v_2), \ldots x_t(v_n)\}$ be the global vector state after round t, where $x_t(v_1)$ represents the value of node $v_i \in V$ after round t. Let X_0 be the vector of initial votes. For $t > 0$ let $S_t(v) = \sum_{u \in N(v)} w_{uv} \cdot x_{t-1}(u)$. Then

$$x_{t+1}(v) = \begin{cases} -1, & S_t(v) < 0, \\ 1, & S_t(v) > 0, \\ x_t(v_i), & S_t(v) = 0. \end{cases} \tag{3.1}$$

For example, consider a node v whose current vote is $+1$ and suppose v has 3 neighbors with weights {1, 3, 7} and votes {1, 1, -1} respectively. The updated vote for v in this case is $\text{sign}(1 + 1 \cdot 1 + 3 \cdot 1 + 7 \cdot (-1)) = -1$. The update phase continues until convergence is reached (if at all[2]). The final result of the voting process is obtained by counting uniformly the final votes of all nodes. The value voted for by the majority wins the poll.

In the above definitions the coalition plays fairly (i.e., with the same rules as the rest of the player); this is referred to as the *reversible* mode. In the *irreversible* version, the coalition members play unfairly, in the sense that in the *update* phase they make one-sided moves: they "answer" when being polled for their vote but they keep their original values without executing the update rule specified in Eq. (3.1).

[2]The period of the voting process is defined as $Period(X_t) = \min\{k | X_{t^*+k} = X_{t^*} \text{ for some } t^*\}$. Then for finite graphs, the period can be 1 or 2, cf. [47]. Here we only consider cases where the period is 1 (steady state).

Chapter 4

Empirical Results

4.1 Data Sets and Models

The WWW serves as a habitat for many social networks, whose users publicly post their view and communicate directly. We have used data on social networks from several sources. The Stanford Large Network Dataset Collection (SNAP) [1] was used for the graphs of Slashdot, Twitter, Brightkite, Gowalla and Citation (cit-HepPh). We used the collection[1] of Mislove et al. [42, 44, 52] for the graphs of Facebook, Youtube and Flickr. Finally, we have used The Koblenz Network Collection (KONECT) [2] for the graphs of Email, Epinions, Wikiusers and Slashdot2.

Table 4.1 lists the networks we have used and some of their basic properties. Several of these networks were used for both static and dynamic (i.e. over time) analysis. Below is a short explanation about the nature of each network and the way it was obtained. In all cases, where the original data set was directed or signed, and/or contained multiple edges, we have converted it to be undirected and simple by adding opposite edges where missing and removing parallel edges.

Youtube: Youtube is a video-sharing site that includes a social network. The graph consists of links based on a who-follows-who relationship. The network was obtained on

[1] available at `http://socialnetworks.mpi-sws.org`

network name	n (WCC%)	m	avg. degree.	effective diameter	$\sqrt{m}$ (% of nodes)	influence factor rever.	influence factor irrev.	duration (months)
Youtube [44]	1138499 (99.7)	2990443	5.25	6.31	1729 (0.15)	6.79	1.98	-
Slashdot [39]	82168 (100)	543381	13.23	4.70	737 (0.90)	3.64	1.48	-
Twitter [42]	81306 (100)	1242397	61.82	5.60	1115 (2.77)	3.42	0.98	-
Flickr [44]	197042	15555040	157.98	5.24	3944 (2.00)	3.32	1.20	-
Facebook [52]	63731 (99.4)	817090	25.64	5.59	904 (1.42)	6.01	2.67	29
Gowalla [14]	196591 (100)	950327	7.35	5.68	975 (0.50)	3.67	1.98	20
Brightkite [14]	58228 (97.4)	214078	7.35	5.90	463 (0.80)	4.48	1.98	30
Citation [38]	34546 (99.6)	420899	24.37	5.00	649 (1.88)	7.64	1.67	124
Slashdot2 [30]	51083 (100)	116975	4.58	5.23	342 (0.67)	-	-	12
Email [36]	87273 (96.7)	298338	6.83	5.87	546 (0.09)	-	-	73
Epinions [41]	131828 (90.4)	711496	10.79	5.49	844 (0.64)	-	-	32
Wikiusers [10]	118100 (95.8)	2036794	34.49	3.81	1427 (1.21)	-	-	8

Table 4.1: Examined networks and their basic properties

January 2007.

Slashdot: Slashdot is a technology-related news website. In 2002 Slashdot introduced the Slashdot Zoo feature, which allows users to tag each other as friends or foes. The network contains friend/foe links between the users of Slashdot. The static network was obtained in February 2009. The network used for the dynamic analysis (Slashdot2) was obtained over the period of Aug. 2005 - Aug. 2006.

Twitter: Twitter is a well-known online microblogging service. A directed edge from node i to node j represents the fact that user i is “following” user j.

Flickr: Flickr is a photo-sharing site based on a social network. The graph consists of links based on a who-follows-who relationship. The network was obtained on January 2007.

Facebook: Facebook is the largest social network in the Internet. Its graph consists of a list of all of the user-to-user links from the Facebook New Orleans networks crawled between Sep. 2006 and Jan. 2009.

Gowalla: Gowalla is a location-based social networking website where users share their locations by checking in. The friendship network is undirected and based on the check-ins of these users over the period of Feb. 2009 - Oct. 2010. For the analysis of this network over time, we considered the first check-in time as the user's joining time to the network. As we've found that about 50% of Gowalla users never checked in, they were not taken in account for the over time analysis.

Brightkite: Same as Gowalla, except that almost 100% of the users did checke in. The data was obtained over the period of Apr. 2008 - Oct. 2010.

Citations: Arxiv's HEP-PH (high energy physics phenomenology) citation graph is taken from the e-print arXiv and covers all the citations within its data-set. If a paper i cites paper j, then the graph contains a directed edge from i to j (papers cited by a paper out of the data-set or citing such a paper are not contained in the graph). The data covers papers in the period from January 1993 to April 2003.

Email: The Enron email network consists of emails sent between employees of Enron between Jan. 1998 and Feb. 2004. The nodes in the network represent individual employees and the edges correspond to individual emails. The data was obtained over the period of Aug. 2001 - Mar. 2009.

Epinions: Epinions is a consumer review website where the members can decide whether to "trust" each other. The network is based on a who-trusts-who relationship. The data was obtained over the period of Jan. 2001 - Aug. 2003.

Wikiusers: This network indicates positive and negative conflicts between users of Wikipedia, for example, users involved in an edit-war. A node represents a user. An edge represents a positive or negative conflict between two users.

The SNAP Collection was also used for its software package for the creation of *Erdős–Rényi* and the Preferential Attachment models graphs. The first parameter for the generation of *Erdős–Rényi* and Preferential Attachment graphs is the number of nodes. The second parameter for *Erdős–Rényi* is the number of edges and for Preferential Attachment is the average degree, so for a given real network both models generate a random graphs with the same number of nodes and the same average degree as the real network.

