The Random Triangle Model

Johan Jonasson
Chalmers University of Technology

January 19, 1998

Abstract

The random triangle model is a Markov random graph model which, for parameters $p \in (0,1)$ and $q \geq 1$ and a graph $G = (V, E)$, it assigns to a subset, η, of E, a probability which is proportional to $p^{|\eta|}(1-p)^{|E|-|\eta|}q^{t(\eta)}$, where $t(\eta)$ is the number of triangles in η. It is shown that this model has maximum entropy in the class of distributions with the given edge and triangle probabilities.

It is proved that a modification of the Swendsen-Wang correspondence between the Fortuin-Kasteleyn random cluster model and the Potts model is valid for the random triangle model, i.e. it corresponds to a certain distribution of "spins" on the set of all triangles of G. Using this correspondence, the asymptotic behavior of the random triangle model on the complete graph is examined for p of order $n^{-\alpha}$, $\alpha > 0$, and different values of q, where q is written on the form $q = 1 + h(n)/n$. It is shown that the model exhibits an explosive behavior in the sense that if $h(n) \leq c \log n$ for $c < 3\alpha$, then the edge probability and the triangle probability are asymptotically the same as for the ordinary $G(n, p)$ model, whereas if $h(n) \geq c' \log n$ for $c' > 3\alpha$, then these entities both tend to 1. For critical values, $h(n) = 3\alpha \log n + o(\log n)$, the probability mass divides betweeen these two extremes.

Moreover, if $h(n)$ is of higher order than $\log n$, then the probability that $\eta = E$ tends to 1, whereas if $h(n) = o(\log n)$ and $\alpha > 2/3$, then, with a probability tending to 1, the resulting graph can be coupled with a graph resulting from the $G(n, p)$ model. In particular these facts mean that for values of p in the range critical for the appearance of the giant component and the connectivity of the graph, the way in which triangles are rewarded can only have a degenerate influence.

1 Introduction

Since Erdős and Rényi introduced the subject in 1959 (see [5]), random graphs have attracted much attention in various disciplines of science. One class of such disciplines are the social sciences where random graph models are used for describing the structure of social networks. See e.g. the book by Faust and Wasserman [6] or the survey paper by Frank [8]. For instance, an edge between two vertices could mean that the corresponding individuals are friends. It is this example that provides the background for this paper.

It is a well recognized property of such "friendship graphs" that they exhibit so called *transitivity*, i.e. that "a friend of a friend of mine is often also a friend of mine".

Therefore one wants to study random graph models which, unlike the simplest model with edges present or absent independently of each other, reward transitivity. In terms of the graph structure rewarding transitivity means that we must create a model where the probability that a certain fixed edge is present given the absence or presence of all other edges is larger if its end vertices are second neighbours than if they are not. There is a number of different random graph models for which this holds. Two examples of such models which have been used in a social network context are random hypergraph models (see [12] and the references therein) and the random intersection graph model (see [13]). Both these examples, however, exhibit some undesirable properties in the conditional probability that a certain edge is on given the rest of the graph.

So what *is* the "most natural" prototype for a random graph model with transitivity? If we drop our desire for transitivity and just want a model where the probability for edge presence is p, for some $p \in [0,1]$, then the the model with independent edges seems to be most natural in some sense. Likewise, if we want a model with a fixed number of edges, we let these spread out uniformly. These two models have in common that they are in some sense the most random ones under the corresponding conditions. Mathematically speaking they maximize the *entropy* in the corresponding classes of random graph distributions, i.e. they maximize the function $-\sum_s \mu(s)\log\mu(s)$ where μ ranges over the possible random graph distributions and the sum is over the possible outcomes.

In the present situtation it is therefore natural to look for a maximum entropy model with the property that the probability that a given triangle is present takes on some desired value which is strictly higher than the product of the three corresponding edge presence probabilities. (We define a triangle as a set of three edges such that each two of these have a common end vertex.) It turns out that the model we are then looking for is the following.

DEFINITION. Let $G=(V,E)$ be a finite graph. The *random triangle measure* on $\mathcal{P}(E)$ is given by

$$\mu_{t,G}^{p,q}(\eta) = \frac{1}{Z_{t,G}^{p,q}} p^{|\eta|}(1-p)^{|E|-|\eta|} q^{t(\eta)} \tag{1}$$

where $t(\eta)$ is the number of triangles in η and $Z_{t,G}^{p,q}$ is a normalizing constant. We require that $p \in [0,1]$ and $q \geq 1$.

Consider for a moment the case where G is the complete graph and let A_e be the event that $\{e \in Y\}$ where Y is understood to be a random graph chosen according to $\mu_{t,G}^{p,q}$ and e is some given edge, and let A_t be the corresponding event for some given triangle t. Since the probabilities $\mu_{t,G}^{p,q}(A_e)$ and $\mu_{t,G}^{p,q}(A_t)$ are continuous functions of p and q it is easily seen that for any a and b such that $a \in [0,1]$ and $a^3 \leq b < a$, we can find p and q in (1) such that $\mu_{t,G}^{p,q}(A_e) = a$ and $\mu_{t,G}^{p,q}(A_t) = b$. In fact this property holds for all transitive graphs, G, e.g. a finite part of a d-dimensional triangular lattice with the proper torus convention.

The desired property of maximum entropy is the result of the following proposition.

PROPOSITION 1.1 *The entropy of $\mu_{t,G}^{p,q}$ is maximal in the class of probability measures, μ, on $\mathcal{P}(E)$ with $\mu(A_e) = \mu_{t,G}^{p,q}(A_e)$ and $\mu(A_t) = \mu_{t,G}^{p,q}(A_t)$ for all $e \in E$ and all $t \in T$.*

Proof. Regard this as a sampling situation where the set of "individuals" is the set of all edges and triangles of G. Choose a sample from this set of individuals by picking each edge with probability p and each triangle with probability $q/(q+1)$ independently for different individuals. Denote this sample X. It is clear that X has maximum entropy in the class of sample distributions on this set of individuals with the given inclusion probabilities.

Now, by conditioning on the event that the triangles resulting from our choice of edges exactly coincides with the triangles that were actually chosen in the sample, we get the random triangle measure. Let Y be the indicator of this event and use the standard entropy formula:

$$H(X,Y) = H(Y) + H(X|Y=0)Pr(Y=0) + H(X|Y=1)Pr(Y=1)$$

where H denotes entropy. Since Y is $\sigma(X)$-measurable we have that $H(X,Y) = H(X)$ and since $H(X)$ is maximal, changing the distribution of X on $\{Y=1\}$ without changing the edge and triangle probabilities cannot increase $H(X)$. The result follows. □

Now imagine for a while that we replace $t(\eta)$ of the definition (1) with $k(\eta)$, the number of connected components of the graph. Then we get the so called *random cluster model*. This is a well known model from statistical mechanics and some of the methods used to ananlyze it will turn out helpful for us, so let us take a closer look at it.

