Edge union of networks on the same vertex set

Chuan Wen Loe and Henrik Jeldtoft Jensen

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

E-mail: c.loe11@imperial.ac.uk and h.jensen@imperial.ac.uk

Received 25 December 2012, in final form 3 May 2013
Published 30 May 2013
Online at stacks.iop.org/JPhysA/46/245002

Abstract
Random network generators such as 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, $G(V,E_1 \cup E_2)$.

PACS numbers: 02.10.Ox, 89.75.Hc

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 network generators are Erdős–Rényi (ER), Watts–Strogatz (WS) and Barabási–Albert (BA) models. These models generate networks with only one type of edge and hence may be thought of as relating to agents interacting under a homogeneous relation [1–3].

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. For example, think of social interactions between people: one type of relation may consist of family bonds while another could be shared interests, such as chess, the 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 [5–8], multi-relational [9–13], multi-dimensional [14–19], multiplex [11, 20–23], 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 composed of layers of different relations (dimensions). Each layer has the same set of vertices, and these vertices are connected through different relationships.

There are various proposals to solve classic problems such as 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 the statistical properties of the union of the different types of edges in this paper. 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 eight 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 eight different types of relationships. One of such relations is the association in and between academic classes at the university. The eight different relationships are then combined as edges (of different types) on the same vertex set. Other real-world applications of these kinds of networks can be found in [27] and [28].

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

This paper is a study of some of the statistical properties of a random network that is formed by the union of two random networks. Different families of a 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 summarizes the construction of ER, WS and BA 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 the 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 the vertex set and the edge set, respectively.

2.1. Erdős–Rényi

A realization of an ER network [29] is selected with equal probability from the set of all possible graphs with \( n = |V| \) vertices and \( |E| \) edges. However, generating a huge random ER 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 each vertex pair is connected with the probability \( p \), where \( p \approx |E|/(\binom{n}{2}) \). The slight difference is that ER has precisely \( |E| \) edges, while the Gilbert graph has approximately \( |E| \) edges with a high probability.

2.2. Watts–Strogatz

A WS network [31], \( W_{w,q} \), is parameterized by \( w \) and \( q \) for the mean degree\(^1\) 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 the probability \( q \). The rewiring replaces \( \{v_i, v_j\} \) with \( \{v_i, v_b\} \), where \( v_b \) is chosen uniformly in \( V \) and the resultant network remains a simple network.

\(^1\) We replace the commonly used variable \( k \) with \( w \) to avoid confusion with the variable \( k \) commonly used in the degree distribution function.
The key property of the WS network is that it varies between a regular ring lattice \((q = 0)\) and an ER network \((q = 1)\). As \(q \to 1\), WS can be expressed as an ER graph, i.e. \(W_{q=1} = G_{w/(n-1)}\).

2.3. Barabási–Albert

A BA 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 \(\text{deg}(v_i)\) as the degree of vertex \(v_i\). The probability that the new vertex is connected to vertex \(v_i\) is given by

\[
    p_i = \frac{\text{deg}(v_i)}{\sum_j \text{deg}(v_j)}.
\]

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: \(C\)). A network refers to either an ER, WS or BA network. The composition of all six pairwise edge unions of these networks is \([\text{ER, ER}], [\text{ER, WS}], [\text{ER, BA}], [\text{WS, WS}], [\text{WS, BA}]\) and \([\text{BA, BA}]\).

**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 \(C(V, E_1 \cup E_2)\).

**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 \(E_1 \cap E_2\).

**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 \(\epsilon\) common edges is

\[
    P(|E_1 \cap E_2| = \epsilon) = \mathcal{H}(\epsilon; \binom{n}{2}, |E_1|, |E_2|).
\]
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.

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

We note that equation (1) is the probability that there are $\epsilon$ 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):

$$E[|E_1 \cap E_2|] = E[H] = \frac{|E_1| \cdot |E_2|}{\binom{n}{2}}.$$

In fact, we can generalize 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 of 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 $\epsilon$ common $c$-cliques is

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

**Proof.** Choose $|K_2|$ $c$-cliques out of the 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 $\epsilon$ blue $c$-cliques is defined by the hypergeometric function.

