Voting to the Link: a Static Network Formation Model

Róbert Pethes and Levente Kovács

Physiological Controls Research Center
Research and Innovation Center of Óbuda University
Bécsi út 96/b, H-1034 Budapest, Hungary
pethes.robert@phd.uni-obuda.hu, kovacs.levente@nik.uni-obuda.hu

Abstract: It is well known that the structure of social, organization and economic networks have a huge effect on the behaviour of the underlying system. This structure is often considered as a network, and modelling the formation of these networks is an active research area of complex systems. In this paper we present a simple network generation model: there is a closed population of agents, and the agents are voting to the connections according to their fixed preferences. These preferences are denoted by real numbers, and can be considered as simply the desired number of connections or social capital.

Keywords: static random graph model;inhomogeneous random graphs; social networks

1 Introduction

Networks play a central role in determining the outcome of many social and economic relationships as well in the effectiveness of organizations such as companies. The fact, that we live in hierarchies [1] is a fundamental feature of our society and one of the basic organization rule of structures such as companies, the government, schools, army, etc. Social networks are also important in determining how diseases spread [2] ,which products we buy [3], how we vote [4] and so on. These fields are very interesting in their own right, but better understanding of them can have several practical consequences. Avoiding the spreading of fake news [5], better infection control [6], improving traffic and designing smart cities [7] are just a few examples of many. At this point it is useful to define two crude categories where networks are important in sociology and economy: In the first one, the network structure is a distribution or service network, in the second, the network connects different individuals (persons, companies, countries). In the former case the structure is usually a result of some kind of design, in the latter case however the formation of the network depends on the independent and often self-interested agents. This latter case is much more interesting for us. In case of modelling social and economic network formation, the models come mainly from two fundamental disciplines: the first is the theory of random networks, where the network formation is considered as a

random process, and the second is game theory, where strategic models of how networks are formed are developed. The static voting model is in the random network model family. Here we are not going to review the huge area of random networks, but just mention some important milestones. The field of random graphs was first introduced by the famous paper of Erdos and Rényi [8]. The Erdos-Rényi $ER_n(p)$ random graph has n vertices and each pair of vertices is independently connected with probability p. Despite the fact that $ER_n(p)$ is the simplest imaginable model of a random network, it has fascinating phase transition when p varies. Looking at many real world networks [9] such as social network, Internet, transport networks, biological networks, etc., we can see that their degree distribution is power-law, and they have the so called small-world property. Since the ER model has Poisson degree distribution, it is not a good model of real world networks. There are many extensions of this random graph model, for example the inhomogeneous version of the ER model [9] that we will use later in this paper, the configuration model [9], the generalized random graph model [9], and the most general form of the inhomogeneous random graphs [10] defined by kernel functions. In this paper we only deal with static models, but we have to mention one network growth model family: the class of preferential attachment models are such that new elements are more likely to attach to elements with high degree compared to elements with small degree. This phenomenon was first published by Barabási and Albert [11] and their model is called Barabási - Albert (BA) model. Networks generated by the BA model show power-law degree distribution. In this paper we continue and extend the work on static edge voting models [12]. The structure of this paper is as follows: after an introduction to the necessary notions and definitions, in Section 2 we turn to the Inhomogeneous Random Graphs. In this section we have two main points: the Chernoff-bounds and the fact that we can use the Gauss distribution to approximate the distribution of the degree of a given node in the network. In Section 3 we define the static edge voting model (SEV), introduce the Mixture of Gaussians (MOG) method to estimate the degree distribution of the graph (Subsection 3.1) and discuss how we generate the parameters of the SEV models to be able to execute numerical tests (Subsection 3.2). In Subsections 3.3, 3.4 and 3.5 we define and analyse 3 special cases of the SEV model: the proportional, the Poission and the biased voting model. We will demonstrate the usability of the MOG approximation with the proportional model, and in the following models we will use only this.

1.1   Notations and definitions

We denote the set $\{1,...,n\}$ with $[n]$. In our models there are only simple graphs. A G simple graph is given by a pair: $G = (V(G), E(G))$, where $V(G)$ is the vertex set of the G graph and $E(G)$ contains the edges of the graph. The vertices are marked with integer numbers, so the vertex set of the graph with n vertices is $V = [n]$. The degree of a node in the graph is definded as the number of neighbours of the node. For a given $a \in V$ we denote the degree of a by $d(a)$ or simply by d_a. We can enumerate the degrees of all the nodes in a sequence: $(d_1,...,d_n)$. This sequence is called the degree sequence of the G graph. We define the degree distribution of the G graph as the distribution of $d(U)$, where U is a randomly and uniformly chosen node. Therefore $d(U)$ is a random variable even in case of a deterministic graph.

In deterministic case we can express $\mathbb{P}(d(U) = k)$ as $\#\{a | a \in V, d_a = k\}/n$. If G is random we can handle d_a as a random variable, and we refer to the distribution of d_a as the distribution of the degree of the node a. We will use the abbreviation: $V_i = V \setminus \{i\}$. Now we define the inhomogeneous random graphs (IRG) as it is defined in [9]. Lets denote $P = \{p_{ij}\}$, $1 \le i < j \le n$, the edge probabilities in the graph, and $IRG_n(P)$ for the inhomogeneous random graph for which the edges are drawn independently, and the probability that the edge $\{i, j\}$ is present in the graph equals p_{ij}. All the models in this paper have an equivalent IRG representation, and we will use either the original definition or the equivalent $IRG_n(P)$ depending on which one is more suitable.

In order to compare discrete probability distributions, we use two methods: total variation distance [9] and Jensen-Shannon divergence [13]. In general, for two probability measures μ and ν, the total variation distance is defined as:

$$d_{TV}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|$$

For discrete probability distributions with the same $\mathbb{X}$ support $(p_x)_{x \in \mathbb{X}}$ and $(q_x)_{x \in \mathbb{X}}$, these measures are given by $\mu(A) = \sum_{a \in A} p_a$ and $\nu(A) = \sum_{a \in A} q_a$. In this case it becomes:

$$d_{TV}(p, q) = \frac{1}{2} \sum_x |p_x - q_x|$$

The Jensen-Shannon divergence is the symmetrized and smoothed version of the Kullback–Leibler divergence, so first we have to define it. For discrete probability distributions $(p_x)_{x \in \mathbb{X}}$ and $(q_x)_{x \in \mathbb{X}}$ defined on the same probability space:

$$KL(p, q) = \sum_{x \in \mathbb{X}} p_x \log \left(\frac{p_x}{q_x} \right)$$

The Kullback–Leibler divergence is defined only if for all x, $q_x = 0$ implies that $p_x = 0$. Whenever p_x is zero the contribution of the corresponding term is interpreted as zero. The KL divergence is not symmetric. To make it symmetric define the Jensen-Shannon divergence as:

$$JS(p, q) = \frac{1}{2} KL(p, m) + \frac{1}{2} KL(q, m)$$

where the $(m_x)_{x \in \mathbb{X}}$ distribution is defined as: $m_x = \frac{1}{2}(p_x + q_x)$

2 Properties of Inhomogeneous Random graphs

It would be worthwhile to know the distribution of the d_i degree of an arbitrary i vertex in the IRG model. However, computationally it is very expensive to compute the proper distribution. Generally:

$$\mathbb{P}(d_i = k) = \sum_{S \subseteq V_i, |S| = k} \prod_{j \in S} p_{ij} \prod_{j \in V_i \setminus S} (1 - p_{ij}) \tag{1}$$

In our preliminary paper [12] we have formalized the following proposition, and the proof is also given there. This is not a new theorem, but only the reformulation of the Chernoff bound [9] for the case of inhomogeneous random graphs.