4.2 Experimental Setup

We examined the results of majority voting on social networks. We looked at the basic case of binary votes and two coalitions, one formed by the elite (considered as the k-rich club) and the other formed by all other (common) nodes. The parameters of the setting considered are the elite size, the influence factor granted to the elite, and the reversibility of the elite. In each simulation we considered either the reversible or the irreversible mode. Then, one of the parameters – influence-factor or size – was fixed, while different values were sampled for the second parameter. By using binary search, we determined the minimum value for the free parameter that guarantees a majority for the elite, making the elite coalition a monopoly. In the following sections we discuss our findings.

4.3 The Power of the Elite

Figure 4.1 presents the basic results for the minimum elite sizes and influence factors that are needed for the elite coalition to become a monopoly in the *all-reversible* mode for all real social networks studied. The results should be compared with Figure 4.2, which demonstrates a few possible outcomes of majority voting for different types of general networks. In both figures the X-axis shows the influence factor granted to the elite (in log scale) while the Y-axis shows the value of the exponent y for which $n^y = k$ is the size of the minimum k-rich-club coalition that yields a monopoly. For example the point (4,0.6) indicate that if we would grant the elite a influence factor of 4 then the $n^{0.6}$-rich-club is the smallest rich-club coalition that become a monopoly. Each line in the figures represents a different network. The function y is in general monotonically non-increasing, since granting the elite additional influence can only decrease the minimum elite size required to win.

In Figure 4.1, the colored circles mark the special case where $k = \sqrt{m}$ and the mini-

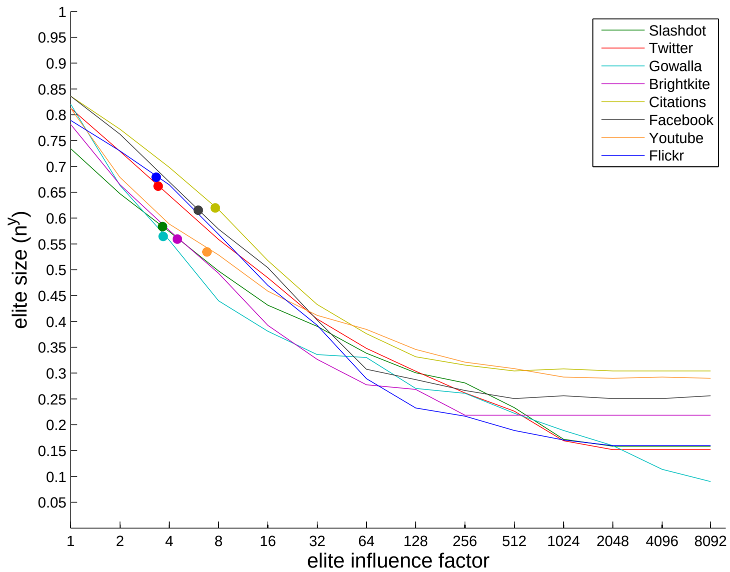
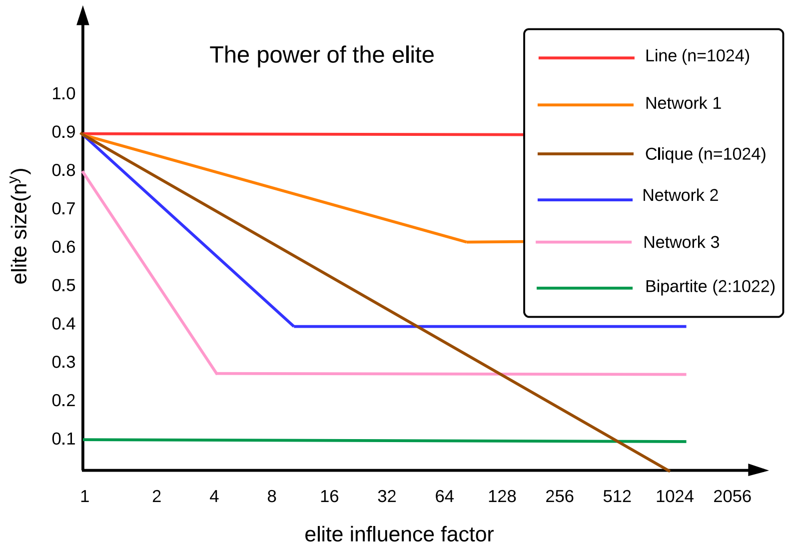


Figure 4.1: The power of elite. Relatively small elite coalition win the majority voting. An elite of size $\sqrt{m}$ requires a maximum influence factor of 7.64 (and 4.78 on average) to win the local majority voting.

mum influence factor required by the $\sqrt{m}$-rich-club to be a monopoly. We observe similar characteristics for all real social networks depicted in the figure: (i) a relatively small influence factor granted to the elite leads to a large decrease in the elite monopoly size, (ii) granting an elite of size $\sqrt{m}$ a relatively small influence factor (between 3.32 and 7.68; see reversible column in Table 4.1 for the full list) makes it a monopoly, and (iii) from some point on, granting the elite coalition higher influence factor does not further reduce the minimum elite size required to win the vote.

These three characteristics point out that the elite has disproportionate power w.r.t. the rest of the society in the majority voting game. Looking at Table 4.1, one can see that the result are obtained for a variety of networks with distinct characteristics (i.e., number of nodes, number of edges, etc.). The results are particularly interesting for an elite of size $\sqrt{m}$, since previous studies [54] considered k-rich-clubs of linear number of nodes [5].

Such disproportionate power of a $\sqrt{m}$ elite (or rich club) may not be the case for



.

Figure 4.2: How to read the results

arbitrary networks. Figure 4.2 demonstrates several possible outcomes of majority voting on networks of size 1024, from ones with the weakest possible elite structure (red line – topmost) to those with the strongest (green line – lowest). Networks with the weakest elite structure require a large elite to win the poll, and granting the elite additional influence does not reduce the minimum elite size (i.e., the elite is saturated with power). A representative example for such a network and elite can be the *line graph* (or path), where the elite resides sequentially on the path[2] (i.e., the elite is a sub-path within the network). The elite of Network 1 (orange line) is stronger: although it requires the same elite size as the red network when the influence factor is 1, granting it extra influence decreases the minimum elite size, up to some point where adding more influence does not empower the elite anymore. An interesting example is the clique graph (brown line): when its elite is granted no extra influence, half of the nodes are required to form a monopoly (in this case $1024^{0.9} = n/2$). However, granting the elite additional influence reduces the required elite size linearly, until

[2]Since all the nodes have the same degree, we are free to select the worst case elite.

a single node granted with an influence factor of 1024 can monopolize the entire network. The blue line represents Network 2 where the elite is stronger, since the slope is sharper than this of the orange network (although it becomes saturated for a smaller elite size). In Network 3 (the pink line), the elite is stronger than in the blue one, because it is always below it (meaning, requires smaller elite size for any influence factor). The elite of the network represented by the green line is the strongest – it requires the smallest elite size to win, already with influence factor 1, although it becomes saturated from that point forward. A representative example for such elite coalitions can be a complete $(2, 1022)$-bipartite graph having 2 nodes at one side (the elite) and 1022 at the other *when the elite is irreversible*. Intuitively we may say that the smaller the area below the graph, the stronger is the elite illustrated by it.