**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):

$$E[|K_1 \cap K_2|] = E[H] = \frac{|K_1| \cdot |K_2|}{\binom{n}{c}}.$$

To compute the degree distribution of the composite network accurately, we have to minimize any double counting of the common edges. For example, in figure 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 $a$. 

4
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 $d_2 = \deg(v_{i,G^2})$; then the probability that there are $\epsilon$ common edges between $v_{i,G^1}$ and $v_{i,G^2}$ is

$$P_{\epsilon}(\epsilon | d_1, d_2) = H(\epsilon; n - 1, d_1, d_2).$$

(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. □

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 $d_2 = \deg(v_{i,G^2})$; then the expected number of common edges between $v_{i,G^1}$ and $v_{i,G^2}$ is

$$E[H] = \frac{|d_1| \cdot |d_2|}{n - 1}.$$ 

(6)

In this paper, we only consider the case where both networks have approximately the same edge set size. Unequal edge set size introduces an 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 $G^1_p$ and $G^2_p$ be two independent Gilbert graphs. The probability that an edge does not exist in both $G^1_p$ and $G^2_p$ 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

$$G^1_p \cup G^2_p = G_{1,p} - (1 - p)(1 - p').$$

(7)

4.2. Erdős–Rényi with Watts–Strogatz

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

For $q \to 0$, 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

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

(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 $P^E(k)$ and $P^W(k)$, respectively. To account for the common edges for $P^C(k)$, we have to consider the following.

(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 $\epsilon$ common edges between $v_{er}$ and $v_{ws}$.
From lemma 3.3, the probability that there are \( \epsilon \) common edges between \( v_{er} \) and \( v_{ws} \) is 
\[ P_{ce}(\epsilon | j, k + \epsilon - j). \]
Thus the improvement on equation (8) is
\[ P^C(k) \sim \sum_{\epsilon} \sum_{j} P_{ce}(\epsilon | j, k + \epsilon - j) P^W(j) P^G(k + \epsilon - j). \]  
(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
\[ P^C(k) \sim \sum_{\epsilon} P_{ce}(\epsilon | w, k + \epsilon - w) P^G(k + \epsilon - w). \]  
(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 led to various generalized BA models [33–35] 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:
\[ P^C(k) \sim \sum_{\epsilon} \sum_{j} P_{ce}(\epsilon | j, k + \epsilon - j) P^G(j) P^B(k + \epsilon - j), \]  
(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 equation (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 \( \binom{m}{2}p - nmp \) uncommon 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 1 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 1 if the new edge attaches to a degree \( k \) vertex. This leads to the following equation:
\[ u(k, t + 1) = u(k, t) + u(k - 1, t)/n - u(k, t)/n. \]  
(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 \( k \):
\[ \frac{\partial u}{\partial t} + \frac{1}{n} \frac{\partial u}{\partial k} = 0. \]  
(13)

Using the initial condition \( u(k, 0) = P^B(k) \cdot n \sim 2m^2k^{-3} \) \( (P^B(k) = 0 \text{ for } k < m) \), we can solve for \( u(k, t) \) at \( t = 2 \cdot \binom{m}{2}p - nmp \approx np(n - m) \) (twice the number of uncommon edges
because there are two vertices that change at each time step) and find
\[ P^C(k) = u(k, t) = \frac{2nm^2}{(k + 2n(p - m^2))^3}. \]  
(14)

In figure 2, we compare this asymptotic expression to the analytic expression derived by
iterating equation (12) and to simulation results. The continuum approximation of equation (13)
Figure 2. Parameters: $n = 75,000$, $m = 10$ and $p = 2m/(n-1)$ (chosen such that both networks have an equal number of edges). The solid line plots the iterative method, equation (12). The dotted line plots the closed form (equation (14)). The crosses represent the distribution obtained from simulations. The degree-distribution log-binning of the crosses is plotted with squares.

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.

4.4. Watts–Strogatz with Watts–Strogatz

Let $W_{w,q}$ and $W_{w',q'}$ be two WS networks. If $q \approx q' \to 1$, the two WS can be expressed as two ER (section 4.1). In contrast, the limits $q \to 1$ and $q' \to 0$ are 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 the WS and BA graphs, respectively. There are $wn/2$ edges in the WS and $nm$ edges in BA. It is most interesting to consider the case when the graphs have an approximately equal number of edges, i.e. $w \approx 2m$. We ignore the case $q \to 1$, as it is similar to section 4.3.