Proposition 1. *If given a $IRG_n(P)$ model where $P = \{p_{ij}\}$, $i \leq i < j \leq n$, and d_i is the degree of the vertex i, then for all $D \in \mathbb{N}$:*

$$\mathbb{P}(d_i \geq D) \leq e^{-tD} M_i(t) \text{ for all } t > 0$$

and

$$\mathbb{P}(d_i \leq D) \leq e^{-tD} M_i(t) \text{ for all } t < 0$$

where

$$M_i(t) = \prod_{[n]\setminus i}(p_{ij}e^t + 1 - p_{ij})$$

Now, we prove the following corollary of Proposition 1:

Corollary 1. *If given a $IRG_n(P)$ model where $P = \{p_{ij}\}$, $i \leq i < j \leq n$, and d_i the degree of the vertex i, then for all $D \in \mathbb{N}$:*

$$\mathbb{P}(d_i \geq D) \leq (L(D))^{n-1} \text{ if } D > (n-1)\overline{p_i} \tag{2}$$

and

$$\mathbb{P}(d_i \leq D) \leq (L(D))^{n-1} \text{ if } D < (n-1)\overline{p_i} \tag{3}$$

where:

$$L(D) = \overline{p_i}\left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1-\frac{D}{n-1})}\right)^{1-\frac{D}{n-1}} + (1-\overline{p_i})\left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1-\frac{D}{n-1})}\right)^{-\frac{D}{n-1}} \tag{4}$$

and

$$\overline{p_i} = \frac{1}{n-1}\sum_{j \in V_i} p_{ij}$$

Proof. Using the notations of Proposition 1:

$$M_i(t) = \left(\left(\prod_{[n]\setminus i}(p_{ij}e^t + 1 - p_{ij})\right)^{\frac{1}{n-1}}\right)^{n-1}$$

Applying the inequality $(\prod x_i)^{\frac{1}{n}} \leq \frac{1}{n}\sum x_i$ and rearranging the terms, we have:

$$e^{-tD}M_i(t) \leq (L(D,t))^{n-1}$$

where:

$$L(D,t) = e^{-t\frac{D}{n-1}} \left(\overline{p_i}e^t + 1 - \overline{p_i} \right)$$

$$\frac{\partial L(D,t)}{\partial t} = -\frac{D}{n-1} e^{-t\frac{D}{n-1}} \left(\overline{p_i}e^t + 1 - \overline{p_i} \right) + e^{-t\frac{D}{n-1}} \overline{p_i}e^t =$$

$$= e^{t(1-\frac{D}{n-1})} \overline{p_i} \left(1 - \frac{D}{n-1} \right) + (\overline{p_i} - 1) \frac{D}{n-1} e^{-t\frac{D}{n-1}}$$

And the second derivative:

$$\frac{\partial^2 L(D,t)}{\partial^2 t} = e^{t(1-\frac{D}{n-1})} \overline{p_i} \left(1 - \frac{D}{n-1} \right)^2 - (\overline{p_i} - 1) \left(\frac{D}{n-1} \right)^2 e^{-t\frac{D}{n-1}}$$

Since $\frac{\partial^2 L(D,t)}{\partial^2 t}$ is positive for all t therefore $L(D,t)$ is convex for any fixed D, so $L(D,t)$ has a unique minimum, what we can get solving $\frac{\partial L(D,t)}{\partial t} = 0$ for t:

$$t^* = argmin_t L(D,t) = ln \left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1 - \frac{D}{n-1})} \right)$$

After substitution back to $L(D,t)$:

$$L(D) = L(D,t^*) = \overline{p_i} \left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1-\frac{D}{n-1})} \right)^{1-\frac{D}{n-1}} + (1-\overline{p_i}) \left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1-\frac{D}{n-1})} \right)^{-\frac{D}{n-1}}$$

We still have to show that in case of (2) $t^* > 0$ and similarly in case of (3) $t^* < 0$. Lets start with (2) and assume that $\frac{D}{n-1} > \overline{p_i}$

$$t^* = ln \left(\frac{(1-\overline{p_i})\frac{D}{n-1}}{\overline{p_i}(1 - \frac{D}{n-1})} \right) > 0$$

if and only if

$$(1-\overline{p_i})\frac{D}{n-1} > \overline{p_i} \left(1 - \frac{D}{n-1} \right)$$

After rearrange the terms, we get:

$$\frac{D}{n-1} > \overline{p_i}$$

Similarly, if we assume that $\frac{D}{n-1} < \overline{p_i}$ we can see that $t^* < 0$ in the same way.

$\square$

Corollary 1 is a less strict, but easier to use theorem when $L(D) < 1$ holds. In ref [12], we showed that we can find an approximation of the degree distribution of node i with the $\mathcal{N}(\mu, \sigma)$ Gauss distribution, setting the parameters to $\mu = \mathbb{E}[d_i] = \sum_{j \in V_i} p_{ij}$ and $\sigma^2 = Var[d_i] = \sum_{j \in V_i} p_{ij}(1 - p_{ij})$. The Lyapunov variant of the central limit theorem [14] gives the theoretical background for this. In this paper we refer to this approximation method as the Gauss approximation of the degree distribution of the nodes.

3 The static edge voting model

We divide the set of agents to disjoint groups: $V = [N] = S_0 \cup ... \cup S_M$, and suppose that the behaviour of the agents belonging to a group is the same. Denote V_{ab} a random variable that means the vote of the a node to the $\{a, b\}$ edge. For all $a \in S_i$ and $b \in S_b$, V_{ab} has the same probability distribution. For any different nodes a and b the probability that there will be an edge between the nodes a and b in the graph is up to the incoming votes and it is equal to $s(V_{ab}, V_{ba})$, where s is the edge probability function.