4.4 Reversible vs. Irreversible Mode

In this section we address the influence of reversibility on majority voting. Figure 4.3 illustrates the impact of reversibility on the outcome of majority voting for all networks. Solid (respectively, dashed) lines represent the reversible (resp., irreversible) mode. Recall that in the irreversible mode, the elite ignores its neighbors and always maintains its initial votes. The circles markers have the same meaning as in Figure 4.1. We observe similar characteristics for all networks: (i) the slopes in both modes are similar, (ii) for small influence factors granted to the elite, the irreversible slope is below that in the reversible mode, but from some influence and forward, both are asymptotically bounded. (iii) an elite of size $\sqrt{m}$ requires smaller influence factor to win in the irreversible mode (between 1 and) than in the reversible mode (between 3.32 and 7.68). See Table 4.1 for the full list of influence factors. These three characteristics point out that the differences between the two modes are quantitative rather than qualitative.

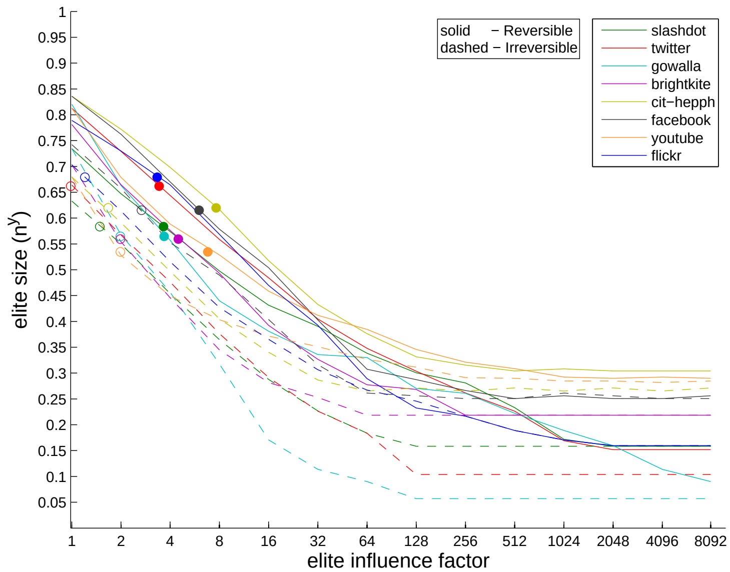
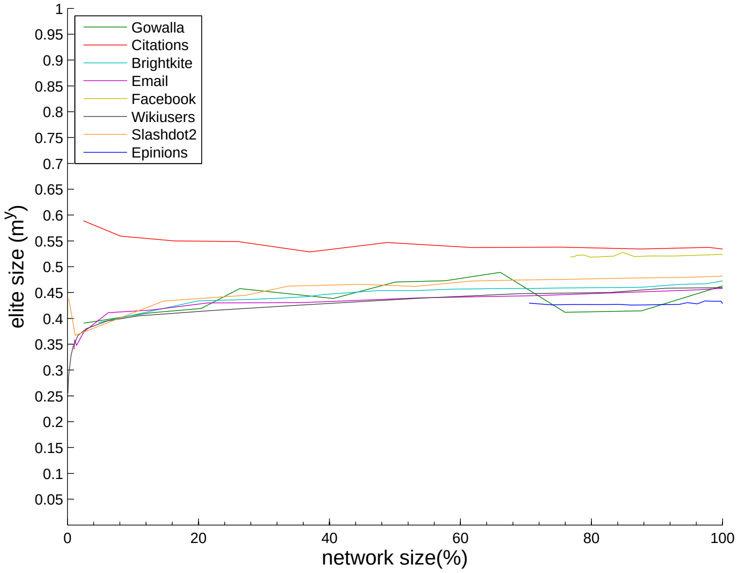


Figure 4.3: The power of elite: reversible vs. irreversible. The differences between reversible and irreversible modes are quantitive rather than qualitative.

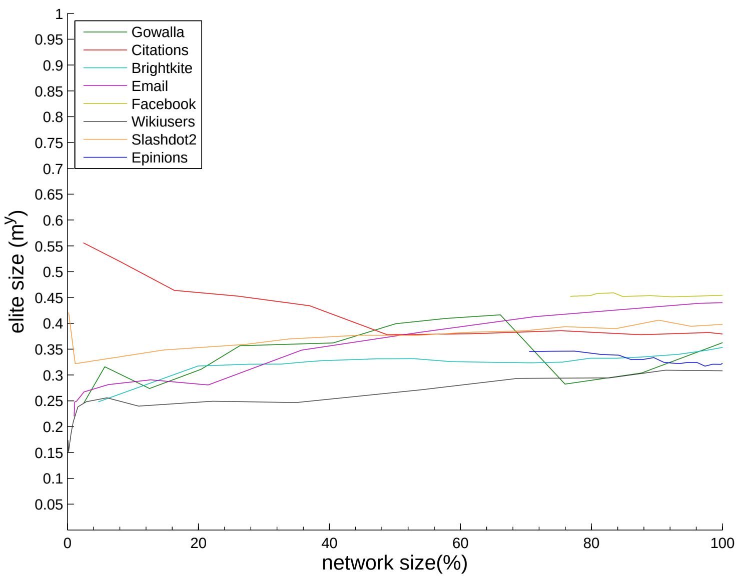
4.5 The Power of the Elite Over Time

In the previous sections we have discussed the power of elite as reflected in static snapshots of the networks under study. In this section we address the power of the elite along the life-time of the social network. For this purpose we took multiple snapshots of the studied networks in several points in time from the start (i.e., the network creation time) to the end (i.e., the current time). Each network snapshot contained only nodes that have already joined the network by the time that snapshot was taken.

For each graph we performed two types of measurements. For the first, we granted the elite a constant influence factor throughout the network's lifetime. Then, at each point, we used binary search to find the minimum elite size required to win the vote. We observed a similar behavior for all networks: when granting the elite with a constant, unchanged, influence factor throughout its life-time, the size of the elite monopoly appears to converge. Figure 4.4 shows the result for all networks when the elite is granted influence factor 4 for



(a) Reversible



(b) Irreversible

Figure 4.4: The power of the elite over time. For a constant influence factor ($\phi = 4$), the influence factor deriving the minimum sub-linear elite required to win the majority voting is asymptotically constant.

the reversible (a) and irreversible (b) modes. The X-axis shows the evolution of the number of nodes in the network (as a percentage of the curent size). Note that the first snapshot of some networks already contains a large fraction of the total number of nodes. The Y-axis shows the value of the factor y where $m^y = k$. We observe that different networks converge to different values between 0.45 and 0.55, i.e., the required elite size is in the order of $\sqrt{m}$.