2 Preliminaries on the Random Cluster Model

The random cluster model can be seen as a third example of a random graph model which has an element of transitivity but which shares the drawback of the random intersection model mentioned above. The precise definition is the following.

Definition. Let $G = (V, E)$ be a finite graph. The random cluster measure on $\mathcal{P}(E)$ is given by

$$\mu_G^{p,q}(\eta) = \frac{1}{Z_G^{p,q}} p^{|\eta|}(1-p)^{|E|-|\eta|} q^{k(\eta)}$$

where $k(\eta)$ is the number of connected components in η. The numbers p and q are the parameters of the model and are to be chosen so that $p \in [0,1]$ and $q > 0$. The number $Z_G^{p,q}$ is a normalizing constant.

Note that letting $q = 1$ yields the ordinary model with independent edges. As mentioned, the random cluster model arose in a completely different context than social networks. It was introduced in the 70's by Fortuin and Kasteleyn [7] because of its correspondence, for $q = 2, 3, \ldots$, with a q-state *Potts model*, which for $q = 2$ is the *Ising model*, which is in turn a mathematical model of a ferromagnet at a microscopic level. For an introduction to the use of the random cluster model in this context we refer to the survey paper by Häggström [9]. The Potts model and the correspondence will be described below.

The transitivity of the random cluster model is captured by the following proposition.

PROPOSITION 2.1 *Fix an arbitrary edge $e \in E$ and let $Y_e = Y \cap (E \setminus \{e\})$ for a random subset, Y, of E, distributed according to the random cluster measure. Then*

$$\mu_G^{p,q}(e \in Y | Y_e) = \begin{cases} p & \text{if the end vertices of } e \text{ are connected in } Y_e \\ \frac{p}{p+(1-p)q} & \text{otherwise.} \end{cases}$$

By inspecting this expression one can see that the probability of having an edge between two vertices is larger if they are already connected in some way than if they are not, so the model indeed exhibits some transitivity.

The random cluster model clearly gives an intricate dependence between the edges. This makes it hard to analyze directly, but by using the correspondence with the Potts model one can often get around this difficulty.

DEFINITION. Let $q \in \{2, 3, \ldots\}$ and let $G = (V, E)$ be a finite graph. The q-state Potts measure on $\{1, \ldots, q\}^V$ is given by

$$\nu_G^{p,q}(\xi) = \frac{1}{U_G^{p,q}}(1-p)^{d(\xi)}$$

where $d(\xi)$ is the number of neighbours, u and v, such that $\xi(u) \neq \xi(v)$ and $U_G^{p,q}$ is a normalizing constant.

Usually one replaces the parameter p by another parameter β given by $1 - p = e^{-2\beta}$, where β is called the reciprocal temperature of the model. The value $\xi(u)$ is called the *spin* of the vertex u. The correspondence between this model and the random cluster model is captured by the following coupling, which was explicitly introduced by Edwards and Sokal [4] but was earlier implicitly introduced by Swendsen and Wang [15].

PROPOSITION 2.2 *Define a probability measure on $\mathcal{P}(E) \times \{1, \ldots, q\}^V$ by*

$$P_G^{p,q}(\eta, \xi) = \frac{1}{Z} p^{|\eta|}(1-p)^{|E|-|\eta|} I_A(\eta, \xi)$$

where A is the set of outcomes such that $\xi(u) \neq \xi(v)$ implies that u and v belong to different connected components in η. Then the marginal distributions on $\mathcal{P}(E)$ and $\{1, \ldots, q\}^V$ are $\mu_G^{p,q}$ and $\nu_G^{p,q}$ respectively.

In words, the measure $P_G^{p,q}$ is given by letting the edges be present with probability p independently and letting the vertices have spins chosen by uniform distribution on $\{1, \ldots, q\}$ independently and then conditioning on that no two vertices with different spins have an edge between them. A proof of Proposition 2.2 can be found e.g. in [9, page 6]. An immediate consequence of this result is that if one generates a subgraph of G according to the random cluster measure and then assigns all vertices in the same connected component the same uniformly chosen spin, independently for the different connected components, the resulting spin configuration is distributed according to the Potts measure. Vice versa, if one generates a spin configuration according to the Potts measure and then puts edges only between neighbours with the same spin with probability p independently of each other, the resulting graph is distributed according to the random cluster measure. The last fact means that given the spin configuration, the random cluster graph behaves like q disjoint graphs where the edges are present or absent independently, a fact which is of course helpful when studying its structure. This technique rests on the fact that q is integer valued, but it can be generalized. Such a generalized technique is used in the paper by Bollobas, Grimmett and Janson [3], where the structure of the random random cluster model on the complete graph is studied for p of order n^{-1}.

Before moving on, let us also note that the random cluster measure can be obtained as the stationary distribution of a certain reversible Markov chain, a so called *Gibbs sampler*. The transitions of this Markov chain are such that at each time point one chooses an edge $e \in E$ at random. Then, regardless of whether or not this edge is present for the moment, one lets it be present with probabilty p if its end vertices are connected through the rest of the graph and with probability $p/(p+q(1-p))$ otherwise. There will be a corresponding Gibbs sampler for the random triangle model.

3 General Properties of the Random Triangle Model

We saw in the introduction that the random triangle model maximizes the entropy in the class of distributions with the given edge and triangle probabilities. In this section we give a few other general properties of the model and devote the next section to the complete graph case.

Note that, as opposed to for the random cluster model, the nature of definition (1) is such that the presence or absence of an edge will only depend on edges adjacent to it. The model is therefore of interest not only as a model for social networks but also as a proptotype for a "random cluster"-like model with only local edge dependencies.

The transitivity of the model is described by the following proposition.

PROPOSITION 3.1 *Let Y be a random subset of E distributed according to the random triangle measure and fix an edge $e \in E$. Let $\Delta(Y, e)$ be the number of triangles in $Y \cup \{e\}$ of which e is a part. Then*

$$\mu_{t,G}^{p,q}(e \in Y | Y_e) = \frac{pq^{\Delta(Y,e)}}{pq^{\Delta(Y,e)} + 1 - p}$$

where $Y_e = Y \cap (E \setminus \{e\})$ as in Proposition 2.1.

The proof is just definition chasing. In the light of Proposition 3.1 we can construct the Gibbs sampler corresponding to the random triangle model. This is analogous to the construction of the Gibbs sampler for the random cluster model. We consider a Markov chain on $\mathcal{P}(E)$ where the transitions are done by chosing an edge $e \in E$ at random an then, if η is the present state, letting the chosen edge be present with probability $pq^{\Delta(\eta,e)}/(pq^{\Delta(\eta,e)} + 1 - p)$. This Markov chain describes in a nice way how a random triangle graph emerges dynamically in a population with fixed individuals and also suggests a method for simulation of the model.