For $q \to 0$, the degree distribution of $C$ is straightforward. 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 contribute to an increase in the degree of the vertices of BA by $w$. This gives the approximation for

$$P_C(k) \sim 2m^2(k-w)^{-3}. \quad (15)$$

The degree distribution can be refined by considering the common edges between WS and BA. However, the expression will be in an open form and provides little insight. From simulations ($q \to 0$), equation (15) is sufficient to approximate the distribution (figure 3).
Figure 3. Parameters: $n = 100,000$, $w = 10$, $m = 5$ and $q = 0.1$. The crosses represent the degree distribution obtained from simulations. The dotted line plots equation (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 retains the features. Similar to equation (9), the probability density function is the combined probability of

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

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 $k$.

Unlike in section 4.3, we are not able to apply the same method to the union of two BA networks. In section 4.3, we began with a BA network and then we added uncorrelated random edges onto it. In this 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 have blue edges. We begin with the BA with red edges and assume that 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 that of BA/BA. The union of BA/BA has more high-degree vertices than that 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 a few common edges/triangles in the union,
which can be verified with corollaries 3.1 and 3.2. This simplification is reasonable for sparse networks and it gives us simple relationships similar to equation (19).

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 \( p, q \to 0 \). Define \( T(v_i) \) as the number of subgraphs with three edges and three vertices (triangles) which all share the vertex \( v_i \). Recall the clustering coefficient equation for a network:

\[
\text{Clustering coefficient} = \frac{1}{n} \sum_{i=0}^{n} \text{Clustering coefficient of } v_i
\]

\[
= \frac{1}{n} \sum_{i=0}^{n} T(v_i) \left( \frac{\text{deg}(v_i)}{2} \right)^2.
\]

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 \( C \). Since \( p \approx q \) and \( p, q \to 0 \), 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 \( C \) will be similar. We therefore have for small \( w^1 \):

\[
\text{Clustering coefficient of } C = \frac{1}{n} \sum_{i=0}^{n} \text{Clustering coefficient of vertex } v_{i,c}
\]

\[
= \frac{1}{n} \sum_{i=0}^{n} T(v_{i,c}) \left( \frac{\text{deg}(v_{i,c})}{2} \right).
\]
Figure 5. Edge union of two WS networks for $n = 1000$, $w' = w = 20$ and $q' = q$. The dots are from equation (19) and the crosses are from the simulations.

Because $w$ is small ($w \ll n$), we assume that there are a few common edges. Hence, $T(v_i,c) = T(v_i,w) + T(v_i,w') \approx 2T(v_i,w)$ and $\deg(v_i,c) \approx 2\deg(v_i,w)$. Then,

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

$$\approx \frac{1}{n} \sum_{i=0}^{n} \frac{1}{2} \left( \frac{T(v_i,w)}{\frac{1}{2} \deg(v_i,w)} \right)$$

$$\approx \frac{1}{2} \text{Clustering coefficient of } W^1. \quad (19)$$

Figure 5 compares the results of equation (19) with an actual simulation of the edge union of two WS networks. The dots plot the expected clustering coefficient from equation (19) by halving the clustering coefficient of one of the WS networks. The crosses represent the actual clustering coefficient obtained from simulations. The details of the simulations are explained in the appendix.

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 (figure 6):

$$\text{Clustering coefficient of } C = \frac{1}{n} \sum_{i=0}^{n} \text{Clustering coefficient of vertex } v_i,c$$

$$= \frac{1}{n} \sum_{i=0}^{n} \frac{T(v_i,c)}{\frac{1}{2} \deg(v_i,c)}. $$
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 equation (23). The dotted line is the upper bound from equation (24).