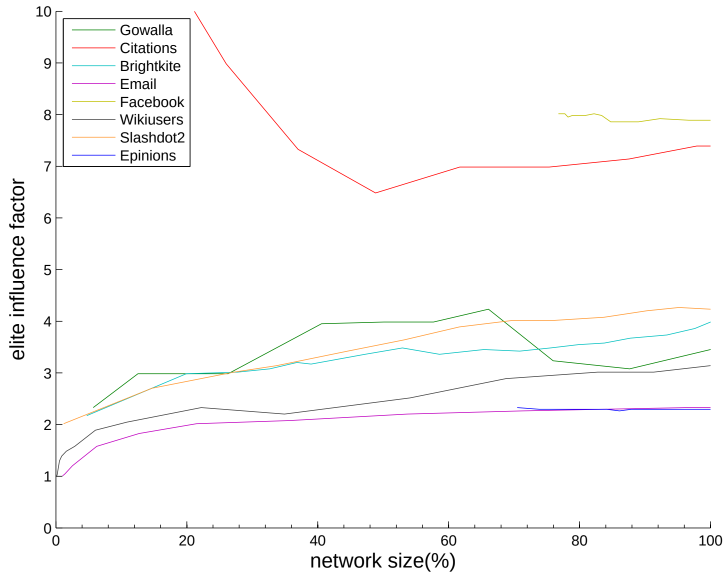
To verify this picture, in the second experiment we took a snapshot of an elite coalition consisting of the $\sqrt{m}$-rich-club at several points of time along the life-time of the network. Then, at each point, we used binary search to find the minimum influence factor required for the elite to become a monopoly. Figure 4.5 shows the result for all networks for the reversible (a) and irreversible (b) modes. We observe similar characteristics for all networks: excluding a short initialization period at the start of the network's life-time, the (small) influence factor an elite of size $\sqrt{m}$ must be granted in order to win the poll converges to a constant with time, along the life-time of the social network.

These two observations indicate that an elite of size $\sqrt{m}$ keeps its control over the rest of the network nodes throughout the life-time of the social network. It monopolizes the outcome of majority voting in every time-step along the way. These results support the claim that $\sqrt{m}$ is the right scaling that should be considered for the size of the elites that "controls" the network.

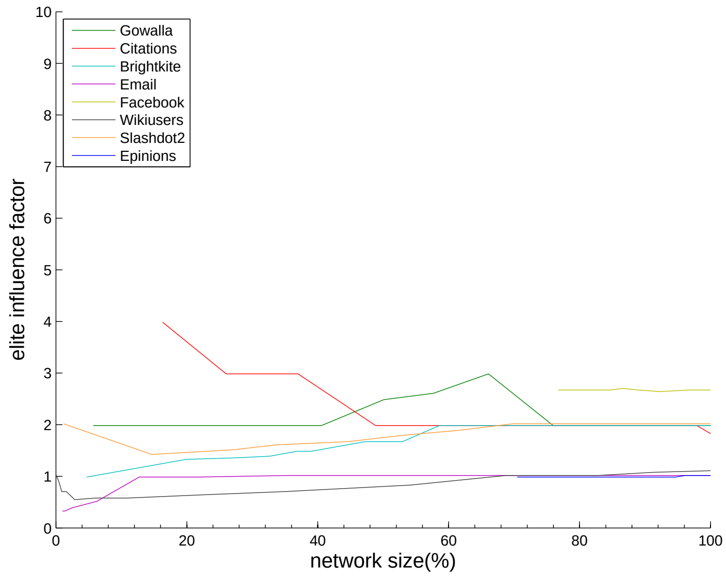
4.6 Power of the Elite in Random Models

We have compared the performance of the elite of the *Erdős–Rényi* (ER) and Preferential Attachment (PA) models when challenged with the same simulations done for the elite of real world networks. We generated a matching model graph for each of the networks shown in Table 4.1. By "matching" we mean it has the same number of nodes and average degree[3].

[3]For the Preferential Attachment model, the average degree is rounded to the nearest natural number for obvious reasons.



(a) Reversible



(b) Irreversible

Figure 4.5: The power of an elite of size $\sqrt{m}$ over time. The (small) influence factor required for such an elite to win the majority voting converges to a constant with time.

The Power of the Elite

We observe similar characteristics for all networks: (i) Both models reflect weaker elites than their corresponding real networks, where the *Erdős–Rényi* elite is the weakest of the three. (ii) Particularly, an *Erdős–Rényi* elite of size $\sqrt{m}$ is unable to win even when granted an asymptotically large influence factor. This may be expected recalling that this model has a binomial degree distribution, unlike the power-law distribution characterizing other models. (iii) In the Preferential Attachment model, the elite is weaker than in real networks, mostly when a small influence factor is granted to the elite. When large influence factors are used, the difference is milder and for some networks, the Preferential Attachment elite is even stronger than that of the corresponding real network. Figure 4.6 shows the difference between the three for all networks.

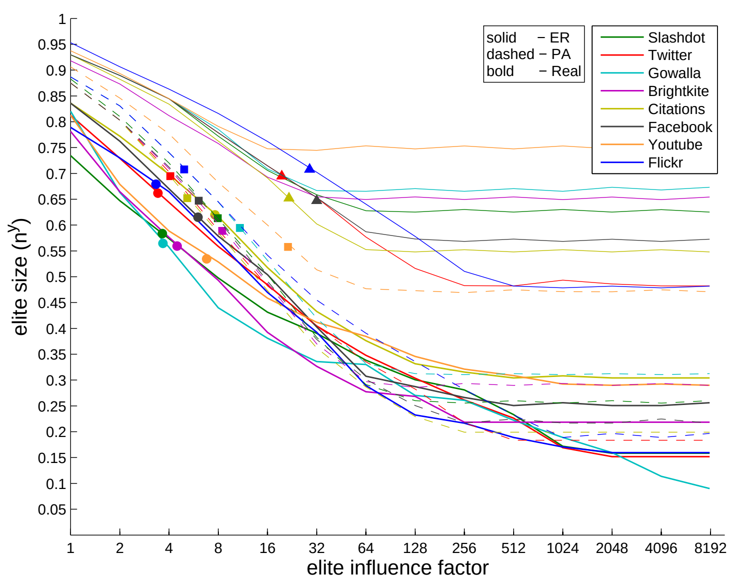


Figure 4.6: The power of elite of real networks vs. theoretical models. Real networks are demonstrating stronger elites than Preferential Attachment for small influence factor. *Erdős–Rényi* model elite fail to win though granted with asymptotically large influence factor.