Just as for the random cluster model, the intricate dependence between edges makes direct analysis hard. Therefore it would be desirable to find a valid modification of the correspondence between the random cluster model and the Potts model. We will do this by assigning spins to the *triangles* of G rather than to the vertices. (The term "spin" is of course not really relevant here, but since the concept is mathematically essentially the same thing, we allow ourselves this small sin.) To start with, we let the spins take on any value in the interval $[0, q]$ independently for different triangles. Then we let the edges of E be present with probabaility p

independently and finally we condition on that all edges that are part of a triangle with spin larger than 1 are present. This procedure yields a probability measure on $\mathcal{P}(E) \times [0,q]^T$, where T is the set of triangles of G, corresponding to the coupling of Proposition 2.2. The marginal distributions are the random triangle measure and a measure on $[0,q]^T$ corresponding to the Potts measure:

THEOREM 3.2 *Let T be the set of triangles of G and define the probability measure $P^{p,q}_{t,G}$ on $\mathcal{P}(E) \times [0,q]^T$ by letting $P^{p,q}_{t,G}$ have density (w.r.t. the product of counting measure and Lebesgue measure)*

$$h^{p,q}_{t,G}(\eta,\xi) = \frac{1}{Z_t} p^{|\eta|}(1-p)^{|E|-|\eta|} I_B(\eta,\xi) \tag{2}$$

where Z_t is a normalizing constant and B is the set of outcomes such that all edges, e, that are part of a triangle $t \in T$ such that $\xi(t) \geq 1$ satisfies $e \in \eta$. Then the marginal distribution of $P^{p,q}_{t,G}$ on $\mathcal{P}(E)$ is $\mu^{p,q}_{t,G}$ and the marginal distribution on $[0,q]^T$ is given by the density

$$\pi^{p,q}_{t,G}(\xi) = \frac{1}{Z_t} p^{f(\xi)}$$

where $f(\xi)$ is the number of edges that are part of some triangle $t \in T$ such that $\xi(t) \geq 1$.

Proof. For the first part we integrate out ξ:

$$\int_{\xi \in [0,q]^T} h^{p,q}_{t,G}(\eta,\xi) d\xi = \frac{1}{Z_t} p^{|\eta|}(1-p)^{|E|-|\eta|} \int_{\xi:(\eta,\xi)\in B} d\xi$$

$$= \frac{1}{Z_t} p^{|\eta|}(1-p)^{|E|-|\eta|} q^{t(\eta)}.$$

For the second part we sum out η:

$$\sum_\eta h^{p,q}_{t,G}(\eta,\xi) = \frac{1}{Z_t} \sum_{\eta:(\eta,\xi)\in B} p^{|\eta|}(1-p)^{|E|-|\eta|}$$

$$= \frac{1}{Z_t} p^{f(\xi)}$$

as the possible η's must contain all edges that are part of a triangle, t, in G with $\xi(t) \geq 1$. $\square$

Letting the spins have a continuous range was comfortable for the proof, but it is not comfortable when working with the model. Therefore we will in the sequel not use the measure $\pi^{p,q}_{t,G}$ of Theorem 3.2 but rather the equivalent measure $\nu^{p,q}_{t,G}$ on $\{0,1\}^T$ given by

$$\nu^{p,q}_{t,G}(\xi) = \frac{1}{Z_t}(q-1)^{|\xi|} p^{f(\xi)}$$

where $|\xi| = \sum_{t\in T} \xi(t)$ is the number of triangles in G with spin 1 and $f(\xi)$ is the number of edges that are part of a triangle with spin 1. The following important corollary is an immediate consequence of Theorem 3.2.

COROLLARY 3.3 *(a) Choose a spin configuration $\xi \in \{0,1\}^T$ by*

(i) Choosing $\eta \in \mathcal{P}(E)$ according to $\mu^{p,q}_{t,G}$.

(ii) Given η, choosing spins for the triangles $t \in T$ independently so that $\xi(t) = 0$ with probability $1/q$ if all three edges of t are present in η and with probability 1 otherwise.

Then ξ has distribution $\nu^{p,q}_{t,G}$.

(b) Choose a subset $\eta \in \mathcal{P}(E)$ by

(i) Choosing $\xi \in \{0,1\}^T$ according to $\nu^{p,q}_{t,G}$.

(ii) Given ξ, letting the edges, $e \in E$, be present or absent in η independently of each other in such a way that e is present in η with probability 1 if e is part of a triangle $t \in T$ with $\xi(t) = 1$ and with probability p otherwise.

Then η has distribution $\mu^{p,q}_{t,G}$.

4 Asymptotics on the Complete Graph

In this section we study the asymptotics for the random triangle model on the complete graph. This subject has been briefly treated before in [14]. The subscript G of the μ's and the ν's will from now on be dropped. The principal question here is how the way in which we reward triangles can affect the structure of the graph compared to an ordinary $G(n,p)$ graph when the number of vertices is large. The $G(n,p)$ model is the random graph model with independent edges on the complete graph with edge probability p. This model has been thoroughly analyzed from various points of view. The interested reader is urged to dig into Bollobas' [2] book on the subject.

Before starting, let us recall some order notation. Let $\{f(n)\}_{n=1}^{\infty}$ and $\{g(n)\}_{n=1}^{\infty}$ be two sequences of nonnegative real numbers. We write $g(n) = o(f(n))$ if $g(n)/f(n) \to 0$ as $n \to \infty$. We write $g(n) = O(f(n))$ if $\{g(n)/f(n)\}$ is bounded. In case $g(n)$ is $O(f(n))$ but not $o(f(n))$ we write $g(n) = \Theta(f(n))$. Finally, $g(n) = \Omega(f(n))$ will mean that $f(n) = O(g(n))$ and $g(n) = \omega(f(n))$ will mean that $f(n) = o(g(n))$.

Let us begin with asking ourselves what a reasonable value for q could be. Heuristically, we would like each person to have about a constant number of friends so that p should be of order n^{-1}. However, given that two persons have a common friend there should be a constant probability that these two persons are also friends. This implies that we should have $q = \Theta(n)$. On the other hand, mathematical heuristics tell us that since the total number of triangles in the complete graph is $\Theta(n^3)$ whereas the total number of edges is $\Theta(n^2)$, we should have $q = 1 + \Theta(n^{-1})$ or something close to that to avoid that the large number of triangles dominates everything else and causes the outcome of the model to be degenerate. After all, the complete graph provides no geographical restrictions, so it might be reasonable that the tendency to make friends with the friends of your friends decreases as the population increases in such an ideal world.

