

Edge Union of Networks on the Same Vertex Set.

Chuan Wen, Loe¹ and Henrik Jeldtoft Jensen²

Department of Mathematics and Complexity & Networks Group
Imperial College London, London, SW7 2AZ, UK

E-mail: (1) c.loe11@imperial.ac.uk, (2) h.jensen@imperial.ac.uk

Abstract. Random networks generators like Erdős-Rényi, Watts-Strogatz and Barabási-Albert models are used as models to study real-world networks. Let $G^1(V, E_1)$ and $G^2(V, E_2)$ be two such networks on the same vertex set V . This paper studies the degree distribution and clustering coefficient of the resultant networks, $G(V, E_1 \cup E_2)$.

1. Introduction

Random networks are used as models to study real-world networks like the Internet and social networks. Some of the families of random networks generators are Erdős-Rényi, Watts-Strogatz and Barabási-Albert models. These models generate networks with only one type of edges and hence may be thought of as relating to agents interacting under a homogeneous relation [1, 2, 3].

However recently research has started to address the limitations of homogeneous networks. In the real world, agents may be related through more than one form of relations. E.g. think of social interactions between people, one type of relations may consist of family bonds another of shared interests, such as chess, a third could be professional relations and so forth.

There are many names for networks with multiple types of connections, different researchers use different names: Multigraph [4], Multi-Layered [5, 6, 7, 8], Multi-Relational [9, 10, 11, 12, 13], Multi-Dimensional [14, 15, 16, 17, 18, 19], Multiplex [11, 20, 21, 22, 23], Multi-Modal [24] or graphs on the same vertex set [25]. We will here use the term multi-dimensional. We may think of these networks as compose of layers of different relations (dimensions). Each layer has the same set of vertices, and in every layer the vertices are connected through different relationships.

There are various proposals to solve classic problems like link prediction and community detection in multi-dimensional networks. However, so far the structural property of the entire network has not been investigated. Given the statistical properties of two layers, we address in this paper the statistical properties of the union of the different types of edges. Combining edges of different types is known as the *edge union of networks on the same vertex set*.

We note that the famous Zachary Karate Club Network [26] is an example of an edge union of 8 networks on the same vertex set. Any two members (vertices) of the karate club are connected if they have consistent interactions outside classes and club meetings. “Outside interactions” are defined in terms of 8 different types of relationships. One of such relations is the association in and between academic classes at the university. The 8 different relationships are then combine as edges (of different types) on the same vertex set. Other real-world applications of these kind of networks can be found in [27] and [28].

As another example of the relevance of multi-dimensional networks think of the network of email exchanges between people. Assume that the preferential attachment of the Barabási-Albert model is true in this context, where a new member tends to email the more popular members. However being “popular” at a personal level is different from being “popular” in the work environment. If the database of email exchange does not distinguish the two contexts, the email exchange network is essentially the union of two Barabási-Albert networks.

The present paper is a study of some of the statistical properties of a random network that is formed by the union of two random networks. The different families

of random network are distinguished by their clustering coefficients and the degree distribution of the vertices and we want to determine the corresponding measures for the union of different types of networks. We first consider the degree distributions and then the clustering coefficients.

2. Network Families

This section summarises the construction of Erdős-Rényi, Watts-Strogatz and Barabási-Albert models. Recall that these random network generators were designed to construct simple graphs. We are going to combine networks with different edge statistics but with the same number of vertices, accordingly we suppress in the notation below the number of vertices in the labelling of the specific network types. For instance, the usual notation for Gilbert Graph is $G_{n,p}$, where n is the size of the vertex set. We simplify the notation to G_p as the size of the vertex set is implied to be n for all the networks in the rest of the paper. Lastly, let V and E be vertex set and edge set respectively.

2.1. Erdős-Rényi

A realisation of an Erdős-Rényi network [29] is selected with equal probability from the set of all possible graphs with $n = |V|$ vertices and $|E|$ edges. However to generate a huge random Erdős-Rényi is difficult. To circumvent this problem one may instead let the number of edges fluctuate slightly and consider a Gilbert graph [30] G_p in which every vertex pair is connected with probability p where $p \approx |E|/\binom{|V|}{2}$. The slight difference is that Erdős-Rényi has precisely $|E|$ edges while Gilbert graph has approximately $|E|$ edges with high probability.

2.2. Watts-Strogatz

A Watts-Strogatz network [31], $W_{w,q}$ is parameterised by w and q for mean degree[‡] and probability of rewiring respectively. The construction begins with a regular ring lattice where each vertex connects to $w/2$ neighbours on each side.

For each vertex $v_i \in V$ and $i < a$, each edge leaving v_i is rewired with probability q . The rewiring replaces $\{v_i, v_a\}$ with $\{v_i, v_b\}$ where v_b is chosen uniformly in V and the resultant network remains a simple network.

The key property of the Watts-Strogatz network is that it varies between a regular ring lattice ($q = 0$) and an Erdős-Rényi network ($q = 1$). As $q \rightarrow 1$, Watts-Strogatz can be expressed as a Erdős-Rényi graph, i.e. $W_{w,q} \rightarrow G_{w/(n-1)}$.

[‡] We replace the commonly used variable k with w to avoid confusion with the variable k commonly used in degree distribution function

2.3. Barabási-Albert

Barabási-Albert network [32] is denoted by B_m where m is the number of new edges at each iteration. The network construction begins with some arbitrary small number of vertices connected randomly.

At each iteration, one new vertex of degree m is added. The edges of the new vertex are connected probabilistically with a probability proportional to the degree of the existing vertices. Define $deg(v_i)$ as the degree of vertex v_i . The probability that the new vertex is connected to vertex v_i is given by:

$$p_i = \frac{deg(v_i)}{\sum_j deg(v_j)}.$$

This is referred to as preferential attachment.

3. Definitions and Preliminaries

3.1. Definitions

For brevity, the resultant network from the edge union of networks is defined as the *composite network* (abbreviation: \mathcal{C}). A *network* refers to either an Erdős-Rényi, Watts-Strogatz or Barabási-Albert network (abbreviation: ER, WS, BA respectively). The composition of all 6 pairwise edge unions of these networks are $\{ER, ER\}$, $\{ER, WS\}$, $\{ER, BA\}$, $\{WS, WS\}$, $\{WS, BA\}$ and $\{BA, BA\}$.

Definition 3.1. Let $G^1(V, E_1)$ and $G^2(V, E_2)$ be two networks on the same vertex set V . The **edge union of the networks** gives the composite network, $\mathcal{C}(V, E_1 \cup E_2)$.

Definition 3.2. Let $G^1(V, E_1)$ and $G^2(V, E_2)$ be two networks on the same vertex set V . An edge $e \in E_1 \cup E_2$ in the composite network is called a **common edge** if $e \in E_1 \cap E_2$. Hence the set of common edges is $E_1 \cap E_2$.