Reversible vs. Irreversible

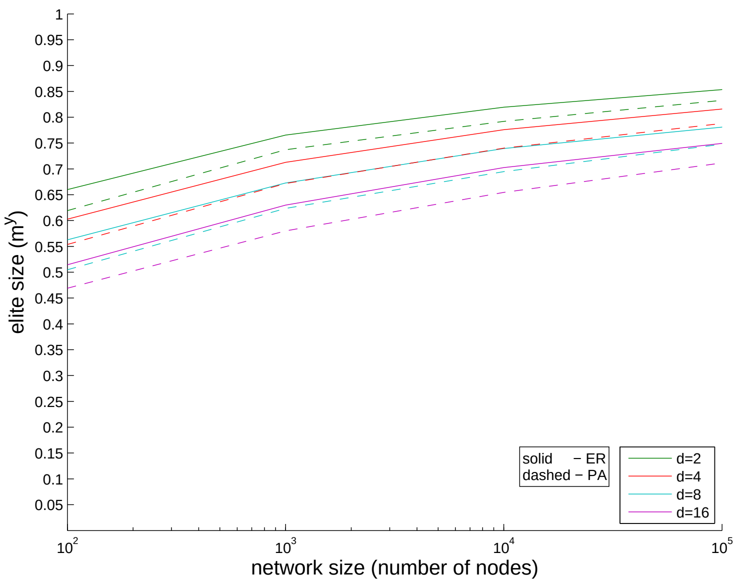
Similar to the behavior of real networks, both models do not present qualitative differences between the reversible and irreversible modes. In addition, both models demonstrate small, sometimes negligible, quantitative differences between the two modes.

Power of the Elite Over Time

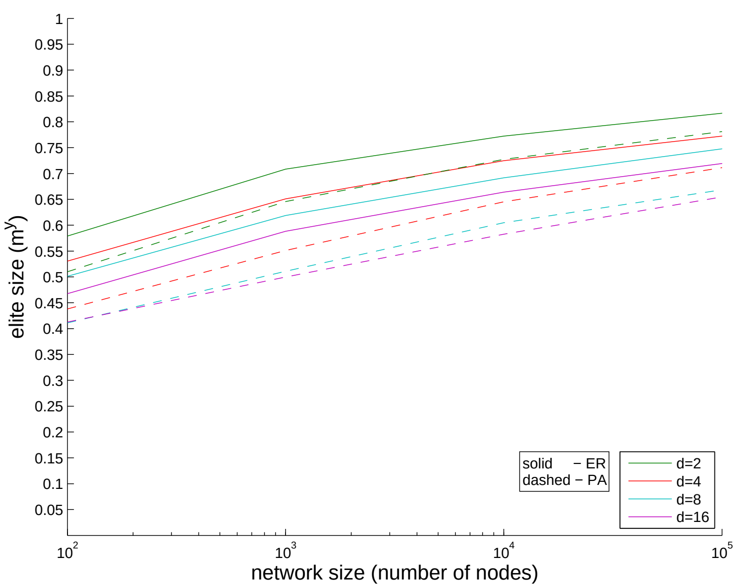
We simulated the network growth for both the *Erdős–Rényi* and Preferential Attachment model graphs, and repeated the two simulations from section 4.5 for each.

For the first simulation, using constant influence factor and searching for the minimum elite size required to win the vote, Figure 4.7 shows the result for both graphs when the elite is granted influence factor 4 for the reversible (a) and irreversible (b) modes. The X-axis shows the graph size (in log scale). The Y-axis shows the value of the exponent y where $m^y = k$. One can see that the required minimum elite size converges for both models but the convergence is slower (and for larger scale) compared to real networks.

Similarly, for the second simulation, using an elite of size $\sqrt{m}$ and searching for the minimum influence factor required to win the vote, Figure 4.8 shows the result for both models for the reversible (a) and irreversible (b) modes. The X-axis shows the graph size (in log scale). The Y-axis shows the influence factor granted to the elite. One can see that the minimum influence factor grows super-linearly with time, unlike the observed growth for real world networks.

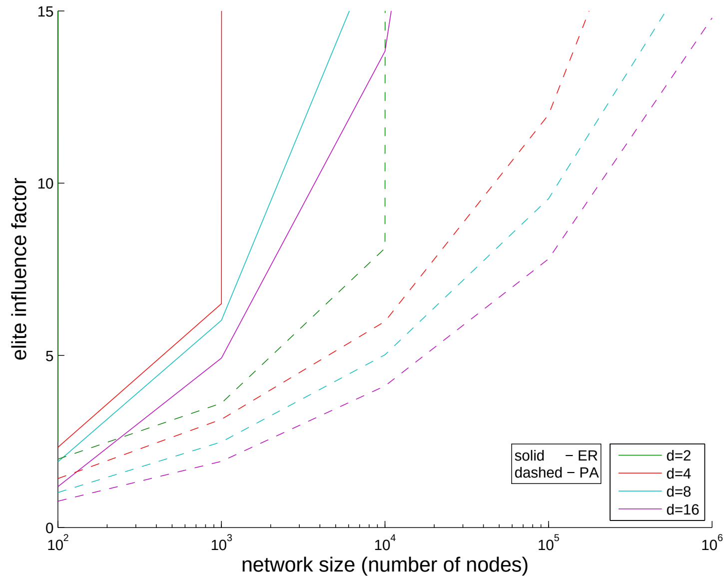


(a) Reversible

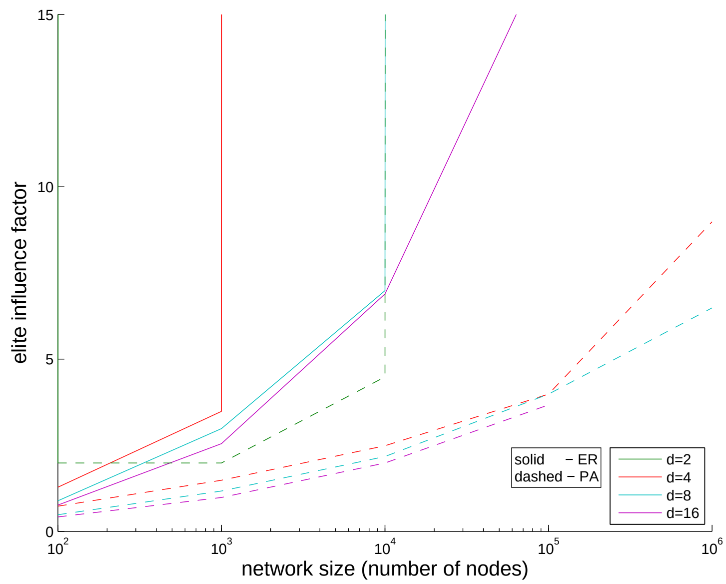


(b) Irreversible

Figure 4.7: The power of the elite of *Erdős–Rényi* and Preferential Attachment models. The minimum elite size to win when granted constant influence factor converges to larger scale compared to real world networks.



(a) Reversible



(b) Irreversible

Figure 4.8: The power of the elite of *Erdős–Rényi* and Preferential Attachment models. The influence factor required for an elite of size $\sqrt{m}$ grows super-linearly with time, unlike the observed growth for real world networks.

Chapter 5

Conclusions and Future Work

Our goal was to investigate the potential of small cooperating groups to control a decision making process based on social influence. Particularly, we checked the “potential for control” of the elite of the society: consider the members of the elite agree on an alternative, and if they all act in some specified way, then that alternative will be adopted by the rest of the society.