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

Clustering coefficient of $\mathcal{C} \approx \frac{1}{n} \sum_{j=0}^{n} \frac{T(v_{i,ws}) + T(v_{i,ba})}{\binom{\deg(v_{i,c})}{2}}$

\begin{equation}
\geq \frac{1}{n} \sum_{j=0}^{n} \frac{T(v_{i,ws})}{\binom{\deg(v_{i,c})}{2}}.
\end{equation}

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 $T(v_{i,ws})$ we 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 $\alpha$ be the average number of triangles for any vertex in WS given $q$. Thus,

\begin{equation}
\text{Clustering coefficient of } \mathcal{C} \geq \frac{1}{n} \sum_{j=0}^{n} \frac{T(v_{i,ws})}{\binom{\deg(v_{i,c})}{2}} \approx \frac{\alpha}{n} \sum_{j=0}^{n} \frac{1}{\binom{\deg(v_{i,c})}{2}} = \alpha \sum_{k=0}^{n} \binom{k}{2} \frac{1}{\binom{\deg(v_{i,c})}{2}}.
\end{equation}

From [36], the clustering coefficient of WS decreases at the rate of $(1-q)^3$ 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:

\begin{equation}
\text{Clustering coefficient of } \mathcal{C} \geq \frac{(1-q)^3}{4} \sum_{k=0}^{n} \binom{k}{2} \frac{1}{\binom{\deg(v_{i,c})}{2}}.
\end{equation}
To obtain an upper bound on the clustering coefficient, we can round up the clustering coefficient contribution originating from BA, i.e.

\[
\text{Clustering coefficient of } C \approx \frac{1}{n} \sum_{i=0}^{n} \binom{T(v_{i,ws}) + T(v_{i,ba})}{\frac{\deg(v_{i,c})}{2}}
\]

\[
= \frac{1}{n} \sum_{i=0}^{n} \frac{T(v_{i,ws})}{\frac{\deg(v_{i,c})}{2}} + \frac{1}{n} \sum_{i=0}^{n} \frac{T(v_{i,ba})}{\frac{\deg(v_{i,c})}{2}}
\]

\[
\leq \frac{1}{n} \sum_{i=0}^{n} T(v_{i,ws}) + \frac{1}{n} \sum_{i=0}^{n} T(v_{i,ba})
\]

\[
\approx \frac{(1 - q)^3 w^2}{4} \sum_{k=0}^{n} \frac{1}{\binom{k}{2}} + \text{Clustering coefficient of } B_m. \quad (24)
\]

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 \( G_p \). Hence, the clustering coefficient of \( G_p \) can be estimated like any ER network.

However, the ‘clustering coefficient’ of ER is often measured by an alternative similar metric, the transitivity ratio \([37]\). It is the ratio of the number of triangles to the number of connected triples:

\[
\text{Transitivity ratio of } G_p = \frac{\binom{n}{3} p^3}{\binom{n}{2} p^2} = p. \quad (25)
\]

5.4. Clustering coefficient of other combinations

In the following section, we show that the clustering coefficients of ER\( \cup \)BA and BA\( \cup \)BA are small in general (figure 8). For this reason, we do not derive any analytical results in this paper.

WS\( \cup \)ER is a more interesting combination, but it is difficult to deduce the clustering coefficient of the union from the known results for the individual families of networks. The clustering coefficient of WS is measured by equation (17) and the ‘clustering coefficient’ (transitivity ratio) of ER is measured by equation (25). Since these metrics are different in general \([38]\), we are only able to deduce a weak estimate of the union (figure 7).

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

With the probability \( p \) from ER, the number of neighbour pairs at \( v_{i,c} \) that are connected to the edges of ER (and not from WS) is

\[
\binom{\deg(v_{i,c})}{2} - T(v_{i,ws}) p.
\]

Hence,

\[
\text{Clustering coefficient of } C \approx \frac{1}{n} \sum_{i=0}^{n} \frac{\binom{\deg(v_{i,c})}{2} - T(v_{i,ws}) p + T(v_{i,ws})}{\binom{\deg(v_{i,c})}{2}}
\]

\[
= \frac{1}{n} \sum_{i=0}^{n} \frac{\deg(v_{i,c})}{2} p + T(v_{i,ws}) (1 - p)
\]

\[
= p + \frac{(1 - p) n}{n} \sum_{i=0}^{n} T(v_{i,ws}).
\]
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 equation (26) for various values of $WS$’s rewiring probability $q$.

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 $WS \cup ER$.