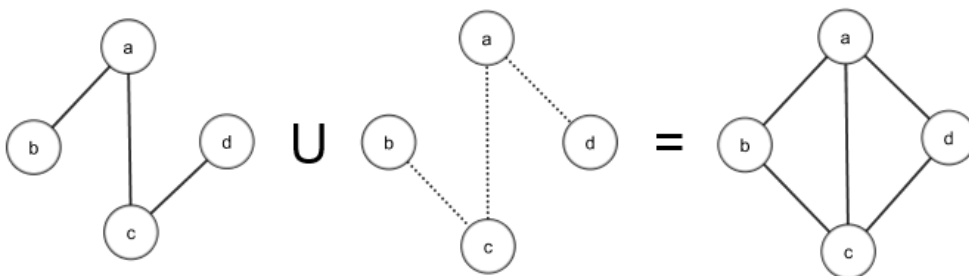


Figure 1: Edge union of two networks on the same vertex set. The solid edges and the dotted edges are from the edge set of different networks. The composite network is the rightmost network. Edge $\{a, c\}$ is a common edge between the two networks.

Definition 3.3. Consider a population of n urns. Choose $a < n$ urns and place one red ball in each. Next choose $b < n$ urns and place one blue ball in each. The probability that exactly i urns contain one red and one blue ball is given the hypergeometric function

$$\mathcal{H}(i; n, a, b) = \frac{\binom{a}{i} \binom{n-a}{b-i}}{\binom{n}{b}}$$

3.2. Preliminaries

The following lemmas and concepts are useful in the later analysis.

Lemma 3.1. *Let E_1 and E_2 be the sets of edges of two networks with n vertices each. The probability that there are ϵ common edges is:*

$$P(|E_1 \cap E_2| = \epsilon) = \mathcal{H}(\epsilon; \binom{n}{2}, |E_1|, |E_2|) \quad (1)$$

Proof. Choose $|E_2|$ edges out of a total possible $\binom{n}{2}$ edges and colour them blue. Next choose $|E_1|$ edges from the set E_1 , the probability of getting ϵ blue edges is defined by the hypergeometric function. \square

We note that equation 1 is the probability that there are ϵ common edges between the two networks. In addition, the expected number of common edges is given by:

Corollary 3.1. *Let E_1 and E_2 be the set of edges of two networks with n vertices each. The expected number of common edges is the expectancy of the hypergeometric function in equation 1:*

$$\mathbb{E}[|E_1 \cap E_2|] = \mathbb{E}[\mathcal{H}] = \frac{|E_1| \cdot |E_2|}{\binom{n}{2}} \quad (2)$$

In fact we can generalise the above to compute the number of common cliques in the union. In particular it is useful to count the number of common triangles (3-clique) as it affects the accuracy of the clustering coefficient in the union to two networks.

Lemma 3.2. *Let K_1 and K_2 be the sets of c -cliques of two networks with n vertices each. The probability that there are ϵ common c -cliques is:*

$$P(|K_1 \cap K_2| = \epsilon) = \mathcal{H}(\epsilon; \binom{n}{c}, |K_1|, |K_2|) \quad (3)$$

Proof. Choose $|K_2|$ c -cliques out of a total possible $\binom{n}{c}$ c -cliques and colour them blue. Next choose $|K_1|$ c -cliques from the set K_1 , the probability of getting ϵ blue c -cliques is defined by the hypergeometric function. \square

Corollary 3.2. *Let K_1 and K_2 be the sets of c -cliques of two networks with n vertices each. The expected number of common c -cliques is the expectancy of the hypergeometric function in equation 3:*

$$\mathbb{E}[|K_1 \cap K_2|] = \mathbb{E}[\mathcal{H}] = \frac{|K_1| \cdot |K_2|}{\binom{n}{c}} \quad (4)$$

To compute the degree distribution of the composite network accurately, we have to minimise any double counting of the common edges. For example in Fig. 1, vertex a of the composite network has degree 3. It is the sum of the degree of the networks (degree 2 from each network) **minus** the number of common edges (= 1) at vertex a .

Lemma 3.3. *Let v_{i,G^1} and v_{i,G^2} be the vertices of networks G^1 and G^2 each with n vertices respectively. Let $d_1 = \deg(v_{i,G^1})$ and $d_2 = \deg(v_{i,G^2})$, the probability that there are ϵ common edges between v_{i,G^1} and v_{i,G^2} is:*

$$P_{ce}(\epsilon|d_1, d_2) = \mathcal{H}(\epsilon; n - 1, d_1, d_2) \quad (5)$$

Proof. There are only $n - 1$ vertices left for v_{i,G^1} and v_{i,G^2} to connect to. Using the same argument as in Lemma 3.1, there are d_1 blue edges and d_2 red edges from v_{i,G^1} and v_{i,G^2} respectively. \square

Corollary 3.3. *Let v_{i,G^1} and v_{i,G^2} be the vertices of networks G^1 and G^2 with n vertices respectively. Let $d_1 = \deg(v_{i,G^1})$ and $d_2 = \deg(v_{i,G^2})$, the expected number of common edges between v_{i,G^1} and v_{i,G^2} is:*

$$\mathbb{E}[\mathcal{H}] = \frac{|d_1| \cdot |d_2|}{n - 1} \quad (6)$$

In this paper we only consider the case where both networks have approximately the same edge set size. Unequal edge sets size introduces additional complexity to the study, since the larger edge set dominates the resultant network.

4. Degree Distribution of Edge Union of Networks

4.1. Erdős-Rényi with Erdős-Rényi

Let G_p^1 and $G_{p'}^2$ be two independent Gilbert Graphs. The probability that an edge does not exist in both G_p^1 and $G_{p'}^2$ is $(1 - p)(1 - p')$, accordingly the probability that the composite network will contain an edge is given by $1 - (1 - p)(1 - p')$ and we have

$$G_p^1 \cup G_{p'}^2 = G_{1-(1-p)(1-p')} \quad (7)$$

4.2. Erdős-Rényi with Watts-Strogatz

Let G_p and $W_{w,q}$ be an ER network and WS network respectively. As $q \rightarrow 1$, $W_{w,q}$ can be approximated as an ER so in this limit the union will be scribed by Eq. (7).

For $q \rightarrow 0$, the degree distribution of the WS network approaches $P^W(k) = \delta(k-w)$, the number of vertices with degree k . Neglecting common edges, a vertex in the composite network will have degree k when the corresponding vertex in ER has degree $k - w$. Hence an approximation to the degree distribution of the composite network is:

$$P^C(k) \sim \frac{(np)^{(k-w)} e^{-np}}{(k-w)!} \quad (8)$$

We can improve the estimate of the degree distribution by explicitly discounting the common edges in the union. This will become of increasing importance as q is increased from 0. Let the degree distribution of ER and WS be $P^G(k)$ and $P^W(k)$ respectively. To account for the common edges for $P^C(k)$, we have to consider:

- (i) Probability that a vertex in WS, v_{ws} has degree j ;
- (ii) Probability that a vertex in ER, v_{er} has degree $k + \epsilon - j$;
- (iii) Probability that there are ϵ common edges between v_{er} and v_{ws} .