Definition 1. *Denote $SEV_N(\{S_0, ..., S_M\}, \{F_{ij} | i, j \in \{0, ..., M\}\}, s)$ the static edge voting model, where:*

- *N is the number of agents and $V = [N]$ is the set of agents.*

- *$\{S_0, ..., S_M\}$ is a partition of the V set and M is the number of groups in the partition. (Here we allow empty sets in the partition.)*

- *Denote V_{ab} the random variable that means the vote of the a node to the $\{a, b\}$ edge candidate.*

- *$\{F_{ij} | i, j \in \{0, ..., M\}\}$ is a set of probability distributions. For all $a \in S_i$ and $b \in S_b$ the law of V_{ab} is F_{ij}.*

- *$s : \mathbb{R} \times \mathbb{R} \mapsto [0, 1]$ is the edge probability function, a non-decreasing integrable function in both arguments. Every $\{a, b\}$ ($a \neq b$) node pair will be part of the graph with probability $s(V_{ab}, V_{ba})$.*

We proved in [12] the following simple proposition:

Proposition 2. *Given a $SEV_N(\{S_0, ..., S_M\}, \{F_{ij} | i, j \in \{0, ..., M\}, s\}$ static edge voting model the expected value of the d_a degree of the a node, if $a \in S_i$:*

$$\mathbb{E}[d_a | a \in S_i] = \sum_{j=0}^{M} k_j \mathbb{E}[s(V_{ab}, V_{ba}) | b \in S_j] - \mathbb{E}[s(V_{ab}, V_{ba}) | b \in S_i, a \neq b]$$

where $k_i = |S_i|$

3.1 Approximation of the degree distribution

It is clear that the equivalent $IRG_n(P)$ of the $SEV_N(\{S_0, ..., S_M\}, \{F_{ij} | i, j \in \{0, ..., M\}\}, s)$ is the *IRG* with edge probabilities:

$$P_{ab} = \mathbb{P}(\{a, b\} \in E)$$

We can use the total probability rule to express the degree distribution of the SEV_N random network:

$$\mathbb{P}(d(U) = k) = \sum_{i=0}^{M} \mathbb{P}(U \in S_i) \mathbb{P}(d(U) = k | U \in S_i) \tag{5}$$

We can get $\mathbb{P}(U \in S_i) = \frac{\#S_i}{N}$ directly from the SEV_N model, and we can compute $\mathbb{P}(d(U) = k | U \in S_i)$ directly using the equivalent $IRG_n(P)$ model and equation (2). However, the direct use of equation (2) is computationally very expensive. In spite of this, we can use the Mixture of Gaussian's (MOG) model: apply the Gauss approximation of $\mathbb{P}(d(U) = k | U \in S_i)$ as defined in Section 2. For arbitrary $a \in V$ denote $\mu_a = \sum_{b=1}^{N} P_{ab}$ and $\sigma_a^2 = \sum_{b=1}^{N} (P_{ab}(1 - P_{ab}))$. It is easy to see that for a fixed S_i and for arbitrary $a, b \in S_i$, $P_{ac} = P_{bc}$ for all $c \in V$. Therefore the μ_a and σ_a^2 values are the same for all $a \in S_i$, that we denote by μ_{S_i} $\sigma_{S_i}^2$. We use the estimation $\mathbb{P}(d(U) = k | U \in S_i) \approx F(k + 0.5; \mu_{S_i}, \sigma_{S_i}^2) - F(k - 0.5; \mu_{S_i}, \sigma_{S_i}^2)$, where $F(x; \mu_{S_i}, \sigma_{S_i}^2)$ is the cumulative distribution function of the Gauss distribution with parameters μ_{S_i} and $\sigma_{S_i}^2$. Plugging these back to equation (5), we have

$$\mathbb{P}(d(U) = k) \approx \sum_{i=0}^{N-1} \mathbb{P}(U \in S_i) \left(F(k + 0.5; \mu_{S_i}, \sigma_{S_i}^2) - F(k - 0.5; \mu_{S_i}, \sigma_{S_i}^2) \right) \qquad (6)$$

3.2 Degree sequences

In the rest of the paper we would like to present numerical examples, and for this, we need to generate $D = (D_1, ..., D_n)$ degree sequences. We use four methods to generate these sequences: constant, uniform, binomial and lognormal. The constant degree sequence is the simplest: we just create a N-length sequence of constant values: $ConstSeq(C, N) = (C, ..., C)$. The second method is the uniform. We denote this with $UniformSeq(MIN, MAX, N)$, where MIN, MAX and N are the minimum, the maximum and the size of the sequence. The last two are the binomial and lognormal, denoted by $BinomSeq(p, N)$ and $LognormalSeq(\mu, \sigma^2, N)$, where p and N are the parameters of the binomial distribution and N is the length of the sequence at the same time, while μ and σ^2 are the parameters of the lognormal distribution. We omit here the algorithm what we used generating these sequences. The main point here is that the values in the generated sequences are approximately obey the uniform, lognormal and binomial laws. After the degree sequence generation, we computed the empirical distribution of each sequence. In case of the sequences $BinomSeq(0.1, 25)$, $LognormalSeq(1.7, 0.4, 25)$, $BinomSeq(0.1, 1000)$, $LognormalSeq(4, 0.6, 1000)$ the result distributions are plotted on Figure 1. Note that except the constant sequence case, the empirical distributions of the sequences are just the approximations of the source distributions.

3.3 The proportional edge voting model

The static proportional edge voting model is a static random network model where the V_{ab} random variables do not depend on b, and for all $a \in S_i$, $V_{ab} \sim Bernoulli(\frac{i}{N-1})$ and $s(x) = x/2$. This model can be interpreted as that the agent a who belongs to the group S_i would like to achieve $D_a = i$ degree. The probability that the a agent gives a vote on the $\{a, b\}$ edge candidate is $\frac{D_a}{N-1}$. This voting rule justifies the "proportional" name. The direct consequence of this setup is that the number of agent groups $M = N - 1$, and S_i is the group of agents who target degree is i. The general definition of the static edge voting model assumes that the distribution of the
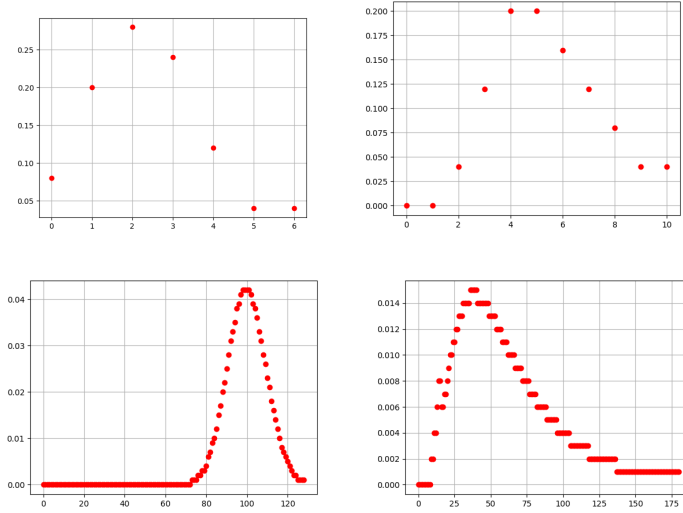
Figure 1
The empirical distribution of the degree sequences: $BinomSeq(0.1, 25)$ (top-left),
$LognormalSeq(1.7, 0.4, 25)$ (top-right), $BinomSeq(0.1, 1000)$ (bottom-left), $LognormalSeq(4, 0.6, 1000)$
(bottom-right)

V_{ab} random variables are given. In case of the proportional model this is equivalent with that the $\{D_a | a \in V\}$ degree sequence is given, where D_a can be considered as the desired degree of agent a. It is clear that:

$$\mathbb{P}(\{a, b\} \in E) = \frac{1}{2} \frac{D_a + D_b}{N - 1}$$

From Proposition 2 we can directly get:

$$\mathbb{E}[d_a | a \in S_i] = \frac{1}{2(N-1)} \left(i(N-2) + \sum_{j=0}^{N-1} k_j j \right)$$