Obviously, it is not at all evident what q should be. Fortunately the following result rules in favor of mathematical heuristics and radically limits the possibilities. To make life a bit easier we will for most of this section assume that p is at most of order $n^{-\alpha}$ for some $\alpha > 0$ and at least of order n^{-2}. This covers most of the interesting range for p for the $G(n,p)$ model; it covers the appearance of the first few edges and goes well beyond the critical point for connectivity.

THEOREM 4.1 *Assume that* $p = p(n) = \Omega(n^{-2})$ *and let* $q = q(n) = 1 + h(n)/n$ *where* $h(n) = \omega(\log n)$. *Then*

$$\mu_t^{p,q}(E) \to 1$$

as $n \to \infty$.

Proof. This will follow from direct inspection of the model along with some combinatorial observations.

Suppose that $\eta \in \mathcal{P}(E)$ has at most $\binom{n-k}{2}$ edges, $k = 1, \dots, n/10$. These edges make at most $\binom{n-k}{3}$ triangles so that

$$\frac{\mu_t^{p,q}(\eta)}{\mu_t^{p,q}(E)} \leq \frac{p^{\binom{n-k}{2}} q^{\binom{n-k}{3}}}{p^{\binom{n}{2}} q^{\binom{n}{3}}} = p^{\binom{n-k}{2}-\binom{n}{2}} q^{\binom{n-k}{3}-\binom{n}{3}}. \tag{3}$$

From ordinary algebra it follows that

$$\binom{n}{2} - \binom{n-k}{2} = \binom{n}{2} - \frac{(n-k)(n-1-k)}{2} = nk - \frac{1}{2}(k-k^2) \leq nk$$

and that

$$\binom{n}{3} - \binom{n-k}{3} = \binom{n}{3} - \frac{(n-k)(n-1-k)(n-2-k)}{6} \geq \frac{n^2k}{4}$$

where the last inequality uses the assumption that $k \leq n/10$. Inserting into (3) yields

$$\frac{\mu_t^{p,q}(\eta)}{\mu_t^{p,q}(E)} \leq p^{-nk} q^{-n^2k/4} = \frac{p^{-nk}}{(1+\frac{h(n)}{n})^{n^2k/4}} = p^{-nk} e^{-nk\omega(\log n)}$$

for n large. Since p is at least at least of order n^{-2}, p^{-nk} is at most of order $e^{2nk\log n + O(n)}$ so that the the ratio $\mu_t^{p,q}(\eta)/\mu_t^{p,q}(E)$ has order at most $e^{-nk\omega(\log n)}$.

Now let us check for what integer values, x, it is true that $\binom{n-k-1}{2} \leq \binom{n}{2} - x \leq \binom{n-k}{2}$. We get

$$nk - \frac{1}{2}k(k+1) \leq x < n(k+1) - \frac{1}{2}(k+1)(k+2) \tag{4}$$

and we see that this is valid for no more than n different values of x. Since these x's satisfy $x < n(k+1)$, the total number of outcomes corresponding to each k is thus bounded by

$$n\binom{\binom{n}{2}}{n(k+1)} \leq n^{1+2n(k+1)} \leq n^{5nk} = e^{nkO(\log n)}.$$

Combining our results we get

$$\frac{\mu_t^{p,q}(\binom{n-k-1}{2} \leq |\eta| \leq \binom{n-k}{2})}{\mu_t^{p,q}(E)} = e^{nkO(\log n) - nk\omega(\log n)} = e^{-nk\omega(\log n)}$$

$$= o(n-1)$$

for all $k = 1, \dots, n/10$. By summation it follows that

$$\frac{\mu_t^{p,q}(\binom{0.9n}{2} \leq |\eta| \leq \binom{n-1}{2})}{\mu_t^{p,q}(E)} = o(1).$$

That the probability for an outcome with less than $\binom{0.9n}{2}$ edges has probability $o(1)$ follows from repeating the arguments in a simplified form noting that for such outcomes there can be at most $\binom{0.9n}{3} < 0.8\binom{n}{3}$ triangles. Finally it has to be checked that the probability for an outcome, η, with $\binom{n}{2} - n + 2 \leq |\eta| \leq \binom{n}{2} - 1$ edges is vanishing. This, however, follows readily from direct inspection of the model. □

Observe that the proof holds for the extreme cases $p = o(n^{-2})$ as well if we put $h(n) = \omega(\log p^{-1}(n))$ unless p is of such low orders as $o(e^{-n})$. (It is, however, not enough to put $h(n) = \omega(\log p^{-1}(n))$ if $p(n)$ is of order close to 1.) For these extreme cases it can also be shown that if $h(n) = o(\log p^{-1}(n))$, then $\mu_t^{p,q}(\emptyset) \to 1$ whereas if $h(n) = \Theta(\log p^{-1}(n))$, then we end up with a distribution which asymptotically divides its mass between $\emptyset$ and E.

Now, let us instead turn to the nontrivial cases. As mentioned earlier, central questions about the structure of a random graph are questions about connectivity and about the appearance of a giant component and also about the nature of the separate components. For the $G(n,p)$ model it is well known that $p = \log n/n$ is a threshold for connectivity of the graph. It is also well known that if $p = b/n$ for a constant, b, then the model can behave in three completely different ways depending on whether $b < 1$, $b = 1$ or $b > 1$. In the case $b < 1$, the graph will with probability tending to 1 consist only of trees and unicyclic components of order at most $\log n$. If $b = 1$, a giant component of order $n^{2/3}$ will appear whereas the other components are still trees or unicyclic and of order at most $\log n$. In the case $b > 1$, the giant component will be of order n, while the remaining components are still trees or unicyclic of the same order as in the other cases. These results were first established by Erdős and Rényi [5] and have later been studied in more detail by several authors. For the interested reader we refer to [2] and for a deeper treatment to the paper by Janson et. al. [11] where a detailed treatment of how the giant component emerges as the edges are added one by one is given.

Now, in what way can our way of rewarding triangles affect these properties? By Theorem 4.1, a random triangle graph shows no similarity whatsoever to a $G(n,p)$ graph for $q = 1 + \omega(\log n)/n$. The following result shows that if $q = 1 + o(\log n)/n$ the situation is quite the opposite, in particular there is no effect *at all* in the critical range for these properties. In fact, it is an immediate consequence of this result and Corollary 3.3 that with probability tending to 1, the random triangle model behaves *exactly* like the $G(n,p)$ model for such p and q.

THEOREM 4.2 *Assume that $q = q(n) = 1 + h(n)/n$ where $h(n) = o(\log n)$ and let $p = p(n) = O(n^{-\alpha})$ for some $\alpha > 2/3$. Then*

$$\nu_t^{p,q}(|\xi| = 0) \to 1$$

as $n \to \infty$.