From lemma 3.3, the probability that there are ϵ common edges between v_{er} and v_{ws} is $P_{ce}(\epsilon|j, k + \epsilon - j)$. Thus the improvement on equation 8 is:

$$P^C(k) \sim \sum_{\epsilon}^k \sum_j^k P_{ce}(\epsilon|j, k + \epsilon - j) P^W(j) P^G(k + \epsilon - j) \quad (9)$$

Since WS is lattice-like, most of the WS vertices have a degree close to w . By assuming $P^W(k) = \delta(k - w)$ we can simplify equation 9 to:

$$P^C(k) \sim \sum_{\epsilon}^w P_{ce}(\epsilon|w, k + \epsilon - w) P^G(k + \epsilon - w) \quad (10)$$

4.3. Erdős-Rényi with Barabási-Albert

It has been noticed that the degree distribution of the BA model often does not represent natural occurring networks. This has lead to various generalised BA models [33, 34, 35] in which preferential and random uniform attachment are combined. This construction is similar to the union of ER with BA, where ER adds some uniform attachment to BA.

To compute the degree distribution, we can use the method from section 4.2:

$$P^C(k) \sim \sum_{\epsilon}^k \sum_j^k P_{ce}(\epsilon|j, k + \epsilon - j) P^G(j) P^B(k + \epsilon - j) \quad (11)$$

where the approximate degree distribution of BA is $P^B(k) \sim 2m^2k^{-3}$ [41].

However the hypergeometric function, $P_{ce}(\epsilon|j, k + \epsilon - j)$ is computationally expensive. It is difficult to calculate the sums in Eq. (11) for large values of k and thereby assess how the additional uniform attachment influence the power law form of the BA network. To extract the asymptotic behaviour, we now make use of a Fokker-Planck approach. The idea is to begin with the BA network and then in each time step add a new uniformly drawn random edge to the BA network. The new edges are taken from the set of ER edges not common with the BA edge set. Since there are nm edges in BA, from corollary 3.1 it follows that there are nmp common edges or $\binom{n}{2}p - nmp$ non-common edges.

Let $u(k, t)$ be the number of vertices of degree k at time step t . At each time step, a new edge will change the degree of 2 vertices. Specifically the number of degree k vertices increases by one if the new edge attaches to a degree $k - 1$ vertex. This has probability $u(k - 1, t)/n$. Similarly the number of k vertices will decrease by one if the new edge attaches to a degree k vertex. This leads to the following equation,

$$u(k, t + 1) = u(k, t) + u(k - 1, t)/n - u(k, t)/n \quad (12)$$

By replacing t and k by continuous variables, we obtain a partial differential equation which will be a good approximation for large values of k .

$$\frac{\partial u}{\partial t} + \frac{1}{n} \frac{\partial u}{\partial k} = 0. \quad (13)$$

Using the initial condition $u(k, 0) = P^B(k) \cdot n \sim 2nm^2k^{-3}$ ($P^B(k) = 0$ for $k < m$), we can solve for $u(k, t)$ at $t = 2 \cdot \binom{n}{2}p - nmp \approx np(n - m)$ (twice the number of non-common edges because there are two vertices that change at each time step) and find,

$$P^C(k) = u(k, t) = \frac{2nm^2}{(k + 2n(p - p^2))^3} \quad (14)$$

In Fig. 2 we compare this asymptotic expression to the analytic expression derived by iterating Eq. (12) and to simulation results. The continuum approximation of Eq. (13) does not really have a region of validity since the finite number of vertices limits the size of the degree and the asymptotic limit cannot be reached. However, the iterated solution of the Fokker-Planck equation (12) matches the simulations well.

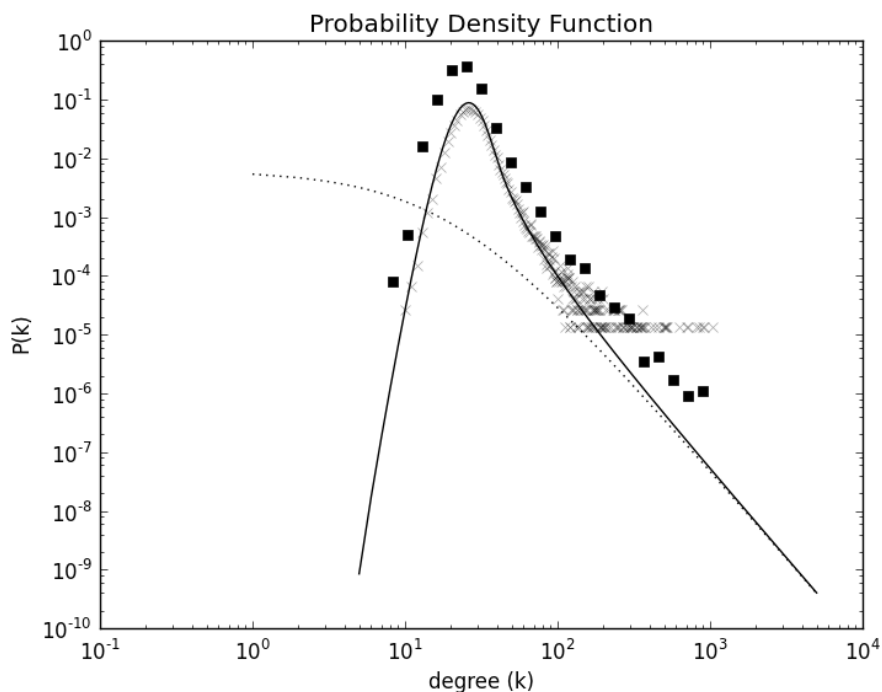


Figure 2: Parameters: $n = 75000$, $m = 10$ and $p = 2m/(n - 1)$ (chosen such that both networks have equal number of edges). The solid line plots the iterative method, Eq. (12). The dotted line plots the closed form Eq. (14). The crosses represent the distribution obtained from simulations. The degree-distribution log-binning of the crosses is plotted with squares.

4.4. Watts-Strogatz with Watts-Strogatz

Let $W_{w,q}^1$ and $W_{w',q'}^2$ be two WS networks. If $q \approx q' \rightarrow 1$, the two WS can be expressed as two ER (section 4.1). In contrast the limit $q \rightarrow 1$ and $q' \rightarrow 0$ is identical to the union of ER and WS considered in section 4.2. We do not discuss the case for general values of q and q' .

4.5. Watts-Strogatz with Barabási-Albert

Let $W_{w,q}$ and B_m be WS and BA graphs respectively. There are $wn/2$ edges in WS and nm edges in BA. It is most interesting to consider the case when the graphs have approximately equal number of edges, i.e. $w \approx 2m$. We ignore the case $q \rightarrow 1$, as it is similar to Sec. 4.3.

For $q \rightarrow 0$, the degree distribution of \mathcal{C} is straight forward. Since the probability of rewiring is low, most of the lattice edges in WS remain unchanged. Hence most of the vertices in WS have degree w , and in turn contributes to an increase in the degree of the vertices of BA by w . This gives the approximation for:

$$P^{\mathcal{C}}(k) \sim 2m^2(k-w)^{-3} \quad (15)$$

The degree distribution can be refined by considering the common edges between WS and BA. However the expression will be in open form and provides little insights. From simulations ($q \rightarrow 0$), Eq. 15 is sufficient to approximate the distribution (Fig. 3).