Using the estimation: $N - 1 \approx N$ and $N - 2 \approx N$ we have:

$$\mathbb{E}[d_a | a \in S_i] \approx \frac{1}{2N} \left(iN + \sum_{j=0}^{N-1} k_j j \right) = \frac{1}{2} \left(i + \sum_{j=0}^{N-1} \frac{k_j}{N} j \right) = \frac{1}{2} \left(i + \overline{D} \right) \tag{7}$$

Where $\overline{D} = \sum_{j=0}^{N-1} \frac{k_j}{N} j$ is the mean target degree. This result shows that this model has a "smoothing" effect on the degrees. Equation 7 also shows that if an agent a has a desired degree D_a close to the average $\overline{D}$ desired degree, then she/he can easy achieve D_a in mean, however when D_a is far from $\overline{D}$, then hitting the aims is much harder: the $\overline{D}$ average value has a gravity.
Lets continue with the degree distribution of the static edge voting model. For

this, use the four degree sequence generator methods to generate the desired degree sequences. To be able to compute the exact degree distributions using equations (2) and (5) first we use a small network with 25 nodes. The desired degree sequences are: $ConstSeq(2,25)$, $BinomSeq(0.1,25)$, $LognormalSeq(1.7,0.4,25)$ and $UniformSeq(0,4,25)$. The degree distribution of each case is plotted on Figure 2. On Figure 3 we also plotted the MOG approximation of the degree distributions. In order to compare the exact degree distributions and the MOG approximations, we computed the total variation distance and the Jensen-Shannon divergence of the distributions. The results are summarised in Table 1. We can conclude that the exact and the MOG approximated distributions are close to each other. We can not compute the exact degree distribution for greater networks, because it is computationally very expensive. However, in one special case, we know the exact distribution: in case of $ConstSeq(C,N)$ desired degree sequence the probability of each edge is $p(N,C) = \frac{C}{N-1}$, therefore this model is equivalent to the Erdos-Renyi random graph: $ER(N, p(N,C))$. We know that the degree distribution of this ER model is binomial with parameters $N-1$ and $p(N,C)$. We computed therefore the total variation distance and the Jensen-Shannon divergence of the binomial distribution with parameters $N-1$ and $p(N,C)$ and the MOG approximation of the degree distribution with desired degree sequence $ConstSeq(C,N)$, where $N \in \{10, 20, ..., 2000\}$ and $C = N/10$. The results are in Figure 4. We can see, that both distances decrease as N increases, however the rate of decline is decreasing. We can consider the proportional edge voting model as a transformation of the distribution of the input parameters, e.g. the distribution of the desired degree sequence. To be able to illustrate this with numerical examples, we set the desired degree sequence to $ConstSeq(100,1000)$, $LognormalSeq(4,0.6,1000)$, $BinomSeq(0.1,1000)$ and $UniformSeq(0,99,1000)$. We used the MOG approximation of the degree distribution, and the results are in Figure 5. We also computed the empirical mean and variance of the desired degree distribution and the approximated degree distribution of the proportional model. These values are collected in Table 2. From Equation (7) we may expect that the model has a shrinking effect to the desired degree distribution. However, this phenomenon is valid only in case of the lognormal and uniform distributions. The constant degree sequence is a special case, but it is not clear why this effect does not work when the input degree distribution is binomial. Finally, we apply Corollary 1 to the IRG equivalent with the proportional edge voting model to bound the probability that the degree of a given node is greater/smaller than a value. In this numerical example, we set the desired degree sequence to $LognormalSeq(4,0.6,1000)$ and apply Corollary 1 to the nodes with degrees 15, 60, 130, 170 respectively. Note that the average desired degree in this case is 61.2919. On Figure 6 we plotted the bound probability in the function of D. In the case, when the desired degree of the node is 15, we can realize that the bound probabilities have a maximum around 38. This observation meets with equation (7). We can do the same observation when the desired degree of the the node is 60, 130 or 170. Looking at the shape of the curves, we may have the idea, that the Chernoff bound and the Gauss approximation of the degree distribution of a given node give similar results.

Table 1

Total variation distance (TV) and Jensen-Shannon divergence (JS) of the exact degree distributions and the MOG approximation of the degree distributions when the desired degree sequences are: $ConstSeq(2, 25)$, $BinomSeq(0.1, 25)$, $LognormalSeq(1.7, 0.4, 25)$ and $UniformSeq(0, 4, 25)$

	ConstSeq	BinomSeq	LognormalSeq	UniformSeq
TV	0.0668770019	0.0459734622	0.0257139778	0.0535480364
JS	0.0035851277	0.0015726680	0.0008468790	0.0022105596
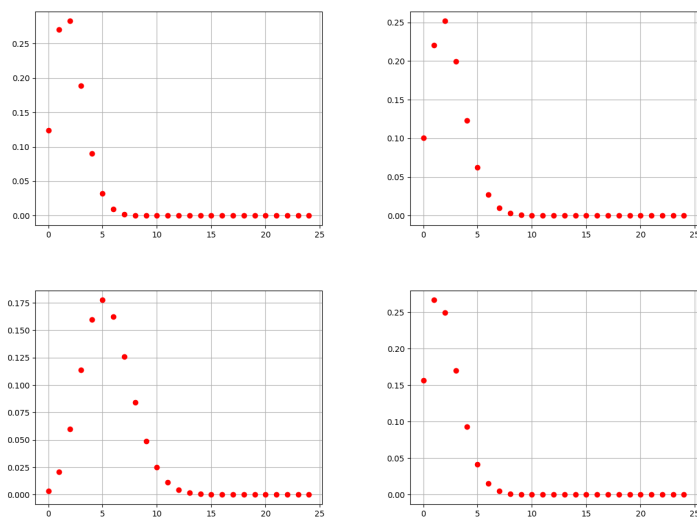
Figure 2

The exact degree distributions of the proportional edge voting model when the desired degree sequences are: $ConstSeq(2, 25)$ (top-left), $BinomSeq(0.1, 25)$ (top-right), $LognormalSeq(1.7, 0.4, 25)$(bottom-left), $UniformSeq(0, 9, 25)$ (bottom-right)

Table 2

The mean and the variance of the desired degree distribution (μ_D, σ_D) and the degree distribution of the proportional edge voting model using the Mixture of Gaussians approximation $(\mu_{d(U)}, \sigma_{d(U)})$