Proof. We are going to use Corollary 3.3 by first choosing η_0 according to $\mu_t^{p,q}$ and then jump between $\mathcal{P}(E)$ and $\{0,1\}^T$ as indicated by Corollary 3.3. This will generate random elements $\xi_1, \eta_1, \xi_2, \eta_2, \ldots$ such that the η_i's all have distribution $\mu_t^{p,q}$ and the ξ_i's all have distribution $\nu_t^{p,q}$.

First, however, let r be the probability that a fixed edge, $e \in E$, is present in η_0, i.e.

$$r \equiv \mu_t^{p,q}\{e \text{ is present}\}.$$

Then, as e is part of $n-2$ different triangles in G,

$$p \le r \le \frac{p(1+\frac{h(n)}{n})^{n-2}}{p(1+\frac{h(n)}{n})^{n-2}+1-p} \le pe^{o(\log n)} = O(n^{-\alpha})O(n^{\delta}) = O(n^{-(\alpha-\delta)})$$

for arbitrary $\delta > 0$. Choose δ small enough to ensure that $\beta \equiv \alpha - \delta > 2/3$. With this β we have

$$\mathbf{E}|\eta_0| = \binom{n}{2} r = O(n^{2-\beta}).$$

Fix a small $\epsilon > 0$. By Markov's inequality we can pick a constant, $A < \infty$, such that

$$\mu_t^{p,q}(|\eta_0| \ge An^{2-\beta}) \le \epsilon.$$

Since $An^{2-\beta}$ edges make at most $Bn^{3-3\beta/2}$ triangles ($B < \infty$) we get

$$\mu_t^{p,q}(t(\eta_0) \ge Bn^{3-3\beta/2}) \le \epsilon. \tag{5}$$

Conditioning on $t(\eta_0) \le Bn^{3-3\beta/2} \equiv Bn^{\gamma}$ (i.e. $\gamma = 3 - 3\beta/2 < 2$) we have for ξ_1 that

$$\mathbf{E}|\xi_1| \le \frac{h(n)}{n+h(n)} Bn^{\gamma} = \frac{o(\log n)}{n} Bn^{\gamma} \le Bn^{\gamma'}$$

for some $\gamma' < 1$. (To be correct, this is, as noted above, a conditional expected value. We allow ourselves to be a bit sloppy here and for the rest of the proof in order not to burden the notation.) Combining this with Markov's inequality and (5) it follows that for some $C < \infty$

$$\nu_t^{p,q}(|\xi_1| \ge Cn^{\gamma'}) \le 2\epsilon. \tag{6}$$

Next we condition on that $|\xi_1| \le Cn^{\gamma'}$ and consider η_1. For η_1 we have that at most $Cn^{\gamma'}$ triangles automatically follow from ξ_1 and these triangles consist of no more than $O(n^{\gamma'})$ edges which in turn can make at most $Dn^{3\gamma'/2}$ triangles, where D is yet another large but finite constant. Apart from this, new triangles can appear in the following three ways.

(a) *One* new edge ties together two of the edges which follow from ξ_1 and thereby produces a triangle.

(b) *Two* new edges make a triangle together with an already given edge.

(c) *Three* new edges make a triangle.

The expected contribution from (c) is $\binom{n}{3}O(n^{-3\alpha}) = O(n^{3-3\alpha})$. From (b) we expect to get at most

$$3Cn^{\gamma'}nO(n^{-2\alpha}) = O(n^{1+\gamma'-2\alpha})$$

triangles, as $Cn^{\gamma'}$ triangles consist of at most $3Cn^{\gamma'}$ edges and each of these edges can be made into a triangle in no more than $n-2 \leq n$ different ways, each with probability at most $O(n^{-2\alpha})$. The expected contribution from part (a) is no more than

$$\binom{3Cn^{\gamma'}}{2}O(n^{-\alpha}) = O(n^{2\gamma'-\alpha})$$

as the number of pairs of edges with a common end vertex is at most

$$\min(3Cn^{\gamma'}2n, \binom{3Cn^{\gamma'}}{2}).$$

Since $Dn^{3\gamma'/2}$ is the dominating contribution it follows that for n large enough

$$\mathbf{E}[t(\eta_1)] \leq (D+1)n^{3\gamma'/2}$$

and in the same way as before that

$$\mu_t^{p,q}(t(\eta_1) \geq En^{3\gamma'/2}) \leq 3\epsilon$$

for some $E < \infty$. For ξ_2 we repeat the same arguments as we used for ξ_1 to see that for some $F < \infty$

$$\nu_t^{p,q}(|\xi_2| \geq Fn^{\gamma''/2}) \leq 4\epsilon$$

for some γ'' between γ' and 1. Now proceed as we did for η_1; the $Fn^{\gamma''/2}$ triangles can implicitly make at most $Gn^{3\gamma''/4}$ triangles ($G < \infty$). From (c) we expect at most $O(n^{3-3\alpha})$ triangles, from (b) at most $3Fn^{\gamma''/2}nO(n^{-2\alpha}) = O(n^{3\gamma''/2-2\alpha})$ triangles and from (a) no more than $O(n^{\gamma''-\alpha})$ triangles. The largest one of the exponents in these expressions is strictly less than 1. Thus we can conclude that for some $H < \infty$ and some $\gamma''' < 1$ we have that

$$\mu_t^{p,q}(t(\eta_2) \geq Hn^{\gamma'''}) \leq 5\epsilon$$

and by using Markov's inequality once again we can find $J < \infty$ such that

$$\nu_t^{p,q}(|\xi_3| < Jn^{\gamma''''-1}) \geq 1-6\epsilon$$

for some $\gamma'''' < 1$. However, since $\gamma'''' < 1$ and ϵ was arbitrary, we are done. □

The technique of the above proof can be used to make further conclusions. First, we consider the case when p is exactly of order $n^{-2/3}$. For two finite measures, M_1 and M_2, on $\mathbf{Z}_+$, let $||M_1 - M_2||$ be the *total variation norm*, i.e. $||M_1 - M_2|| = \frac{1}{2}\sum_{k=0}^{\infty}|M_1(k) - M_2(k)|$.

THEOREM 4.3 *Let $q = 1 + h(n)/n$ with $h(n) = o(\log n)$ and let $p = an^{-2/3}$ for some positive constant a. Then, for $k = 1, 2, \ldots$*

$$||\nu_t^{p,q}(|\xi| \in \cdot) - Poi(\frac{a^3h(n)}{6})|| \to 0$$

as $n \to \infty$, where $Poi(\alpha)$ denotes the Poisson distribution with expectation α. Moreover, the same conclusion is valid if $|\xi|$ is replaced with $v(\xi)/3$, where $v(\xi)$ denotes the number of vertices which are part of a triangle, t, with $\xi(t) = 1$.