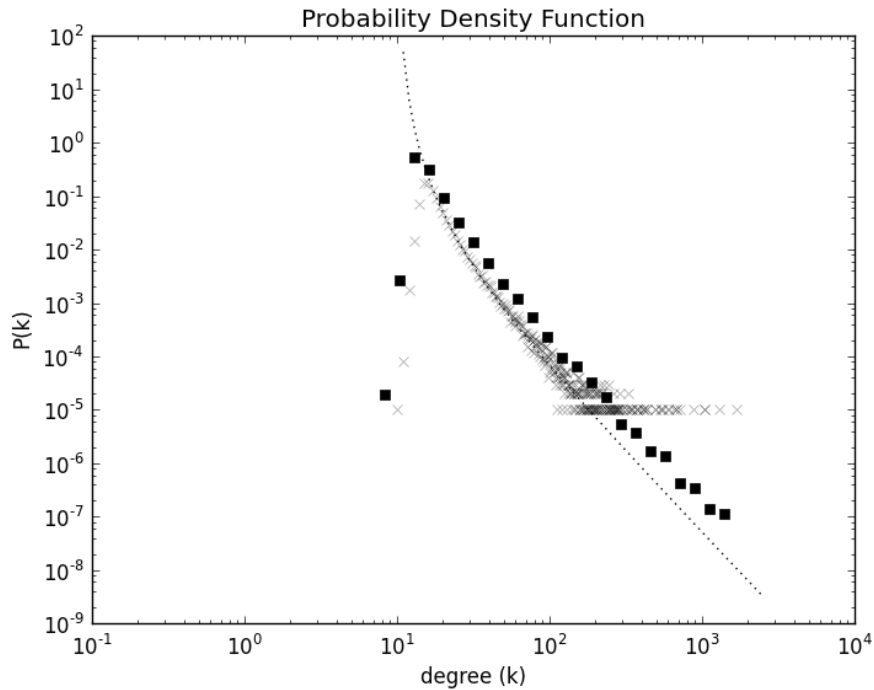


Figure 3: Parameters: $n = 100000$, $w = 10$, $m = 5$ and $q = 0.1$. The crosses represents the degree distribution obtained from simulations. The dotted line plots Eq. 15. The degree-distribution log-binning of the crosses is plotted with squares.

4.6. Barabási-Albert with Barabási-Albert

The key feature of BA is that the degree distribution is scale-free, i.e. follows a power-law in the degree. Hence we want to know if the union of two BA retain the features. Similar to equation 9, the probability density function is the combined probability of:

$$P^C(k) \sim \sum_{\epsilon} \sum_j^k P_{ce}(\epsilon|j, k + \epsilon - j) P^B(j) P^B(k + \epsilon - j) \quad (16)$$

Unfortunately, we are not able to simplify equation 16 to derive the asymptotic behaviour. However simulations (figure 4), indicate a power-law distribution for large k .

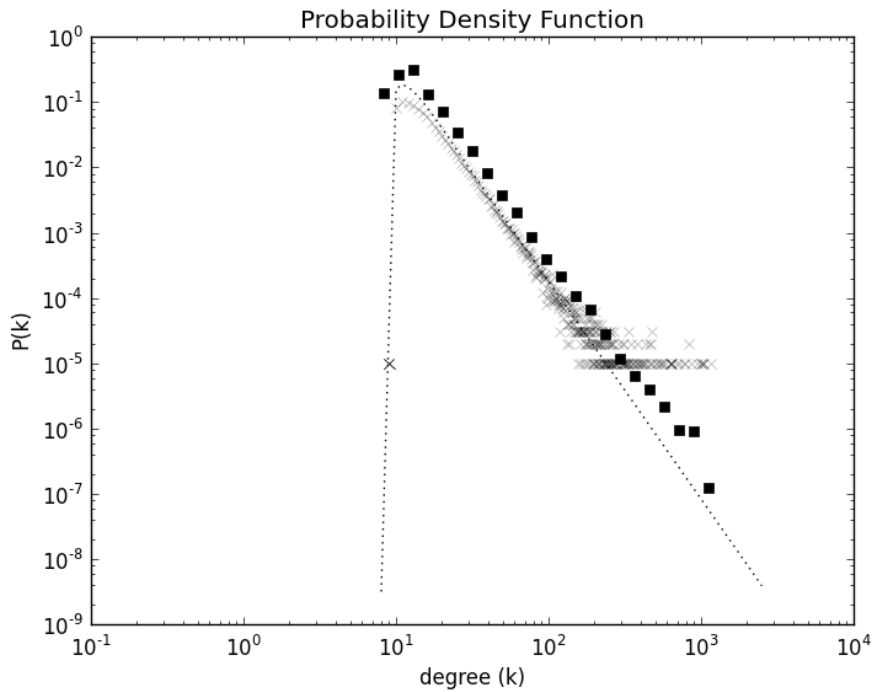


Figure 4: The dashed line represents Eq. 16 for $n = 100000$ and $m = 5$. The crosses represent the results of simulations. The degree-distribution log-binning of the crosses is plotted with squares.

Unlike in Sec. 4.3, we are not able to apply the same method to the union of two BA networks. In Sec. 4.3 we began with a BA and then we added uncorrelated random edges onto it. In the present case the preferential attachment induces correlations which cannot be neglected. To illustrate this better, let the first BA have red edges, and the second BA to have blue edges. Begin with the BA with red edges, and assume we can put the first few blue edges randomly like ER. However the subsequent blue edges have a preferential bias for **only the blue** edges. Since $u(k,t)$ in section 4.3 does not differentiate the colours on the edges, the method does not apply.

Figure 9 shows the degree rank plot of different combinations of the network union. It is clear that the distribution of ER/BA is slightly different from BA/BA. The union of BA/BA has more high degree vertices than those of BA/ER.

5. Clustering Coefficients of Edge Union of Networks

Let us now consider the clustering coefficient of the composite network. While we are not always able to analytically determine the clustering coefficient, we are able to derive lower and upper bounds. Again we assume that there are few common edges/triangles in the union, which can be verified with corollary 3.1 and corollary 3.2. This simplification is reasonable for sparse networks and it gives us simple relationships similar to Eq. (20).

5.1. Watts-Strogatz with Watts-Strogatz

The clustering coefficient of such a union can be computed analytically in the limit $p \approx q$ and $p, q \rightarrow 0$. Define $T(v_i)$ as the number of subgraphs with 3 edges and 3 vertices (triangles) which all share the vertex v_i . Recall the clustering coefficient equation for a network:

$$\text{Clustering Coefficient} = \frac{1}{n} \sum_{i=0}^n \text{Clustering Coefficient of } v_i \quad (17)$$

$$= \frac{1}{n} \sum_{i=0}^n \frac{T(v_i)}{\binom{\text{deg}(v_i)}{2}} \quad (18)$$

$$(19)$$

Let v_{i,w^1} and v_{i,w^2} be the vertices of W^1 and W^2 respectively, which are mapped to the i^{th} vertex ($v_{i,c}$) of \mathcal{C} . Since $p \approx q$ and $p, q \rightarrow 0$, almost every vertex of W^1 and W^2 are similar in structure. i.e. $T(v_{i,w^1}) \approx T(v_{i,w^2})$ and $\text{deg}(v_{i,w^1}) \approx \text{deg}(v_{i,w^2})$. Hence even with random pairing, every vertex of the composite network \mathcal{C} will be similar. We therefore have for small w :

$$\begin{aligned} \text{Clustering Coefficient of } \mathcal{C} &= \frac{1}{n} \sum_{i=0}^n \text{Clustering Coefficient of vertex } v_{i,c} \\ &= \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,c})}{\binom{\text{deg}(v_{i,c})}{2}} \end{aligned}$$