We chose the k-rich club to simulate the elite as these two groups are sharing common characteristics (small, well connected). We analyzed list of social networks to try and measure the “power of elite”. We used local majority voting to simulate the impact of social influence on decision taking when it is isolated from other sources of influence. Observations shown characteristic shared by all social networks: a relatively small elite, size of square root of the number of connections, has disproportionate influence on the rest of the society. As opinions are exchanged between connections, and given the elite is controlling large amount of the connections in the network and also robust against influence from the ”outside”, the elite is able to force its opinion on the rest of the society. In this case we say the elite has a “monopoly” over the decision to be taken. We found that this monopoly also holds over time as the society grows and newcomers are joining. We also observed that the Preferential Attachment model is unable to predict these amount of control the elite has over the rest of the society.

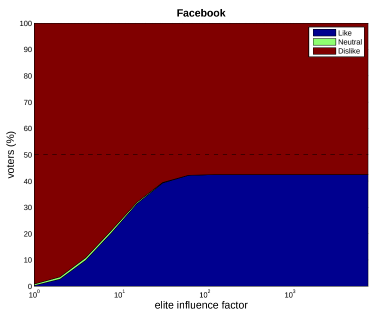
The results are raising several interesting directions for future work: we find it interesting to analyze a version of local majority voting when the decision is not taken by some clear cut but by some probability function (e.g. uniform), as this may better capture a real life decision model. In the same context, we suggest to analyze a model allowing individuals to behave in less "opportunistic" manner. Meaning, one will be querying his neighbors for their opinions only if the number of his friends which have changed their mind, since he last made himself mind, is over some threshold (unlike the repetitive model we used in this thesis). As for the critical mass in diffusion of innovation, it will be interesting to model this application using local majority voting and irreversibility as defined Peleg [47]: when one become "infected", he does not change his mind back. As the Preferential Attachment model is not predicting the elite monopoly phenomena, we think of providing a theoretical model for social network which is also generating an elite with monopoly characteristics.

Appendix A

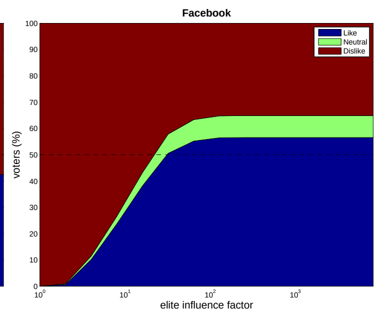
Local Majority Voting Propagation

As discussed in the introduction, the local voting process may be used for simulating diffusion (e.g. of opinions) or an infectious process in social networks. That is, a node is being considered as "infected" as the number of his infected neighbors is over some threshold. In the local *majority* voting model we used in this thesis, the threshold is 50%. We sampled the results of local majority voting after some rounds and examined the progress of votes along the process. Figure A.0 is showing the progress of local majority voting for several social networks from our data set. X-axis is showing the influence factor and the Y-axis is showing the percentage of nodes voting with the elite ("Like", blue), against ("Dislike", red) and neutral[1] (green). We observed a similar behavior for all networks: (i) there exist "phase transition" in which granting the elite with some factor is making the elite win the local majority voting. (ii) for a specific round, granting the elite with asymptotically increasing influence factor does not cause the elite to "convince" more nodes. So, recalling the influence factor is granted only for the elite and is not propagating to nodes which have become "convinced", then in fact, we can see that the ability of the elite to gain monopoly depends in its ability to create critical mass in the first round.

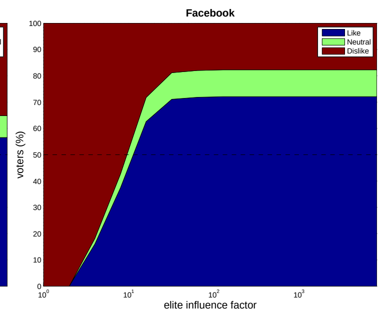
[1]We allowed neutral votes for this particular simulation.