Proof. By letting $\alpha = 2/3$ in the proof of Theorem 4.2 we would in the last step end up with a constant $J < \infty$ such that

$$\nu_t^{p,q}(|\xi_3| \leq Jh(n)) \geq 1 - 6\epsilon.$$

For η_3 this means that the only substantial contribution to $t(\eta_3)$ comes from part (c). In other words, if we fix a small $\delta > 0$,

$$\mu_t^{p,q}(t(\eta_3) \leq (1 + n^{-1/4})\frac{a^3 n}{6}) \geq 1 - 7\epsilon$$

for large n, by Chebyshev's inequality, noting that $\mathbf{Var}(t(\eta_3)) = O(n)$. (Again, like in the proof of Theorem 4.2 this is, to be correct, a conditional variance.) This implies that with probability $1 - 7\epsilon$, $|\xi_4|$ gets a distribution which is stochastically dominated by

$$\mathcal{B}((1 + n^{-1/4})\frac{a^3 n}{6}, \frac{h(n)}{n + h(n)})$$

where $\mathcal{B}(m, s)$ denotes the law of the binomial distribution with parameters m and s.

On the other hand, it is a consequence of Corollary 3.3 that if η is chosen according to the random triangle model and $\hat{\eta}$ is chosen according to the $G(n,p)$ model with the same p, then

$$(|\eta|, t(\eta)) \overset{d}{\geq} (|\hat{\eta}|, t(\hat{\eta})).$$

If we generate corresponding elements, ξ and $\hat{\xi}$, in $\{0,1\}^T$ as indicated by Corollary 3.3(a), then ξ has distribution $\nu_t^{p,q}$ and $\xi \overset{d}{\geq} \hat{\xi}$. Again by Chebyshev's inequality

$$Pr(t(\hat{\eta}) \geq (1 - n^{-1/4})\frac{a^3 n}{6}) = 1 - o(1)$$

so that with probability $1 - o(1)$, $|\xi|$ is stochastically larger than

$$\mathcal{B}((1 - n^{-1/4})\frac{a^3 n}{6}, \frac{h(n)}{n + h(n)}).$$

By standard Poisson approximation of the binomial distribution noting that $n^{-1/4} = o(n^{-1/2})$ and that $n/(n + h(n)) \to 1$, this completes the proof of the first part.

To show the second part we must show that the triangles, t, with $\xi(t) = 1$ are with overwhelming probability disjoint. However, two triangles can be connected to each other in two ways; either by sharing an edge and thereby consist of five edges and four vertices or by sharing a vertex and thereby consist of six edges and five vertices. Taking this into account it is readily seen that the expected number of pairs of triangles connected to each other in η_3 above is $O(n)$. Since the probability that both triangles in a specific pair get $\xi_4(t) = 1$ is $h(n)^2/n^2$ it follows from Markov's inequality that the probability that this happens for any such pair tends to 0. $\square$

The moral of Theorem 4.3 is that for $p = \Theta(n^{-2/3})$ and $q = 1 + o(\log n)/n$, the random triangle graph is approximately the union of a $G(n,p)$ graph and a Poisson distributed number of disjoint, uniformly spread, triangles. Since the $G(n,p)$ graph for p of this kind is already connected in its own right and contains $\Theta(n)$ triangles, the effect of the triangle reward on the edge probability and the triangle probability is asymptotically neglectible. The same thing is true for the cases $\alpha \in (0, 2/3)$ even though we cannot give our statements the same precision as for the other cases.

THEOREM 4.4 *Assume that* $p = \Theta(n^{-\alpha})$ *for some* $\alpha \in (0, 2/3)$ *and let* $q = 1 + h(n)/n$ *for* $h(n) = o(\log n)$. *Then there is a constant* $A < \infty$ *such that*

$$\nu_t^{p,q}(|\xi| \geq Ah(n)n^{2-3\alpha}) \to 0$$

as $n \to \infty$.

Proof. This is essentially just a repetition of the proof Theorem 4.2. Again we generate the sequence $\eta_0, \xi_1, \eta_1, \xi_2, \ldots$. By observing that as soon as $\alpha' < \alpha$, the edge probability, r, is bounded by $n^{-\alpha'}$, we see that with probability $1 - o(1)$ we will have $|\eta_0| = O(n^{2-\alpha'})$ and $t(\eta_0) = O(n^{3-3\alpha'/2})$.

By copying the arguments of the proof of Theorem 4.2 it follows that the contribution of triangles from ξ_2 to η_2's becomes of lower order than the number of "spontaneously" appearing triangles. This means that we can, by Chebyshev's inequality, find $A < \infty$ such that

$$\mu_t^{p,q}(t(\eta_2) \leq \frac{A}{2}n^{3-3\alpha}) = 1 - o(1).$$

Finally, the central limit theorem implies that

$$\nu_t^{p,q}(|\xi_3| \geq Ah(n)n^{2-3\alpha}) = o(1)$$

as desired. □

Theorem 4.4 implies that if an η according to the random triangle measure with $p = \Theta(n^{-\alpha})$ and $h(n) = o(\log n)$ is generated as indicated in Corollary 3.3(b), the contribution from ξ is with overwhelming probability not more than a constant times $h(n)n^{2-3\alpha}$ triangles. Implicitly this means a maximum of $O(h(n)n^{2-3\alpha})$ edges and $O(h(n)^{3/2}n^{3-9\alpha/2})$ triangles. Since these numbers are of lower order than the spontaneously appearing $\Theta(n^{2-\alpha})$ edges and $\Theta(n^{3-3\alpha})$ triangles, the edge and triangle probabilities are not asymptotically different from what they are in the $G(n,p)$ case.

If $h(n) = \Theta(1)$ the proof can be adjusted to be valid as soon as $p = o(1)$ and if we try $h(n) = o(1)$, then the technique works even for $p = O(1)$ showing that in these cases the statement of Theorem 4.4 holds with $\alpha = 0$.