Because w is small ($w \ll n$), we assume there are few common edges. Hence $T(v_{i,c}) = T(v_{i,w^1}) + T(v_{i,w^2}) \approx 2T(v_{i,w^1})$ and $\text{deg}(v_{i,c}) \approx 2\text{deg}(v_{i,w^1})$. Then:

$$\text{Clustering Coefficient of } \mathcal{C} \approx \frac{1}{n} \sum_{i=0}^n \left(\frac{2T(v_{i,w^1})}{\binom{2\text{deg}(v_{i,w^1})}{2}} \right)$$

$$\begin{aligned}
&\approx \frac{1}{n} \sum_{i=0}^n \frac{1}{2} \left(\frac{T(v_{i,w^1})}{\binom{\deg(v_{i,w^1})}{2}} \right) \\
&\approx \frac{1}{2} \text{Clustering Coefficient of } W^1 \quad (20)
\end{aligned}$$

Fig. 5 compares the results of Eq. (20) with actual simulation of the edge union of two WS. The dots plot the expected clustering coefficient from Eq. (20) by halving the clustering coefficient of one of the WS. The crosses represent the actual clustering coefficient obtained from simulations. The details of the simulations are explained in the appendix.

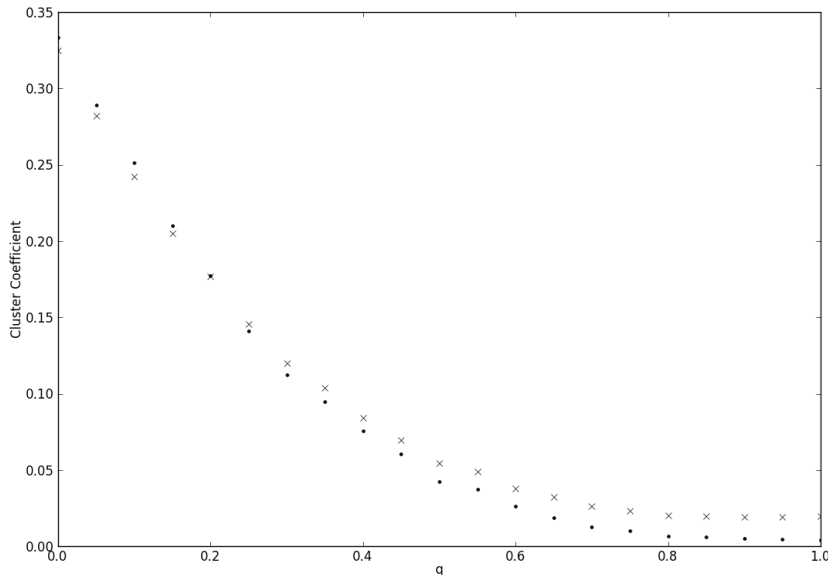


Figure 5: Edge Union of two WS networks for $n = 1000$, $w' = w = 20$ and $q' = q$. The dots are from Eq. (20) and the crosses are from the simulations.

5.2. Watts-Strogatz with Barabási-Albert

The clustering coefficient of the composite is difficult to approximate analytically as there is no simple way to approximate the number of triangles for BA. Instead, we derive a lower bound and an upper bound for the composite:

$$\begin{aligned}
\text{Clustering Coefficient of } \mathcal{C} &= \frac{1}{n} \sum_{i=0}^n \text{Clustering Coefficient of vertex } v_{i,c} \\
&= \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,c})}{\binom{\deg(v_{i,c})}{2}}
\end{aligned}$$

Assume $T(v_{i,c}) \approx T(v_{i,ws}) + T(v_{i,ba})$.

$$\begin{aligned} \text{Clustering Coefficient of } \mathcal{C} &\approx \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws}) + T(v_{i,ba})}{\binom{\text{deg}(v_{i,c})}{2}} \\ &\geq \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} \end{aligned}$$

Since we are considering $q \rightarrow 0$, we can make the following two observations: 1) The number of triangles generated by WS is generally larger than BA, hence choosing $T(v_{i,ws})$ will get a tighter bound. 2) Most of the vertices in WS have the same number of triangles attached due to the assumed low rewiring probability. Let α be the average number of triangles for any vertex in WS given q . Thus:

$$\text{Clustering Coefficient of } \mathcal{C} \geq \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} \quad (21)$$

$$\approx \frac{\alpha}{n} \sum_{i=0}^n \frac{1}{\binom{\text{deg}(v_{i,c})}{2}} \quad (22)$$

$$= \alpha \sum_{k=0}^n P^{\mathcal{C}}(k) \frac{1}{\binom{k}{2}} \quad (23)$$

From [36], the clustering coefficient of WS decreases at the rate $(1 - q)^3$ and $\alpha \approx (w/2)^2$ for $q = 0$. This is also equal to the rate of decrease to the number of triangles since $(1 - q)^3$ is the probability that none of the edges of a triangle is rewired. Thus we have the following lower bound on the clustering coefficient of the composite network:

$$\text{Clustering Coefficient of } \mathcal{C} \geq \frac{(1 - q)^3 w^2}{4} \sum_{k=0}^n P^{\mathcal{C}}(k) \frac{1}{\binom{k}{2}} \quad (24)$$

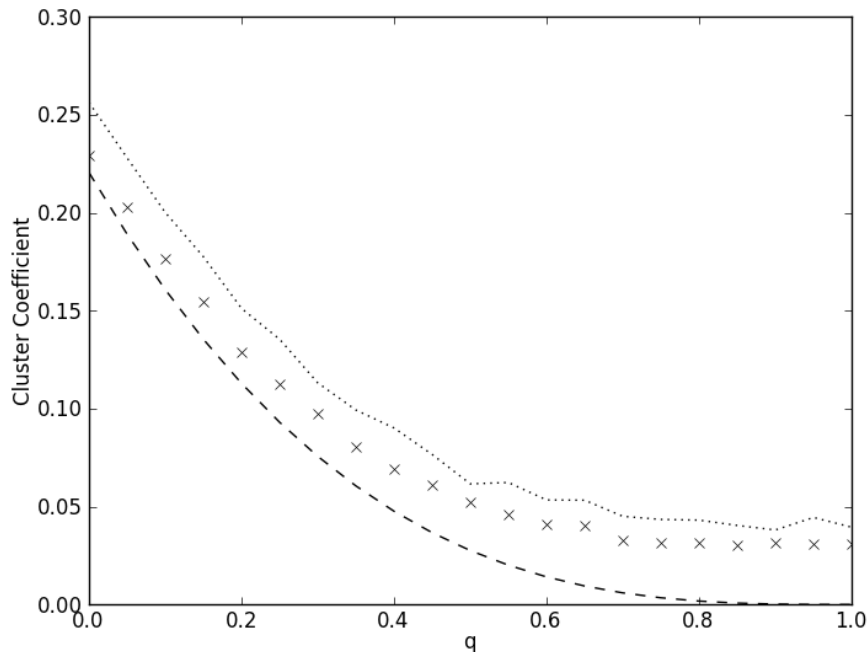


Figure 6: Parameters: $n = 1000$, $w = 10$ and $m = 5$. The crosses plot the clustering coefficient from the simulation of the composite network. The dashed line is the lower bound from Eq. 24. The dotted line is the upper bound from Eq. 25