Together with the approximation from equations (20) and (23),

\[
\text{Clustering coefficient of } C \approx p + \frac{1 - p}{n} \sum_{j=0}^{n} \left( \frac{(1 - q)^j w^2}{2} \right) \sum_{k=0}^{\min(n,j)} x(k) \frac{1}{k!}. \]  

(26)
6. General observations

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

The first observation is that the union with lattice-like WS yields a composite network with the clustering coefficient being 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 coefficients of $v_{i,er}$, $v_{i,ba}$ and $v_{i,ws}$ be the proper lowest term fractions $a/b$, $c/d$ and $e/f$, respectively. 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 edges, 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,

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

Hence, unions with lattice-like WS (i.e. $(c + e)/(d + f)$) will yield a higher clustering coefficient than other combinations. In fact, given the 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 figure 8).

From section 4, the degree distribution of unions with BA has a power-law-like tail distribution. The intuition is that with a 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. Figure 9 compares the degree–rank plot of the different union combination.
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. the 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 the vertices \( v_i \) 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 \( 2T \), equation (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 obtain from equation (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 \( \text{ER} \cup \text{ER} \). This paper addresses the problem to understand the statistical properties for such cases. In addition, the methods provide a framework for future studies on the combinations of other families of network generators such as the 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 three 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 defined on a common vertex set. In this way, we can observe that the edge union of all three 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.

Acknowledgments

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

Appendix. Simulation details

Simulations are implemented in Python 2.7. They use the existing numerical and network libraries such as NumPy and NetworkX to minimize the implementation errors. The trade off is slower runtime, resulting in a smaller sample size.

For simulations, two networks with the same number of vertices are generated with the NetworkX network library. The size of the vertex set is chosen such that we can maximize the sample size when it is computationally feasible (less than 1 h). For instance, the size of the vertex set in figure 2 is 10 000, as the iterative method (equation (12)) is computationally expensive. In contrast, to generate figure 3, it is possible to use the vertex set of size 100 000,
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 to study the connected network is arbitrary from an analysis point of view. However, it is useful if we want to study other metrics like the 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 an approximately equal number of edges. Thus for WS and BA to have an 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.