Thus far we have seen that for $p = O(n^{-\alpha})$, $\alpha > 0$, our way of rewarding triangles does not have any essential effect when $q = 1 + o(\log n)/n$ whereas if $q = 1 + \omega(\log n)/n$, then every edge will be present. Now, can anything else happen if $q = 1 + \Theta(\log n)/n$? Yes, since it is readily verified that the edge probability, r, is for fixed p and n a continuous function of q, it is clear that q can be chosen in such a way that r assumes any desired value in its range $[p, 1)$. However, what happens is that for such a q the probability mass will just divide between on one hand the outcomes where the effect of the triangle reward is neglectible and on the other hand the outcomes where all or almost all edges are present. As a matter of fact, if we let $p = \Theta(n^{-\alpha})$, then this division will only take place when $q = 1 + 3\alpha \log n/n + o(\log n)/n$:

THEOREM 4.5 *Let* $p = \Theta(n^{-\alpha})$ *for some* $\alpha > 0$ *and assume that* $q = 1 + c\log n/n + o(\log n)/n$ *for some* $c > 0$. *Let* η *be chosen by* $\mu_t^{p,q}$ *and let* ξ *have distribution* $\nu_t^{p,q}$ *and fix* $\epsilon > 0$. *Then there is a constant* $A < \infty$ *such that*

$$\nu_t^{p,q}(|\xi| \leq An^{2-3\alpha}\log n) + \mu_t^{p,q}(|\eta| \geq (1-\epsilon)\binom{n}{2}) \to 1$$

as $n \to \infty$. *If* $c < 3\alpha$, *then the first of these terms tends to* 1 *whereas if* $c > 3\alpha$, *then the second term tends to* 1.

Note. It is a consequence of Corollary 3.3 that the sum of the two terms in the theorem cannot asymptotically be larger than 1.

Proof. We shall first focus on showing that $\mu_t^{p,q}(n^{2-\alpha/2} \le |\eta| \le \rho\binom{n}{2}) \to 0$ if the constant ρ has been chosen small enough. When this has been done we can use the proof of Theorem 4.4 for the outcomes with $|\eta| \le n^{2-\alpha/2}$ by letting $\eta_0 = \eta$, to show that givens such an outcome the first term tends to 1. For the other outcomes we will have to do some more work.

The method to show this is "brute force". Let us try to bound $\mu_t^{p,q}(|\eta| = k)$ for $k = n^{2-\alpha/2}, \ldots, \rho\binom{n}{2}$. Write k as $\rho(n)\binom{n}{2}$, where $\rho(n)$ is thus either a constant ($\le \rho$) or a function tending to 0. Assume for simplicity that $p = n^{-\alpha}$ and $h(n) = c\log n$. It will be obvious from the expressions below that this does not upset things. For an outcome η' with $\rho(n)\binom{n}{2}$ edges we have

$$Z_t^{p,q}\mu_t^{p,q}(\eta') \le \frac{1}{n^{\alpha\rho(n)\binom{n}{2}}}(1 + \frac{c\log n}{n})^{\rho(n)^{3/2}\binom{n}{3}} \tag{7}$$

as $\rho(n)\binom{n}{2}$ edges make at most $\rho(n)^{3/2}\binom{n}{3}$ triangles. (Remember that $Z_t^{p,q}$ is the normalizing constant for the random triangle measure.) As $n \to \infty$ the right hand side of (7) is bounded by

$$\frac{1}{n^{\alpha\rho(n)\binom{n}{2}}} n^{c\rho(n)^{3/2}\binom{n}{3}/n} = n^{c\rho(n)^{3/2}\binom{n}{3}/n - \alpha\rho(n)\binom{n}{2}} \tag{8}$$

and, since $\rho(n) \le \rho$, it is clear that by letting ρ be small enough the exponent is less than $-\frac{3}{4}\alpha\rho(n)\binom{n}{2}$. There are $\binom{\binom{n}{2}}{\rho(n)\binom{n}{2}}$ different outcomes with $\rho(n)\binom{n}{2}$ edges. By Stirling's formula we have

$$\binom{\binom{n}{2}}{\rho(n)\binom{n}{2}} \le \frac{\binom{n}{2}^{\rho(n)\binom{n}{2}}}{(\rho(n)\binom{n}{2})!} \le e^{\rho(n)\binom{n}{2}}(\frac{1}{\rho(n)})^{\rho(n)\binom{n}{2}}$$

$$= (\frac{e}{\rho(n)})^{\rho(n)\binom{n}{2}} = n^{\log(\frac{e}{\rho(n)})\rho(n)\binom{n}{2}/\log n}.$$

Since $\rho(n) \ge n^{-\alpha/2}$, this exponent is at most $(\frac{1}{2}\alpha + \frac{1}{\log n})\rho(n)\binom{n}{2} \le \frac{5}{8}\rho(n)\binom{n}{2}$ so that by multiplying with the expression in (8) we get

$$Z_t^{p,q}\mu_t^{p,q}(|\eta| = \rho(n)\binom{n}{2}) \le n^{-\alpha\rho(n)\binom{n}{2}/8} \le n^{-\alpha\binom{n}{2}n^{-\alpha/2}/8}.$$

By multiplying by the number of possible values for k, which is not more than $\rho\binom{n}{2}$, it follows that

$$Z_t^{p,q}\mu_t^{p,q}(n^{2-\alpha/2} \le |\eta| \le \rho\binom{n}{2}) \to 0$$

as desired.

Next, we take a look at how the probability mass is distributed over the outcomes with more than $\rho\binom{n}{2}$ edges in case $c \ge 3\alpha$. Note first that

$$Z_t^{p,q}\mu_t^{p,q}(E) = \frac{(1 + \frac{c\log n}{n})^{\binom{n}{3}}}{n^{\alpha\binom{n}{2}}} = \frac{n^{c\binom{n}{3}/n + o(n^2)}}{n^{\alpha\binom{n}{2}}} = n^{c\binom{n}{3}/n - \alpha\binom{n}{2} + o(n^2)} \tag{9}$$

Now, pick γ such that $\rho \leq \gamma < 1$ and consider an edge configuration, η'', with $\gamma\binom{n}{2}$ edges. Then

$$Z_t^{p,q}\mu_t^{p,q}(\eta'') \leq n^{\gamma^{3/2}c\binom{n}{3}/n-\gamma\alpha\binom{n}{2}}$$

and since $\gamma^{3/2} < \gamma$ and $c \geq 3\alpha$, this expression is for $\gamma \leq 1-\epsilon$ at least $n^{\delta n^2}$ times smaller than $Z_t^{p,q}\mu_t^{p,q}(E)$ for some $\delta > 0$. Since the number of outcomes under consideration is certainly smaller than total number of outcomes which in turn is $2^{\binom{n}{2}}$, it follows for arbitrary $\rho > 0$ that

$$\frac{\mu_t^{p,q}(\rho\binom{n}{2} \leq |\eta| \leq (1-\epsilon)\binom{n}{2})}{\mu_t^{p,q}(E)} \to 0.$$

This shows that the mass on the outcomes with many edges is concentrated on the outcomes where almost all edges are present and thereby proves the first part of the theorem for the cases $c \geq 3\alpha$.