To obtain an upper bound on the clustering coefficient, we can round up the clustering coefficient contribution originating from BA. I.e.

$$\begin{aligned}
\text{Clustering Coefficient of } \mathcal{C} &\approx \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws}) + T(v_{i,ba})}{\binom{\text{deg}(v_{i,c})}{2}} \\
&= \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} + \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ba})}{\binom{\text{deg}(v_{i,c})}{2}} \\
&\leq \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} + \frac{1}{n} \sum_{i=0}^n \frac{T(v_{i,ba})}{\binom{\text{deg}(v_{i,ba})}{2}} \\
&\approx \frac{(1-q)^3 w^2}{4} \sum_{k=0}^n P^{\mathcal{C}}(k) \frac{1}{\binom{k}{2}} \\
&\quad + \text{Clustering Coefficient of } B_m. \tag{25}
\end{aligned}$$

5.3. Erdős-Rényi with Erdős-Rényi

From section 4.3, the union of two ER is an ER network. For simplification, let the composite network be G_p^c . Hence the clustering coefficient of G_p^c can be estimated like any ER network.

However the “clustering coefficient” of ER is often measured by an alternative similar metric, Transitivity Ratio [37]. It is the ratio of the number of triangles to the number of connected triples:

$$\text{Transitivity Ratio of } G_p^c = \frac{\binom{n}{3}p^3}{\binom{n}{3}p^2} = p. \quad (26)$$

5.4. Clustering coefficient of other combinations

In the next section, we show that the clustering coefficient of ERUBA and BAUBA are small in general (Fig. 8). For this reason we don't derive any analytical results in this paper.

WSUER is a more interesting combination, but is difficult to deduce the clustering coefficient of the union from the know results for the individual families of networks. The clustering coefficient of WS is measured by Eq. 17 and the “clustering coefficient” (transitivity ratio) of ER is measured by Eq. 26. Since these metrics are different in general [38], we are only able to deduce a weak estimate of the union (Fig. 7).

Firstly in the composite network, the maximum number of triangles at vertex $v_{i,c}$ is $\binom{\text{deg}(v_{i,c})}{2}$. Among them, $T(v_{i,ws})$ triangles are from WS. The rest of the triangles exist if the edges of ER connect the neighbors of $v_{i,c}$ that are not connected from the edges of WS.

With probability p from ER, the number of neighbors pairs at $v_{i,c}$ that are connected by the edges of ER (and not from WS) is:

$$\left(\binom{\text{deg}(v_{i,c})}{2} - T(v_{i,ws}) \right) p$$

Hence,

$$\begin{aligned} \text{Clustering Coefficient of } \mathcal{C} &\approx \frac{1}{n} \sum_{i=0}^n \frac{\left(\binom{\text{deg}(v_{i,c})}{2} - T(v_{i,ws}) \right) p + T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} \\ &= \frac{1}{n} \sum_{i=0}^n \frac{\binom{\text{deg}(v_{i,c})}{2} p + T(v_{i,ws})(1-p)}{\binom{\text{deg}(v_{i,c})}{2}} \\ &= p + \frac{(1-p)}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} \end{aligned}$$

Together with approximation from Eq. 21 and Eq. 24,

$$\begin{aligned} \text{Clustering Coefficient of } \mathcal{C} &= p + \frac{(1-p)}{n} \sum_{i=0}^n \frac{T(v_{i,ws})}{\binom{\text{deg}(v_{i,c})}{2}} \\ &\approx p + \frac{(1-p)(1-q)^3 w^2}{4} \sum_{k=0}^n P^c(k) \frac{1}{\binom{k}{2}} \quad (27) \end{aligned}$$

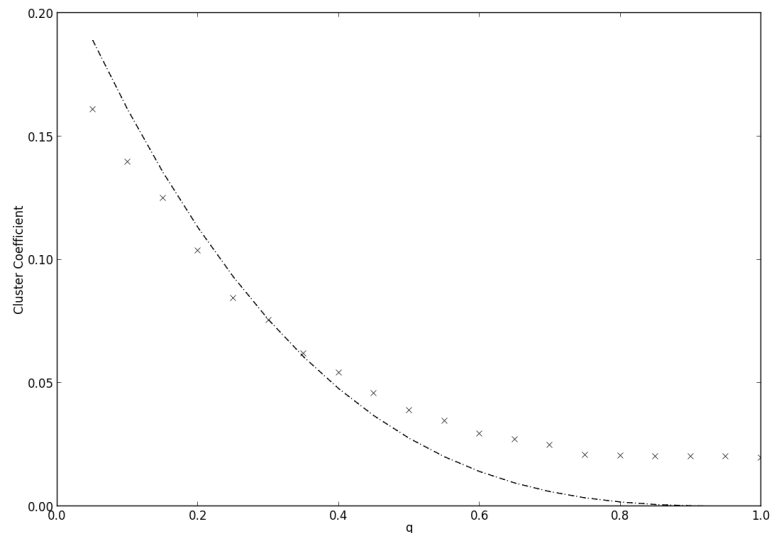


Figure 7: Parameters: $n = 1000$, $w = 10$ and $p \approx 0.01$. The crosses plot the clustering coefficient from the simulation of the composite network, $WS \cup ER$. The dashed line is the approximation from Eq. 27 for various values of WS's rewiring probability q .

6. General Observations

Our study is not exhaustive because of the large parameters space. The main constraint on the discussion in the present paper consist in the assumption that the two networks being added together have approximately equal number of vertices and edges. When this assumption is fulfilled we have obtained the following results.

The first observation is that the union with lattice-like WS, yields a composite network with clustering coefficient higher than the other combinations (Figure 8). This is because there are more triangles at the vertices in a WS network.

For example the clustering coefficient of $v_{i,er}$, $v_{i,ba}$ and $v_{i,ws}$ be proper lowest-term fraction a/b , c/d and e/f . The numerator is the number of triangles and the denominator is the number of triples. Since the clustering coefficient of $ER < BA < WS$, we have

$$\frac{a}{b} < \frac{c}{d} < \frac{e}{f}$$

Assuming few common edge, the clustering coefficient of the composite vertex $v_{i,c}$ is the ratio of the total number of triangles to the total number of triples. i.e. for $ER \cup BA$, the clustering coefficient is $(a + c)/(b + d)$. By the mediant inequality,

$$\frac{a}{b} < \frac{a + c}{b + d} < \frac{c}{d} < \frac{c + e}{d + f} < \frac{e}{f}$$

Hence unions with lattice-like WS (i.e. $(c + e)/(d + f)$) will yield higher clustering coefficient than other combinations. In fact given equal number of edges, the clustering coefficient of $WS \cup WS$ is greater than $WS \cup BA$, which is greater than $WS \cup ER$ (See Fig. 8).