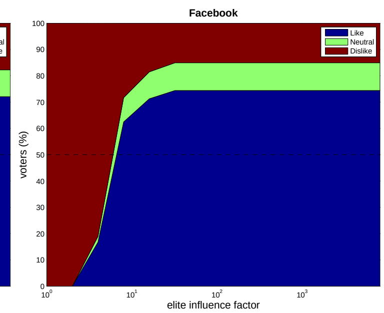
(a) Round 1



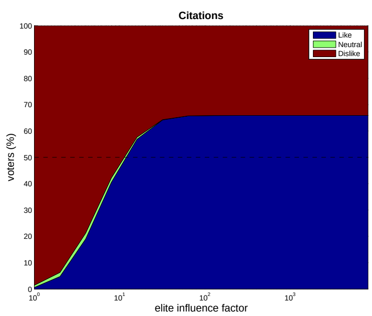
(b) Round 2



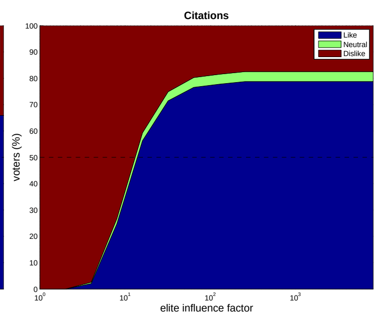
(c) Round 8



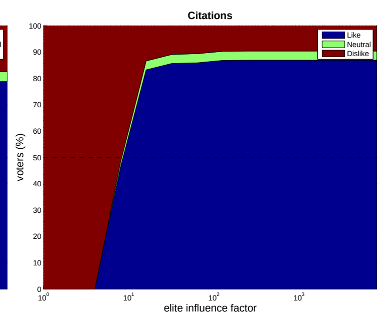
(d) Round 64



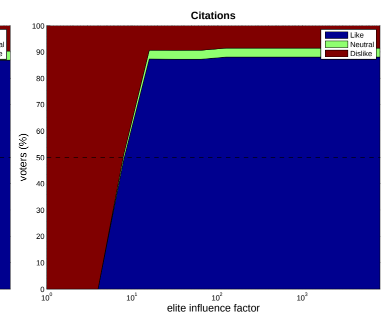
(e) Round 1



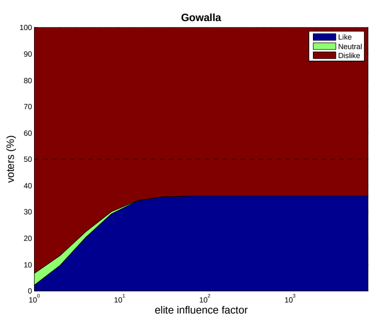
(f) Round 2



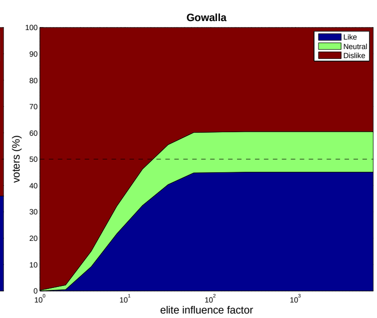
(g) Round 8



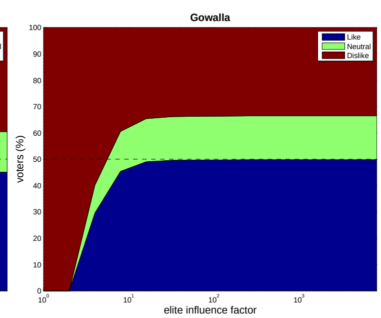
(h) Round 64



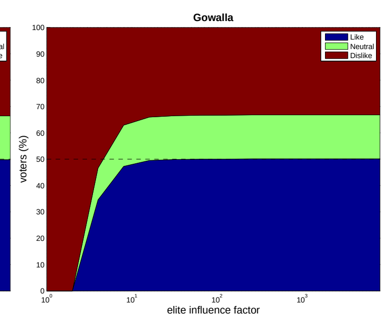
(i) Round 1



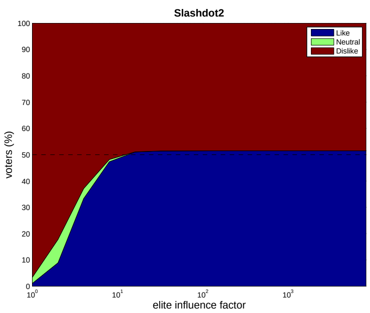
(j) Round 2



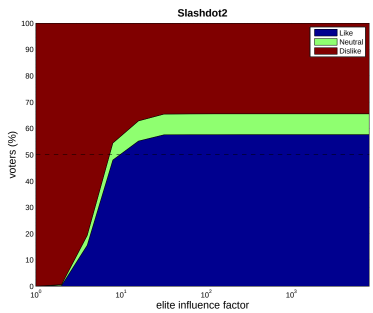
(k) Round 8



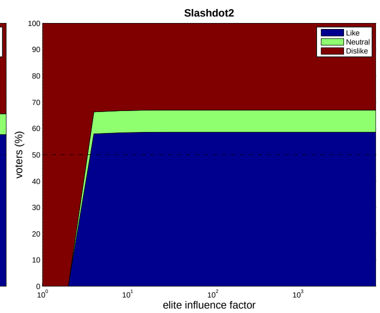
(l) Round 64



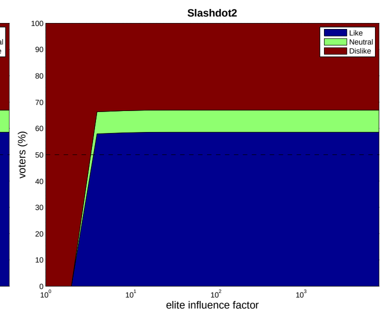
(m) Round 1



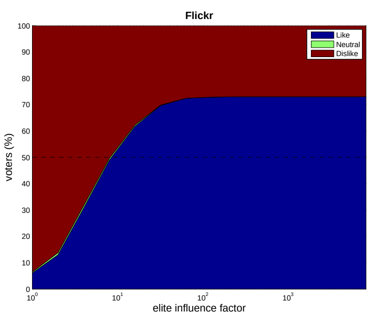
(n) Round 2



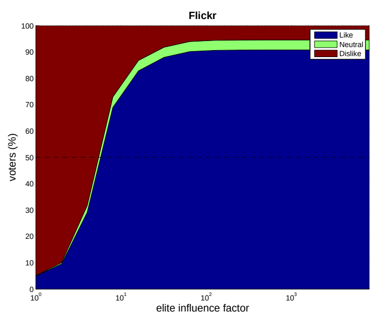
(o) Round 8



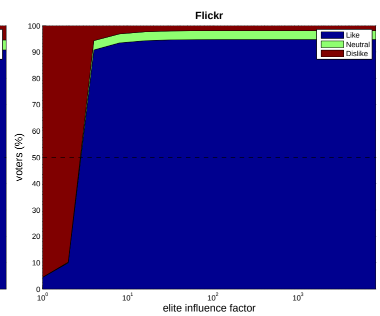
(p) Round 64



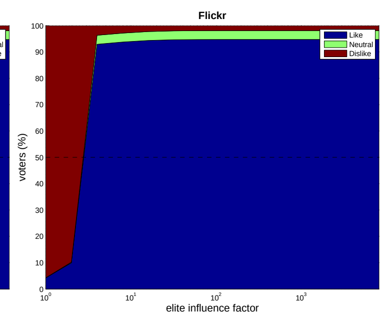
(q) Round 1



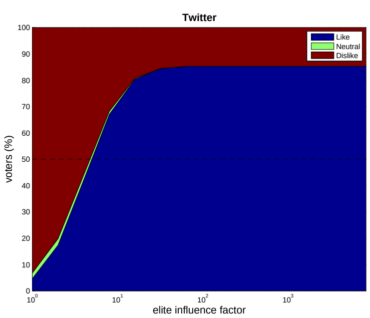
(r) Round 2



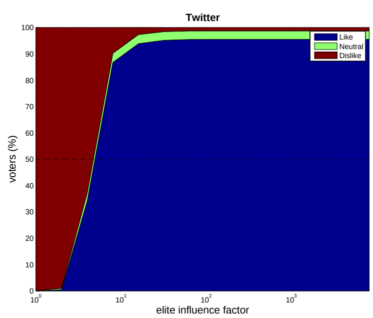
(s) Round 8



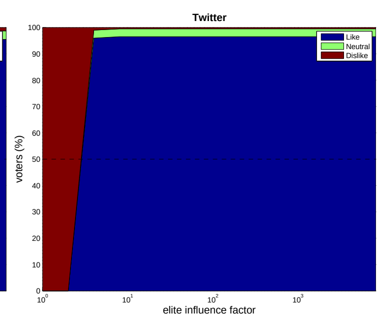
(t) Round 64



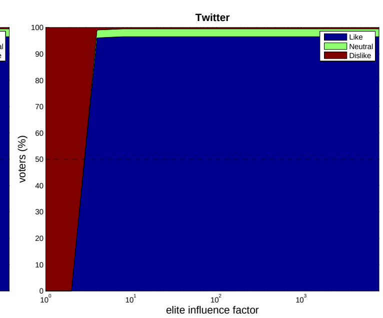
(u) Round 1



(v) Round 2



(w) Round 8



(x) Round 64

Figure A.0: Local majority voting progress (elite size $= \sqrt{m}$)

Bibliography

[1] The Stanford Large Network Dataset Collection (SNAP). `http://snap.stanford.edu/data/`, 2009.

[2] The Koblenz Network Collection (KONECT). `http://konect.uni-koblenz.de/`, 2012.

[3] E. Ackerman, O. Ben-Zwi, and G. Wolfovitz. Combinatorial model and bounds for target set selection. *Theor. Comput. Sci.*, 411:4017–4022, 2010.

[4] R. Albert and A.L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47–97, 2002.

[5] R. Albert, H. Jeong, and A.L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[6] H. Anderson. Epidemic models and social networks. *Mathematical Scientist*, 24:128–147, 1999.

[7] C. Avin, Z. Lotker, Y.-A. Pignolet, and I. Turkel. From caesar to twitter: Structural properties of elites and rich-clubs. *CoRR*, abs/1111.3374, 2012.