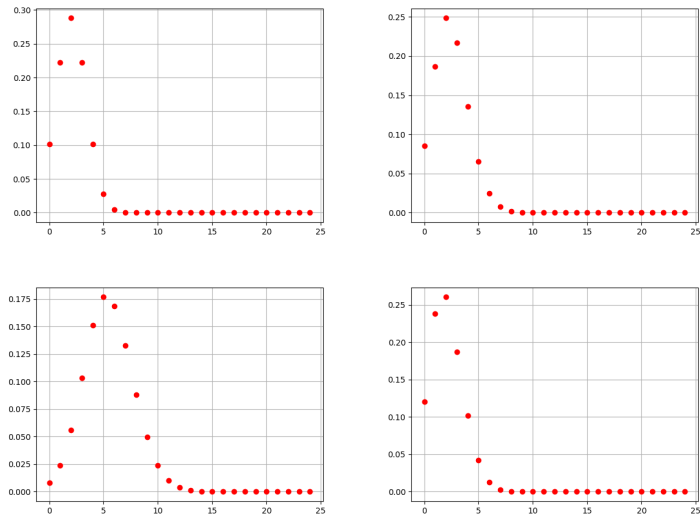|	ConstSeq	BinomSeq	LognormalSeq	UniformSeq
μ_D	100	99.88	61.2919	49.5
σ_D	0	87.8295	1242.1627	833.25
$\mu_{d(U)}$	100	99.88	61.292	49.5
$\sigma_{d(U)}$	90.0733	111.8468	366.913	254.6096

Figure 3
The degree distributions of the proportional edge voting model computing with the MOG
approximation, when the desired degree sequences are: $ConstSeq(2,25)$ (top-left), $BinomSeq(0.1,25)$
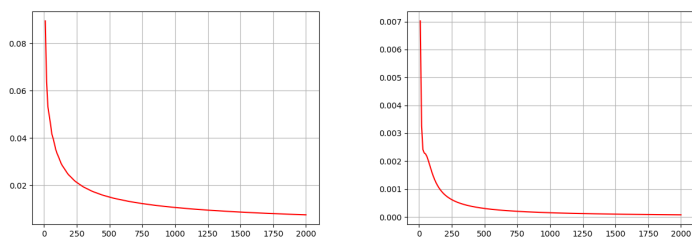(top-right), $LognormalSeq(1.7,0.4,25)$(bottom-left), $Uniform(0,9,25)$ (bottom-right)

Figure 4
Total variation distance (left) and the Jensen-Shannon divergence (right) of the binomial distribution
with parameters $N-1$ and $p(N,C)$ and the MOG approximation of the degree distribution with desired
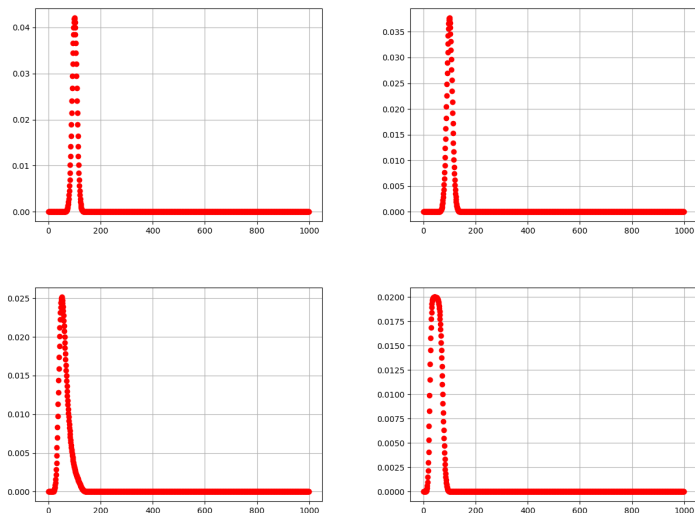degree sequence $ConstSeq(C,N)$, where $N \in \{10, 20, ..., 2000\}$ and $C = N/10$

Figure 5

The degree distributions of the proportional edge voting model computing with the Mixture of Gaussians approximation, when the desired degree sequences are: $ConstSeq(100, 1000)$ (top-left), $BinomSeq(0.1, 1000)$ (top-right), $LognormalSeq(4, 0.6, 1000$(bottom-left), $Uniform(0, 99, 1000)$ (bottom-right)
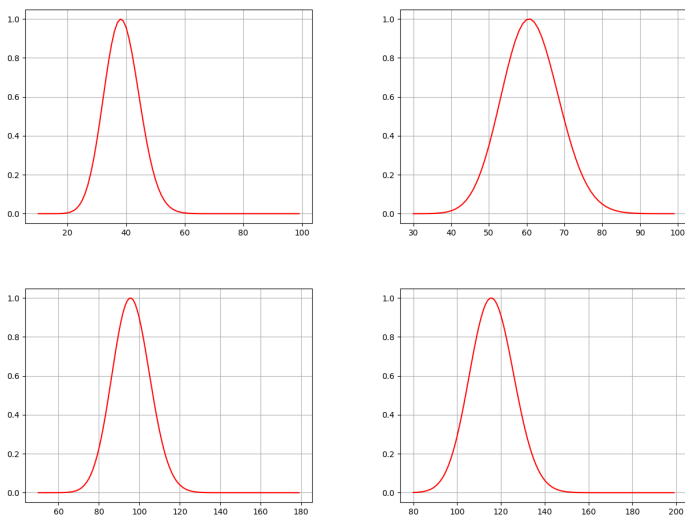
.

Figure 6

Applying the additive version of the Chernoff bound (Corollary 1) to the IRG equivalent with the proportional edge voting model with $LognormalSeq(4, 0.6, 1000)$ desired degree sequence. We plotted the bound probability in the function of D. The desired degree of the selected nodes are 15 (top-left), 60 (top-right),130 (bottom-left) and 170 (bottom-right).

3.4   Voting model with Poisson voting law

Let's discuss an example when the nodes may have more then one vote to a link. Consider the case when $V_{ab} \sim$ Poisson (λ_i) if $a \neq b$ and $a \in S_i$ and the edge probability function is given as $s(x) = 1 - e^{-\lambda x}$, where λ is the control parameter. We have shown [12] that the probability of the $\{a,b\}$ edge is

$$\mathbb{P}(\{a,b\} \in E | a \in S_i, b \in S_j) = 1 - e^{(\lambda_i + \lambda_j)(e^{-\lambda} - 1)} \tag{8}$$

And the expected degree of the a node is:

$$\mathbb{E}[d_a | a \in S_i] = N - 1 + e^{2\lambda_i(e^{-\lambda} - 1)} - \sum_{j=0}^{N-1} e^{(\lambda_i + \lambda_j)(e^{-\lambda} - 1)} \tag{9}$$

The parameters of the Poission voting model are $\{\lambda_i | i \in V\}$ and the λ control parameter. We would like to compare the behaviour of the Poission model to the proportional model. To be able to do this we use the same sequences to set the λ_i parameters of the model: $ConstSeq(100, 1000)$, $BinomSeq(0.1, 1000)$, $LognormalSeq(4, 0.6, 1000)$ and $UniformSeq(0, 99, 1000)$, but in this case we apply these values to set the λ_i parameters. However, we have an additional model parameter: λ, that we have to set somehow. It is reasonable to set λ such that the mean of the $\{\lambda_i\}$ parameters approximately equals to the mean of the (MOG approximated) degree distribution of the Poisson model. To achieve this, we used the bisection [15] root-finder to find the appropriate parameter. The fitted parameters are collected in Table 3. We also added the variance of the degree distribution of the Poisson model and the proportional edge voting model with the same $\{D_i\}$ sequence. The degree distribution of the Poisson models with the fitted parameter are plotted on Figure 7. According to the plots and the variance numbers we can conclude that the Poisson voting model and the proportional voting models behave in a very similar way in this setup.