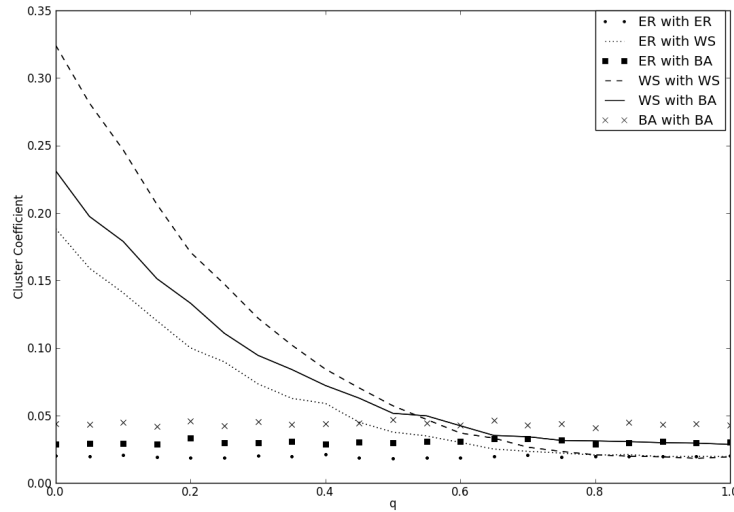


Figure 8: Parameters: $n = 1000$, $w = 10$, $m = 5$ and $p = k/(n - 1)$. The x-axis varies the rewiring probability (q) of WS. For small q , combinations with WS are high. Furthermore, the clustering coefficient of $WS \cup WS$ is greater than $WS \cup BA$, which is greater than $WS \cup ER$

From section 4, the degree distribution of unions with BA have power-law-like tail distribution. The intuition is that with high probability, the high degree vertices in BA have many common edges with the second network. Hence the change at the high degree vertices is small, allowing the tail distribution of the BA network to dominate. Fig. 9 compares the degree-rank plot of the different union combination.

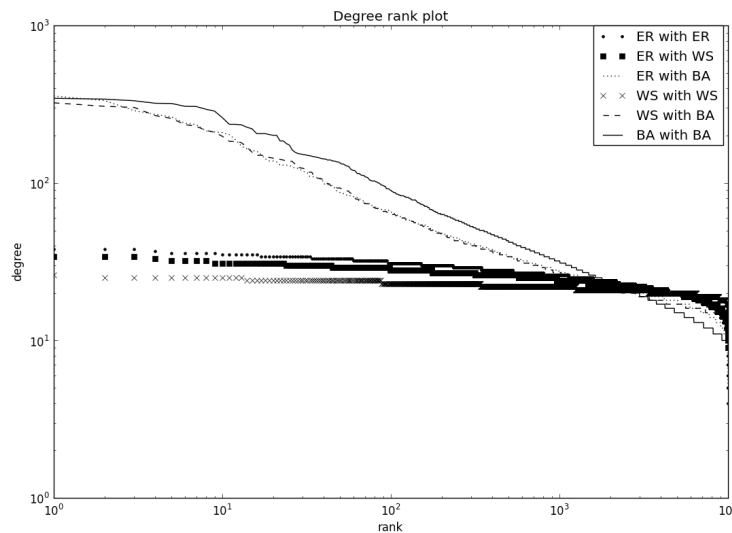


Figure 9: Same parameters as Fig. 8 with $q = 0.1$. Combinations with BA have power-law-like distribution. Note that $WS \cup BA$ is almost on the same line as $ER \cup BA$. Lastly, $BA \cup BA$ has a slightly different distribution than the other BA combinations.

7. Discussions and Future Work

It is common for real-world networks to have similar “layered dimensions”. In temporal networks where connections vary across time, each layer is a time snapshot of the connections of the same vertex set [39]. The edge union of the network snapshots is analogous to extending the time interval in the sampling, i.e. aggregation of a temporal stream of edges over a fixed number of time distance [40].

Aggregation helps to identify the optimal size of the time interval at which to resolve a temporal network. For example assume that vertex v_i and v_j are connected over time T with probability $1 - \exp(-pT)$ where $p \in [0, 1]$ is an arbitrary constant. As we extend (aggregate) the time to $2T$, Eq. (7) gives the edge probability that v_i and v_j are connected over time $2T$ is $1 - \exp(-2pT)$. This is the same result as we would get from Eq. 7 if we add the network at T to the one developed over time T to $2T$.

There are cases where the resultant network cannot be easily expressed like the above example or $ER \cup ER$. This paper address the problem to understand the statistical properties for such cases. In addition, the methods provides a framework for future studies on combinations of other families of network generators like Kronecker Graph [42], Exponential Random Graph [43] and Multifractal Network Generator [44].

Moreover the present investigation can be further developed to study a closely related model, which considers the percolation across one network layer of an effect induced by an adjacent layer. Such networks are known as interdependent networks [45, 46].

For example, two networks $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$ form an interdependent network if there exists an edge between a vertex in V_1 and V_2 . Let the set of such edges between G_1 and G_2 be E_3 . Hence an interdependent network is formally defined as $G(V_1 \cup V_2, E_1 \cup E_2 \cup E_3)$ if $E_3 \neq \emptyset$. However the resulting network can also be considered as the edge union of 3 networks on the same vertex set.

From the previous example, we can add isolated vertices to both networks such that they form the same vertex set $V = V_1 \cup V_2$. In this way we extend all three networks in order to consider them as define on a common vertex set. In this way we can obtain that the edge union of all 3 networks form the interdependent network $G(V, E_1 \cup E_2 \cup E_3)$.

These studies are beyond the scope of this paper but we hope that our paper may help to motivate further theoretical studies on more complex, yet closer to real-world network models.

8. Acknowledgment

We would like to thank Tim Evans and the anonymous referees for their comments and suggestions to improve this paper.