[8] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. *CoRR*, abs/1111.4570, 2011.

[9] A.L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.

[10] U. Brandes and J. Lerner. Structural similarity: Spectral methods for relaxed block-modeling. *J. Classification*, 27(3):279–306, 2010.

[11] R. C. Brigham, R. D. Dutton, T. W. Haynes, and S. T. Hedetniemi. Powerful alliances in graphs. *Discrete Mathematics*, 309:2140–2147, 2009.

[12] S. Brunetti, G. Cordasco, L. Gargano, E. Lodi, and W. Quattrociocchi. Minimum weight dynamo and fast opinion spreading. In *38th Int. Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, pages 249–261, 2012.

[13] C. C. Centeno, M. C. Dourado, L. Draque Penso, D. Rautenbach, and J. L. Szwarcfiter. Irreversible conversion of graphs. *Theor. Comput. Sci.*, 412:3693–3700, 2011.

[14] E. Cho, S.A. Myers, and J. Leskovec. Friendship and mobility: User movement in location-based social networks. In *Proc. Int. Conf. on Knowledge Discovery and Data Mining*, pages 1082–1090, 2011.

[15] F. Cicalese, G. Cordasco, L. Gargano, M. Milanic, and U. Vaccaro. Latency-bounded target set selection in social networks. In *9th Conf. on Computability in Europe (CiE)*, pages 65–77, 2013.

[16] R. A. Dahl. A critique of the ruling elite model. *The American Political Science Review*, 52(2):463–469, 1958.

[17] M.H. Degroot. Reaching a consensus. *J. American Statistical Association*, 69:118–121, 1974.

[18] P. Donnely and D. Welsh. Finite particle systems and infection models. *Proc. Camb. Phil. Soc.*, 94:167–182, 1983.

[19] C. Doob. *Social Inequality and Social Stratification in U.S Society.* Pearson Education, 2012.

[20] M. C. Dourado, L. Draque Penso, D. Rautenbach, and J. L. Szwarcfiter. Reversible iterative graph processes. *Theor. Comput. Sci.*, 460:16–25, 2012.

[21] T. R. Dye. *Who's Running America? The Bush Restoration.* Prentice Hall, 2002.

[22] D. Easley and J. Kleinberg. Networks, Crowds, and Markets @ONLINE, 2010.

[23] O. Favaron, G. Fricke, W. Goddard, S. Mitchell Hedetniemi, S. T. Hedetniemi, P. Kristiansen, and R. Duane Skaggs. Offensive alliances in graphs. *Discussiones Mathematicae Graph Theory*, 24:263–175, 2004.

[24] P. Flocchini, F. Geurts, and N. Santoro. Optimal irreversible dynamos in chordal rings. *Discrete Applied Mathematics*, 113:23–42, 2001.

[25] P. Flocchini, R. Kralovic, P. Ruzicka, A. Roncato, and N. Santoro. On time versus size for monotone dynamic monopolies in regular topologies. *J. Discrete Algorithms*, 1:129–150, 2003.

[26] P. Flocchini, E. Lodi, F. Luccio, L. Pagli, and N. Santoro. Dynamic monopolies in tori. *Discrete Applied Mathematics*, 137:197–212, 2004.

[27] J.R.P. French. A formal theory of social power. *Psych. Review*, 63:181–194, 1956.

[28] E. Goles and J. Olivos. Periodic behavior of generalized threshold functions. *Discrete Mathematics*, 30:187–189, 1980.

[29] E. Goles and J. Olivos. Periodic behaviowr functions. 30:187–189, 1980.

[30] V. Gómez, A. Kaltenbrunner, and V. López. Statistical analysis of the social network and discussion threads in Slashdot. In *Proc. Int. World Wide Web Conf.*, pages 645–654, 2008.

[31] A. Harutyunyan. Some bounds on global alliances in trees. *Discr. Appl. Math.*, 161:1739–1746, 2013.

[32] Y. Hassin and D. Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Inf. Comput.*, 171:248–268, 2001.

[33] T. W. Haynes, S. T. Hedetniemi, and M. A. Henning. Global defensive alliances in graphs. *Electr. J. Comb.*, 10, 2003.

[34] R. Holley and T.M. Ligget. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.*, 3:643–663, 1975.

[35] J. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845, August 2000.

[36] B. Klimt and Y. Yang. The Enron Corpus: A new dataset for email classification research. In *Proc. European Conf. on Machine Learning*, pages 217–226, 2004.

[37] J. Leskovec and E. Horvitz. Planetary-scale views on a large instant-messaging network. In *Proc. 17th Int. Conf. on World Wide Web (WWW)*, page 915, New York, New York, USA, 2008. ACM Press.

[38] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowledge Discovery from Data*, 1(1):1–40, 2007.

[39] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *CoRR*, abs/0810.1355, 2008.

[40] N. Linial, D. Peleg, Y. Rabinovich, and M. Saks. Sphere packing and local majorities in graphs. In *Theory and Computing Systems*, pages 141–149. IEEE, 1993.

[41] P. Massa and P. Avesani. Controversial users demand local trust metrics: an experimental study on epinions.com community. In *Proc. American Assoc. for Artificial Intelligence Conf.*, pages 121–126, 2005.

[42] J. McAuley and J. Leskovec. Learning to discover social circles in ego networks. In *Advances in Neural Information Processing Systems*, pages 548–556. 2012.

[43] S. Milgram. The small world problem. *Psychol. Today*, 2:60–67, 1967.

[44] A. Mislove, M. Marcon, K.P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In *Proc. 7th ACM SIGCOMM Conf. on Internet measurement*, IMC '07, pages 29–42, New York, NY, USA, 2007. ACM.

[45] V. Pareto. *Mind and Society.* Harcourt Brace Jovanovich, 1935.

[46] D. Peleg. Size bounds for dynamic monopolies. *Discrete Applied Mathematics*, 86:263–273, 1998.

[47] D. Peleg. Local majorities, coalitions and monopolies in graphs: a review. *Theor. Comput. Sci.*, 282(2):231–257, June 2002.

[48] S. Poljak and M. Sura. On periodical behavior in societies with symmetric influences. *Combinatorica*, 3:119–121, 1983.

[49] S. Poljak and M. Sůra. On periodical behaviour in societies with symmetric influences. *Combinatorica*, 3(1):119–121, 1983.

[50] S. Poljak and D. Turzík. On an application of convexity to discrete systems. *Discr. Appl. Math.*, 13(1):27 – 32, 1986.

[51] D. Reichman. New bounds for contagious sets. *Discr. Math.*, 312:1812–1814, 2012.

[52] B. Viswanath, A. Mislove, M. Cha, and K.P. Gummadi. On the evolution of user interaction in Facebook. In *Proc. 2nd ACM Workshop on Online Social Networks (WOSN)*, pages 37–42, New York, New York, USA, 2009. ACM Press.

[53] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–2, June 1998.

[54] S. Zhou and R.J. Mondragón. The rich-club phenomenon in the internet topology. *IEEE Commun. Lett.*, 8(3):180–182, 2004.