We can do the same, but fitting to the variance. The results of this approach are in Table 4. We can see that changing the parameter λ we can get similar variability as the input $\{\lambda_i\}$ parameters have, but the price is that the mean is shifted.

We can now ask what is the role of the λ parameter in the model. We can investigate the effect if we fix the $\{\lambda_i\}$ parameters and observe how the mean and the variance is changing when we change λ. We expect that λ behaves as the accelerator of the model, as it is increasing, the mean is increasing, however the variance will reach a maximum and then it will fall. Our experiments confirm these expectations. Fixing the $\{\lambda_i\}$ parameters in $LognormalSeq(4, 0.6, 1000)$, we plotted the results in Figure 8.

3.5   Biased voting model

In this variant of the static edge voting model we modify the proportional edge voting model. Here we consider the D_a values as social capital, where the greater D_a means not only more desired degree but also greater influence. This idea is similar to the preferential attachment rule in the Barabasi-Albert model [11]. To be

Table 3

The fitted λ parameter of the Poisson voting model, when the $\{\lambda_i\}$ parameters setted to $ConstSeq(100, 1000)$, $BinomSeq(0.1, 1000)$, $LognormalSeq(4, 0.6, 1000)$ and $UniformSeq(0, 99, 1000)$. The first and the second lines are the mean and the variance of the λ_i parameters. The third line is the fitted λ parameter for each cases. The forth and the fifth lines are the variance of the (approximated) degree distribution in case of the Poisson and the proportional voting model respectively.

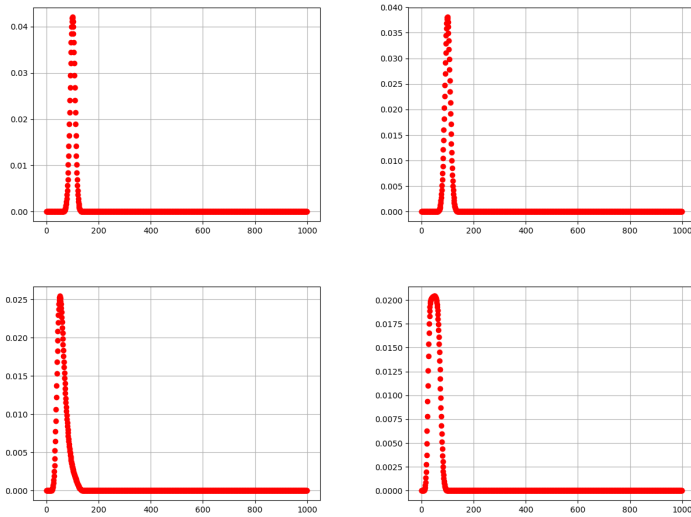|	ConstSeq	BinomSeq	LognormalSeq	UniformSeq
μ_{λ_i}	100	99.88	61.2919	49.5
σ_{λ_i}	0	87.8295	1242.1627	833.25
λ	0.0005280	0.00052802	0.00051923	0.00051572
$\sigma_{d(U)}^{Poi}$	90.1494	109.7385	345.0582	246.0220
$\sigma_{d}^{pro}(U)$	90.0733	111.8468	366.913	254.6096

Figure 7

The degree distribution of the Poisson voting model, when the $\{\lambda_i\}$ parameters are $ConstSeq(100, 1000)$, $BinomSeq(0.1, 1000)$, $LognormalSeq(4, 0.6, 1000)$ and $UniformSeq(0, 99, 1000)$ and the λ parameter of each cases is setted to 0.0005280, 0.00052802, 0.00051923, 0.00051572 respectively

Table 4

The fitted λ parameter of the Poisson voting model, when the λ_i parameters setted to $ConstSeq(100, 1000)$, $BinomSeq(0.1, 1000)$, $LognormalSeq(4, 0.6, 1000$ and $UniformSeq(0, 99, 1000)$. The first line is the variance of the λ_i parameters. The second line is the fitted λ parameter for each cases. The third and the forth lines are the means of the (approximated) degree distribution in case of the Poisson and the proportional voting model respectively.

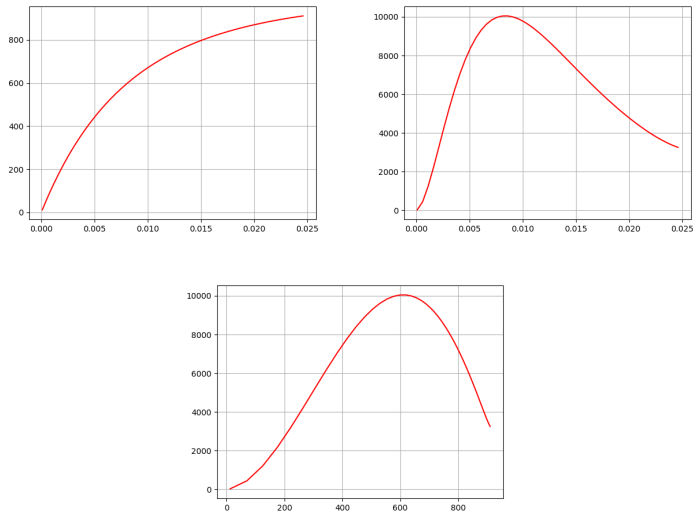|	ConstSeq	BinomSeq	LognormalSeq	UniformSeq
σ_{λ_i}	0	87.8295	1242.1627	833.25
λ	0.05850230	0.00042395	0.00112025	0.00105202
$\mu_{d(U)}^{Poi}$	998.999	81.089	126.781	97.933
$\mu_{d(U)}^{pro}$	100	99.88	61.2919	49.5

Figure 8

Above, the mean and the variance of the Poisson voting model are plotted in the function of parameter λ, when the λ_i parameters are $LognormalSeq(4, 0.6, 1000)$. Below the variance is plotted in the function of the mean

able to formalize this with the edge probability function, first define:

$$v(V_{ab}, V_{ba}) = \frac{D_a}{N-1} V_{ab} + \frac{D_b}{N-1} V_{ba}$$

In the function v the vote of agent a has a weight $\frac{D_a}{N-1}$. Using the v function we can formalize the s function for any fixed $\eta > 0$:

$$s_b(V_{ab}, V_{ba}) = 1 - e^{-\eta v(V_{ab}, V_{ba})}$$

As voting rule we use the Bernoulli voting rule, where V_{ab} does not depend on b and $\mathbb{P}(V_{ab} = 1) = p_i$ if $a \in S_i$

Proposition 3. *In the static edge voting model using the Bernoulli voting rule and the s_b edge probability function, for any $a \in V$ node:*

$$\mathbb{P}[\{a,b\} \in E] = p_{ab}^* + p_{ba}^* - p_{ab}^* p_{ba}^* \tag{10}$$

where:

$$p_{ab} = \mathbb{P}(V_{ab} = 1)$$

$$p_{ab}^* = p_{ab}\left(1 - e^{-\eta \frac{D_a}{N-1}}\right)$$

Proof. Using the definition of s_b and the fact that the expected value of a Bernoulli random variable is its parameter, we have:

$$\mathbb{P}[\{a,b\} \in E] = \mathbb{E}[1 - e^{-\eta v(V_{ab}, V_{ba})}] = 1 - \mathbb{E}[e^{-\eta v(V_{ab}, V_{ba})}]$$