References

[1] Newman M 2010 Networks: An Introduction (New York: Oxford University Press)
[2] Easley D and Kleinberg J 2010 Networks, Crowds, and Markets: Reasoning About a Highly Connected World (New York: Cambridge University Press)
[3] Jackson M O 2008 Social and Economic Networks (Princeton, NJ: Princeton University Press)
[4] Gjoka M, Butts C T, Kurant M and Markopoulou A 2011 Multigraph sampling of online social networks IEEE J. Sel. Areas Commun. 29 1893–905
[5] Kazienko P, Brodka P and Filipowski T 2012 An introduction to community detection in multi-layered social network Information Systems, E-Learning, and Knowledge Management Research (Communications in Computer and Information Science vol 278) (Berlin: Springer) pp 185–90
[6] Nefedov N 2011 Multiple-membership communities detection and its applications for mobile networks Applications of Digital Signal Processing ed C Cuadrado-Laborte (New York: InTech) (doi:10.5772/26469)
[7] Li W, Zhang S, Liu C-C and Zhou X J 2012 Identifying multi-layer gene regulatory modules from multi-dimensional genomic data Bioinformatics 28 2458–66
[8] Rodriguez M A 2007 Social decision making with multi-relational networks and grammar-based particle swarms HCSS 2007-40th Annual Hawaii Int. Conf. on System Sciences pp 39–49
[9] Rodriguez M A and Shinavier J 2009 Exposing multi-relational networks to single-relational network analysis algorithms J. Informetrics 4 29–41
[10] Szell M, Lambiotte R and Thurner S 2010 Multirelational organization of large-scale social networks in an online world Proc. Natl Acad. Sci. USA 107 13636–41
[11] Davis D, Lichtenwalter R and Chawla N V 2011 Multi-relational link prediction in heterogeneous information networksASONAM’11: Int. Conf. on Advances in Social Networks Analysis and Mining pp 281–8
[12] Cai D, Shao Z, He X, Yan X and Han J 2005 Community mining from multi-relational networks PKDD’05: Proc. 9th Eur. Conf. on Principles and Practice of Knowledge Discovery in Databases pp 445–52
[13] Hossmann T, Nomikos G, Spyropoulos T and Legendre F 2012 Collection and analysis of multi-dimensional network data for opportunistic networking research Comput. Commun. 35 1613–25
[14] Berlingerio M, Coscia M, Giannotti F, Monreale A and Pedreschi D 2011 Scalable link prediction on multidimensional networks ICDM’11: IEEE 11th Int. Conf. on Data Mining pp 979–86
[15] Cardillo A, Gómez-Gardeñes J, Zanin M, Romance M, Papo D, del Pozo F and Boccaletti S 2012 Emergence of network features from multiplexity Sci. Rep. 2 620
[16] Yağan O and Gligor V 2012 Analysis of complex contagions in random multiplex networks Phys. Rev. E 86 036103
[23] Brummitt C D, Lee K M and Goh K I 2012 Multiplexity-facilitated cascades in networks Phys. Rev. E 85 045102
[24] Heath L S and Sioson A A 2009 Multimodal networks: structure and operations Comput. Biol. Bioinform. 6 321–32
[25] Daniela K and Osthus D 2007 Maximizing several cuts simultaneously Comb. Probab. Comput. 16 277–83
[26] Zachary W W 1977 An information flow model for conflict and fission in small groups J. Anthropol. Res. 33 452–73
[27] Kim M and Leskovec J 2010 Multiplicative attribute graph model of real-world networks Algorithms and Models for the Web-Graph: Proc. 7th Int. Workshop (Stanford, CA, USA, 13–14 Dec.) (Lecture Notes in Computer Science) vol 6516 ed R Kumar and D Sivakumar (Berlin: Springer) pp 62–73
[28] Stumpf M P H and Ingram P J 2005 Probability models for degree distributions of protein interaction networks Europhys. Lett. 71 152
[29] Erdős P and Rényi A 1959 On random graphs: I Publ. Math. 6 290–7
[30] Gilbert E N 1959 Random graphs Ann. Math. Stat. 30 1141–4
[31] Watts D J and Strogatz S H 1998 Collective dynamics of ‘small-world’ networks Nature 393 440–2
[32] Barabási A L and Albert R 1999 Emergence of scaling in random networks Science 286 509–12
[33] Pennock D M, Flake G W, Lawrence S, Glover E J and Giles C L 2002 Winners don’t take all: characterizing the competition for links on the web Proc. Natl Acad. Sci. USA 99 5207–11
[34] Dorogovtsev S N, Mendes J F F and Samukhin A N 2000 Structure of growing networks with preferential linking Phys. Rev. Lett. 85 4633–6
[35] Dorogovtsev S N, Mendes J F F and Samukhin A N 2001 Anomalous percolation properties of growing networks Phys. Rev. E 64 066110
[36] Barrat A and Weigt M 2000 On the properties of small-world network models Eur. Phys. J. B 13 547
[37] Newman M E J, Watts D J and Strogatz S H 2002 Random graph models of social networks Proc. Natl Acad. Sci. USA 99 2566–72
[38] Schank T and Wagner D 2005 Approximating clustering coefficient and transitivity J. Graph. Algorithms Appl. 9 265–75
[39] Holme P and Saramäki J 2012 Temporal networks Phys. Rep. 519 97–125
[40] Caceres R S, Berger-Wolf T Y and Grossman R 2011 Temporal scale of processes in dynamic networks ICDM’11: IEEE 11th International Conference on Data Mining Workshops pp 925–32
[41] Klemm K and Eguíluz V M 2002 Growing scale-free networks with small-world behavior Phys. Rev. E 65 057102
[42] Leskovec J, Chakrabarti D, Kleinberg J M and Faloutsos C 2005 Realistic, mathematically tractable graph generation and evolution, using Kronecker multiplication Lect. Notes Comput. Sci. 3721 133–45
[43] Frank O and Strauss D 1986 Markov graphs J. Am. Stat. Assoc. 81 832–42
[44] Palla G, Lovász L and Vicsek T 2010 Multifractal network generator Proc. Natl Acad. Sci. USA 107 7640–45
[45] Gao J, Buldyrev S V, Stanley H E and Havlin S 2012 Networks formed from interdependent networks Nature Phys. 8 40–8
[46] Li W, Bashan A, Buldyrev S V, Stanley H E and Havlin S 2012 Cascading failures in interdependent lattice networks: the critical role of the length of dependency links Phys. Rev. Lett. 108 228702