References

- [1] Mark Newman. *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA, 2010.
- [2] Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.
- [3] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.
- [4] Minas Gjoka, Carter T. Butts, Maciej Kurant, and Athina Markopoulou. Multigraph sampling of online social networks. *IEEE Journal on Selected Areas in Communications*, pages 1893–1905, 2011.
- [5] Przemyslaw Kazienko Piotr Brodka, Tomasz Filipowski. An introduction to community detection in multi-layered social network. *CCIS*, 278:185–190, 2012.
- [6] Nikolai Nefedov. Multiple-Membership Communities Detection and Its Applications for Mobile Networks. In *Applications of Digital Signal Processing, Dr. Christian Cuadrado-Laborde*, 2011.
- [7] Maciej Kurant and Patrick Thiran. Layered Complex Networks. *Physical Review Letters* PRL96,138701, 2006
- [8] Wenyuan Li, Shihua Zhang, Chun-Chi Liu and Xianghong Jasmine Zhou. Identifying Multi-Layer Gene Regulatory Modules from Multi-Dimensional Genomic Data In *Bioinformatics Advance Access*, 2012.
- [9] Marko A. Rodriguez. Social Decision Making with Multi-Relational Networks and Grammar-Based Particle Swarms. 2006.
- [10] Rodriguez, Marko A. and Shinavier, Joshua. Exposing Multi-Relational Networks to Single-Relational Network Analysis Algorithms. In *CoRR*, Vol. abs/0806.2274, 2008.
- [11] Michael Szell, Renaud Lambiotte and Stefan Thurner. Multirelational organization of large-scale social networks in an online world. 2010.
- [12] Darcy Davis, Ryan Lichtenwalter, Nitesh V. Chawla. Multi-Relational Link Prediction in Heterogeneous Information Networks 2011.
- [13] Cai, Deng and Shao, Zheng and He, Xiaofei and Yan, Xifeng and Han, Jiawei. Community Mining from Multi-relational Networks. In *PKDD'05*, pages 445–452, 2005.
- [14] Theus Hossmann and George Nomikos and Thrasyvoulos Spyropoulos and Franck Legendre. Collection and analysis of multi-dimensional network data for opportunistic networking research. In *Computer Communications*, Vol. 35, pages 1613–1625, 2012.
- [15] Michele Berlingerio, Michele Coscia, and Fosca Giannotti. Finding and characterizing communities in multidimensional networks. In *ASONAM'11*, pages 490–494, 2011.
- [16] Rossetti, Giulio and Berlingerio, Michele and Giannotti, Fosca. Scalable Link Prediction on Multidimensional Networks. In *ICDM Workshops'11*, page 979–986, 2011
- [17] Berlingerio, Michele and Coscia, Michele and Giannotti, Fosca and Monreale, Anna and Pedreschi, Dino. In *ASONAM'11*, pages 485–489, 2011
- [18] Lei Tang, Xufei Wang and Huan Liu. Community Detection in Multi-Dimensional Networks. 2009.
- [19] Lei Tang and Huan Liu. Uncovering Cross-Dimension Group Structures in Multi-Dimensional Networks. 2010.
- [20] Peter J. Mucha, Thomas Richardson, Kevin Macon, Mason A. Porter, Jukka-Pekka Onnela. Community Structure in Time-Dependent, Multiscale, and Multiplex Networks *Science* 328, 876, 2010.
- [21] Alessio Cardillo, Jesús Gómez-Gardeñes, Massimiliano Zanin, Miguel Romance, David Papo, Francisco del Pozo, Stefano Boccaletti. Emergence of network features from multiplexity *Scientific Reports* 2, 620, 2012
- [22] Osman Yağan; Virgil Gligor. Analysis of complex contagions in random multiplex networks *Phys Rev E* 86, 036103, 2012

- [23] Charles D. Brummitt, Kyu-Min Lee, and K.I. Goh. Multiplexity-facilitated cascades in networks *Phys Rev E* 85, 045102, 2012
- [24] Heath, Lenwood S. and Sioson, Allan A. Multimodal Networks: Structure and Operations In *Comput. Biol. Bioinformatics*, pages 321–332, 2009
- [25] Kühn, Daniela and Osthus, Deryk. Maximizing Several Cuts Simultaneously. *Combinatorics, Probability & Computing* pages 277–283, Vol 16, 2007.
- [26] W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33:452–473, 1977.
- [27] Myunghwan Kim and Jure Leskovec. Multiplicative attribute graph model of real-world networks. In Ravi Kumar and D. Sivakumar, editors, *Algorithms and Models for the Web-Graph - 7th International Workshop, WAW 2010, Stanford, CA, USA, December 13-14, 2010. Proceedings*, volume 6516 of *Lecture Notes in Computer Science*, pages 62–73. Springer, 2010.
- [28] M. P. H. Stumpf and P. J. Ingram. Probability models for degree distributions of protein interaction networks. *EPL (Europhysics Letters)*, 71(1):152, 2005.
- [29] P. Erdős and A. Rényi. On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.
- [30] E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [31] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393:440–442, June 1998.
- [32] A. L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [33] David M. Pennock, Gary W. Flake, Steve Lawrence, Eric J. Glover, and C. Lee Giles. Winners don’t take all: Characterizing the competition for links on the web. In *Proceedings of the National Academy of Sciences*, pages 5207–5211, 2002.
- [34] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin. Structure of growing networks with preferential linking. *Physical Review Letters*, 85(21):4633–4636, 2000.
- [35] SN Dorogovtsev, JFF Mendes, and AN Samukhin. Anomalous percolation properties of growing networks. *Physical Review E*, 64(6):066110, 2001.
- [36] A. Barrat and M. Weigt. On the properties of small-world network models. *EUROP.PHYS.J.B*, 13:547, 2000.
- [37] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks *PNAS*, 99:2566-2572 2002
- [38] Thomas Schank and Dorothea Wagner. Approximating Clustering Coefficient and Transitivity *Journal of Graph Algorithms and Applications*, vol.9, no. 2, 2005
- [39] Petter Holme and Jari Saramki. Temporal networks. *Physics Reports*, 519(3):97 – 125, 2012.
- [40] Rajmonda Sulo Caceres, Tanya Y. Berger-Wolf, and Robert Grossman. Temporal scale of processes in dynamic networks. *ICDM Workshops*, pages 925–932. IEEE, 2011.
- [41] Klemm, Konstantin and Eguíluz, Víctor M. Growing scale-free networks with small-world behavior. *Phys. Rev. E*, 65, 057102, 2002.
- [42] Leskovec, Jure and Chakrabarti, Deepayan and Kleinberg, Jon M. and Faloutsos, Christos. Realistic, Mathematically Tractable Graph Generation and Evolution, Using Kronecker Multiplication. *Lect. Notes. Comput. Sci.* 3721:133145, 2005
- [43] Ove Frank and David Strauss. Markov Graphs *J. Am. Stat. Assoc.* 81:832842, 1986
- [44] Gergely Palla, Lszl Lovszb, Tams Vicseka, Multifractal network generator *PNAS* 107:7640-7645, 2010
- [45] Jianxi Gao, Sergey V. Buldyrev, H. Eugene Stanley and Shlomo Havlin. Networks formed from interdependent networks *Nature Physics* 8: 40-48 (2012)
- [46] Wei Li, Amir Bashan, Sergey V. Buldyrev, H. Eugene Stanley, and Shlomo Havlin. Cascading Failures in Interdependent Lattice Networks: The Critical Role of the Length of Dependency Links. *Phys. Rev. Lett.* 108:228702 (2012))

Appendix A. Simulation Details

Simulations are implemented in Python 2.7. This is to make use of the existing numerical and network libraries like numpy and NetworkX to minimize implementation errors. The trade off is slower runtime, resulting in smaller sample size.

For simulations, two networks with the same number of vertices are generated with NetworkX network library. The size of the vertex set is chosen such that we can maximize the sample size while computationally feasible (less than 1 hour). For instance the size of the vertex set in Fig. 2 is 10000 as the iterative method (equation 12) is computationally expensive. In contrast, to generate figure 3, it is possible to use vertex set of size 100000.

The size of the edge set also determines the computation costs. In this paper, we chose the minimum size edge set such that the generated networks are connected. The decision for studying connected network is arbitrary from an analysis point of view. However it is useful if we want to study other metrics like average shortest distance in the future.

To ensure ER is a connected network, we need the probability, p , that two vertices are connected to satisfy $p > \ln n/n$. Furthermore in this paper we only consider the case where both networks have approximately equal number of edges. Thus for WS and BA to have approximately equal edge set size as ER, $w \approx (n - 1)p$ and $m \approx (n - 1)p/2$ respectively.

We note that unequal size edge sets introduce additional complexity to the study, since the larger edge set will dominate the resultant network. Ignoring this aspect significantly reduces the complexity of the computation.