Since the agents can give 0 or 1 vote, we can express the expected value directly using the definition:

$$1 - \mathbb{E}[e^{-\eta v(V_{ab}, V_{ba})}] = 1 - e^0(1 - p_{ab})(1 - p_{ba}) - e^{-\eta \frac{D_a}{N-1}} p_{ab}(1 - p_{ba}) -$$
$$- e^{-\eta \frac{D_{ba}}{N-1}}(1 - p_{ab})p_{ba} - e^{-\eta \frac{D_a + D_b}{N-1}} p_{ab} p_{ba} =$$

After doing the multiplications, rearranging and grouping the appropriate terms, we get the statement:

$$= p_{ab}(1 - e^{-\eta \frac{D_a}{N-1}}) + p_{ba}(1 - e^{-\eta \frac{D_b}{N-1}}) - p_{ab}(1 - e^{-\eta \frac{D_a}{N-1}})p_{ba}(1 - e^{-\eta \frac{D_b}{N-1}}) =$$
$$= p_{ab}^* + p_{ba}^* - p_{ab}^* p_{ba}^*$$

$\square$

Proposition 4. *In the static edge voting model using the Bernoulli voting rule and the s_b edge probability function, for any $a \in V$ node:*

$$\mathbb{E}[d_a] = \sum_{b \in V_a} (p_{ab}^* + p_{ba}^* - p_{ab}^* p_{ba}^*)$$

where:

$$p_{ab} = \mathbb{P}(V_{ab} = 1)$$

$$p_{ab}^* = p_{ab} \left(1 - e^{-\eta \frac{D_a}{N-1}} \right)$$

Proof. Since the degree of a node is the number of its neighbours in the graph, and the expected value of a Bernoulli random variable is its parameter:

$$\mathbb{E}[d_a] = \sum_{b \in V_a} \mathbb{P}[\{a,b\} \in E]$$

Applying Proposition 3 we get the statement.   □

Using equation (10) we can construct the equivalent IRG model. As before in the case of the Poisson model, we would like to compare the model to this proportional model and to the Poisson model. In order to do this we use the same sequences to set the $\{D_i\}$ parameters of the model: $ConstSeq(100,1000)$, $BinomSeq(0.1,1000)$, $LognormalSeq(4,0.6,1000)$ and $UniformSeq(0,99,1000)$. We use the proportional voting rule: $p_i = \frac{D_i}{N-1}$ but we also have to set the η parameter. We do the same as before: set the η parameter such that the mean of the $\{D_i\}$ parameters approximately equals to the mean of the (MOG approximated) degree distribution. To achieve this, we used the bisection [15] root-finder again. The fitted parameters are collected in Table 5. and the corresponding degree distributions are on Figure 9. As in the case of the Poission model, we also fitted the η parameters to hit the variance of the $\{D_i\}$ parameter set. The results of this method are in Table 6. When we fitted the η parameter to hit the mean of the input $\{D_i\}$ sequence, we can see that we can achieve greater variance in the degree distribution than in the case of the other two models. While we fitted to hit the variance, we can now see that the shift of the mean is not so high then in case of the Poisson model.

Now we investigate the role of the η parameter with numerical experiments. As before, fix the $\{D_i\}$ parameters to $LognormalSeq(4,0.6,1000)$ and compute the mean and the variance of the approximated degree distribution for each $\eta \in 0,1,...,100$ The results are plotted in Figure 10.

4   Discussion

In this paper we continued and extended our preliminary work [12] about static edge voting models. In Section 2 started with reviewing the properties of inhomogeneous random graphs (IRG), because IRGs are our basic tool, and proved a corollary of the Chernoff-bounds what we can use to bound the probability that the degree of a fixed node is below or above of a given D value. In Section 3 we gave

Table 5

The fitted η parameter of biased model, when the D_i parameters setted to $ConstSeq(100,1000)$, $BinomSeq(0.1,1000)$, $LognormalSeq(4,0.6,1000$ and $UniformSeq(0,99,1000)$. The first and the second lines are the mean and the variance of the D_i parameters. The third line is the fitted η parameter for each case. The next lines are the variance of the (approximated) degree distribution in case of the biased, the proportional voting model and the Poisson model respectively.

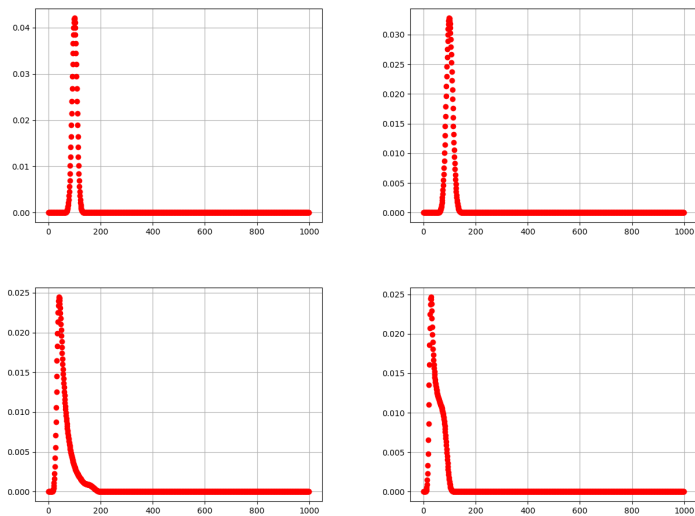|	ConstSeq	BinomSeq	LognormalSeq	Uniform
μ_{D_i}	100	99.88	61.2919	49.5
σ_{D_i}	0	87.8295	1242.1627	833.25
η	7.1914	7.1591	9.5026	11.1728
$\sigma_{d(U)}^b$	90.0733	148.266	884.0557	444.9441
$\sigma_d^{pro}(U)$	90.0733	111.8468	366.913	254.6096
$\sigma_{d(U)}^{Poi}$	90.1494	109.7385	345.0582	246.0220

Figure 9

The degree distribution of biased voting model, when the D_i parameters are $ConstSeq(100,1000)$, $BinomSeq(0.1,1000)$, $LognormalSeq(4,0.6,1000$ and $UniformSeq(0,99,1000)$ and the η parameter of each cases is setted to 7.1914, 7.1591, 9.5026, 11.1728 respectively

Table 6

The fitted η parameter of the biased voting model, when the D_i parameters setted to $ConstSeq(100,1000)$, $BinomSeq(0.1,1000)$, $LognormalSeq(4,0.6,1000)$ and $UniformSeq(0,99,1000)$. The first line is the variance of the λ_i parameters. The second line is the fitted η parameter for each case. The next lines are the means of the (approximated) degree distribution in case of the biased, the proportional voting model and the Poission model respectively.