Also, for any outcome η' with $|\eta'| \leq \theta\binom{n}{2}$ we know that $q^{t(\eta')} \leq n^{c\theta^{3/2}\binom{n}{3}/n}$ and letting θ be small enough, this is of lower order than $n^{\delta n^2}$ for any fixed $\delta > 0$. An upper bound for $Z_t^{p,q}\mu_t^{p,q}(|\eta| \leq \theta\binom{n}{2})$ is therefore given by $n^{c\theta^{3/2}\binom{n}{3}/n}$ as the total mass for these outcomes for the ordinary $G(n,p)$ measure is of course bounded by 1. Together with the above results this yields for $c > 3\alpha$

$$\frac{\mu_t^{p,q}(|\eta| \leq (1-\epsilon)\binom{n}{2})}{\mu_t^{p,q}(E)} \to 0$$

as the exponent of (9) is larger than δn^2 if δ is small enough. This proves that the second term of the sum in the theorem tends to 1 if $c > 3\alpha$.

If on the other hand $c < 3\alpha$, then, with $\gamma < 1$ as before, we have for η' with $|\eta'| = \gamma\binom{n}{2}$

$$Z_t^{p,q}\mu_t^{p,q}(\eta') \leq n^{\gamma^{3/2}c\binom{n}{3}/n-\gamma\alpha\binom{n}{2}} \leq n^{-\gamma\beta n^2}$$

for a small positive β. Thus $Z_t^{p,q}\mu_t^{p,q}(|\eta| \geq \rho\binom{n}{2}) \to 0$ so that since $Z_t^{p,q} > 1$, the proof is complete. $\Box$

It should be noted that for $\alpha = 2/3$ or $\alpha > 2/3$ the stronger conclusions of Theorem 4.2 and Theorem 4.3 are valid if $c < 3\alpha$.

Concluding discussion. What we have seen is that for any $p = O(n^{-\alpha})$ for some $\alpha > 0$ the random triangle model is explosive; depending on q we get nothing or everything. The important moral of this is that for any random graph model with transitivity not degenerate in this sense, the nondegeneracy relies on the extra, and perhaps unintended, structure imposed on the graph. For instance, the random cluster model owes its nondegeneracy to the fact that an edge is more likely to be present given that its two end vertices are connected through the rest of the graph no matter how long that connection is.

Is there no way we can have a nondegenerate effect from rewarding triangles in the complete graph? Yes, if $p = \Theta(1)$ and $q = 1 + c/n$ this will happen. This can be seen by running the Gibbs sampler for which the random triangle measure is stationary distribution. If the sampler is started in a state chosen by the $G(n,p)$

model it is readily seen that with high probability we will after a while have got essentially more edges. On the other hand, if the sampler is started in E, we will soon have lost a large number of edges.

What is the intuitive explanation for the explosiveness of the model? Let us again consider the model from a dynamical point of view, i.e. by looking at it as the stationary distribution of a Gibbs sampler. If the sampler is started from a $G(n,p)$ state, the starting graph will be sparse in the the sense that a vast majority of the edges of E will be absent. This means that there will be very few potential triangles, i.e. pairs of edges with a common end vertex. Therefore the Gibbs sampler will stay for a very long time in states with few edges. Sooner or later, however, it will come to a state with $\Theta(n^2)$ edges. Once in a state like this, removing a present edge will with high probability mean to remove $\Theta(n)$ triangles. If q is large enough, the edges will therefore be very reluctant to be removed. Thus the states with $\Theta(n^2)$ edges are also very stable. It is however the moral of the proof of Theorem 4.2 that the states in between these two extremes are not at all stable. Therefore the Gibbs sampler will spend very little time in these states.

Letting $p = \Theta(1)$ on a large complete graph does not seem like a realistic model for a social network. To increase the realism and still have a significant effect from the transitivity in the random triangle model, it is therefore necessary to impose restrictions of some kind, such as geographical restrictions or restrictions on the degrees of the vertices etc. There are several possibilities along these lines. One is to study the model on graphs where the degrees do not increase too fast as the number of vertices increase. This line of study is taken up in [10], where the random triangle model on the two-dimensional triangular lattice is treated with respect to questions on phase transition and percolation. Another possibility is to make the model doubly random by looking at the random traingle model on a realization of a $G(n,p)$ graph. A third way to go could be to simply bound the degrees of the vertices in some appropriate way. This means to forbid the individuals to have "too many" friends. For instance one could study an appropriate version of the random triangle model on an r-regular graph for $r = 3, 4, 5, \ldots$.

References

1. A. Barbour, L. Holst and S. Janson, "Poisson Approximation," Oxford University Press, New York, 1992.

2. B. Bollobás, "Random Graphs," Academic Press, London, 1985.

3. B. Bollobás, G. Grimmett and S. Janson, The Random Cluster Model on the Complete Graph, *Probab. Theory Rel. Fields* **104** (1996), 283-317.

4. R. G. Edwards and A. D. Sokal, Generalization of the Fortuin-Kasteleyn-Swendsen-Wang Representation and Monte Carlo Algorithm, *Phys. Rev.* **38** (1988), 2009-2012.

5. P. Erdős and A. Rényi, On Random Graphs I, *Publ. Math. Debrecen* **6** (1959), 290-297.

6. K. Faust and S. Wasserman, "Social Network Analysis," Cambridge University Press, Cambridge, 1994.

7. C. M. FORTUIN AND P. W. KASTELEYN, On the Random Cluster Model I, Introduction and Relation to Other Models, *Physica* **58** (1972), 393-418.

8. O. FRANK, Random Sampling and Social Networks: A Survey of Various Approaches, *Math. Sci. Humaines* **104** (1988), 19-33.

9. O. HÄGGSTRÖM, "Random Cluster Representations in the Study of Phase Transitions," Preprint, Chalmers University of Technology and Göteborg University, 1996.

10. O. HÄGGSTRÖM AND J. JONASSON, Phase Transition in the Random Triangle Model, Preliminary version, 1997.

11. S. JANSON, D. KNUTH, T. LUCZAK AND B. PITTEL, The Birth of the Giant Component, *Rand. Struct. Alg.* **4** (1993), 233-358.

12. M. KARONSKI AND T. LUCZAK, Random Hypergraphs, *Combinatorics, Paul Erdös is Eighty*, Vol. 2, 283-293, Janos Bolyai Math Sci, Budapest, 1996.

13. M. KARONSKI, E. SCHEINERMAN AND K. SINGER, On Random Intersection Graphs: The Subgraph Problem, Preliminary version, 1997.

14. D. STRAUSS, On a General Class of Models for Interaction, *SIAM Review* **28** (1986), 513-527.

15. R. H. SWENDSEN AND J. S. WANG, Nonuniversal Critical Dynamics in Monte Carlo Simulations, *Phys. Rev. Lett.* **59** (1987), 86-88.