	ConstSeq	BinomSeq	LognormalSeq	UniformSeq
σ_{D_i}	0	87.8295	1242.1627	833.25
η	0	3.8873	16.348	26.3628
$\mu_{d(U)}^b$	0	63.654	82.00028	76.305
$\mu_{d(U)}^{pro}$	100	99.88	61.2919	49.5
$\mu_{d(U)}^{Poi}$	998.999	81.089	126.781	97.933

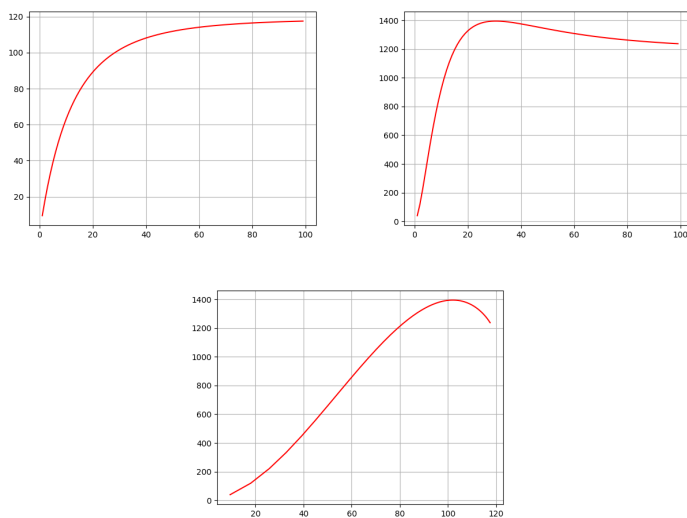Figure 10

Above, the mean and the variance of the biased voting model are plotted in the function of parameter η, when the D_i parameters are $LognormalSeq(4,0.6,1000)$ Below the variance is plotted in the function of the mean

the general definition of the static edge voting models, and introduced the method of Mixture of Gaussians to be able to estimate the degree distribution of the model. We dealt with three special cases: the proportional edge voting model, the Poisson model and the biased voting model. In case of the proportional model any a agent votes for the $\{a,b\}$ edge candidate with probability proportional to its D_a desired degree, and the probability that $\{a,b\}$ becomes an edge is just the average of the incoming votes. We have seen, that the proportional model has a smoothing effect, the expected degree of a node a is approximately the average of the D_a desired degree of a and the $\overline{D}$ mean desired degree. In the Poisson model the agents can give more votes to an edge candidate, and the distribution of the votes given by agent a is Poisson(λ_i) if $a \in S_i$. The edge probability function here is $s(x) = 1 - e^{-\lambda x}$ where λ is a control parameter. The input parameters are in this model the $\{\lambda_a | a \in V\}$ parameters of the Poisson distributions and the λ control parameter. We have seen, that if we set λ such that the mean of the degrees is equal to the mean of $\{\lambda_a | a \in V\}$, then the Poisson edge voting model behaves like the proportional model, however we can get similar variance in the degrees like in the $\{\lambda_a | a \in V\}$ input sequence at the price that the mean of the degree distribution is shifting. The biased voting model is the modification of the proportional model in a way that the votes of agent a is weighted by $\frac{D_a}{N-1}$. In this model D_a can be considered as social capital. This biased behaviour is enforced by the edge probability function: $s_b(V_{ab}, V_{ba}) = 1 - \exp\left(-\eta\left(\frac{D_a}{N-1}V_{ab} + \frac{D_b}{N-1}V_{ba}\right)\right)$. We have seen that we can achieve greater variability in degree as in case of the proportional and the Poission models. Lets do now a final comparison as a summary. Fix again the input parameters of all the three models: in case of the proportional and biased models set the $\{D_i | i \in V\}$ sequences to *LognormalSeq*(4, 0.6, 1000), and in case of the Poisson model set $\{\lambda_i | i \in V\}$ to *LognormalSeq*(4, 0.6, 1000) again. Choose the control parameters λ and η in case of the Poisson and the biased models forcing the mean of the degree distribution equal to the mean of the input *LognormalSeq*(4, 0.6, 1000) sequence. Now compute the expected degree of the nodes for a reference node for each non-empty S_i agent group, and plot $\mathbb{E}[d_a | a \in S_i]$ in the function of the parameter (D_i or λ_i). The results are in Figure 11. We can see that the proportional and the Poission models are behaving almost identically, and we can experience the "smoothing" behaviour in both cases. However, in case of the biased model after a short initial period the curve of the expected degrees becomes close to the identical function, which shows that the biased model more-or-less preserves the variability of the input parameter set in the degree distribution.

We note that the SEV model better describes the process in which agents explore their environment in the social space, but the survival of relationships are affected by other factors. A natural extension is to make the model temporal, and the previous note makes sense only in temporal networks: agents first explore their environment by making connections, but maintaining these connections is a different process. Finally, here we supposed a small population, where every agent can link to any other agent. However, in larger populations this is not true, so we have to include the distance into our framework in some way.
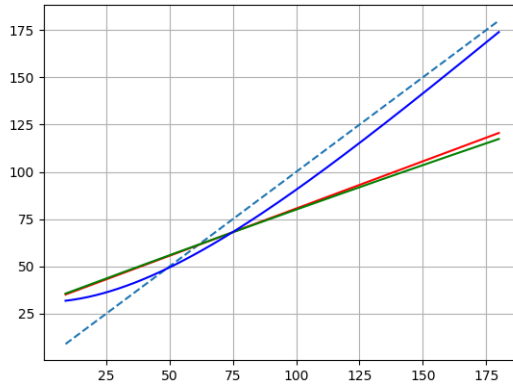
Figure 11
Comparison of the expected degrees of the nodes in the function of the model parameters. The
horizontal axis is the parameter and the vertical axes is the expected degree of the node in the function
of the parameter. The proportional, the Poission and the biased models are represented by the red, green
and blue lines respectively. The blue dashed line is the identical function.

References

[1]   Anna Zafeiris and Tamás Vicsek. *Why We Live in Hierarchies?: A Quantitative Treatise*. Springer, 2017.

[2]   Albert-László Barabási et al. *Network science*. Cambridge university press, 2016.

[3]   Everett M Rogers. Diffusion of innovations. free press. *New York*, 551, 2003.

[4]   Margarita Vázquez-Brage, Ignacio Garcıa-Jurado, and Francesc Carreras. The owen value applied to games with graph-restricted communication. *Games and Economic Behavior*, 12(1):42–53, 1996.

[5]   Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[6]   A Barrat, M Barthélemy, and A Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2012.

[7]   Tamás Péter and Krisztián Szabó. Combined mathematical modeling of different transport networks, considerations and complex analysis. *Acta Polytechnica Hungarica*, 14(2):7–26, 2017.

[8]   Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[9]   Remco Van Der Hofstad. *Random graphs and complex networks*, volume 1. Cambridge University Press, 2016.

[10] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

[11] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[12] Róbert Pethes and Levente Kovács. Static edge voting models. In *IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI 2018) Budapest, Hungary*, pages 223–228. IEEE, 2018.

[13] Bent Fuglede and Flemming Topsoe. Jensen-shannon divergence and hilbert space embedding. In *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.*, page 31. IEEE, 2004.

[14] Hans Fischer. *A history of the central limit theorem: From classical to modern probability theory*. Springer Science & Business Media, 2010.

[15] D Vaughan Griffiths and Ian Moffat Smith. *Numerical methods for engineers*. Chapman and Hall/CRC, 2006.