3 Generative Network Models and Network Evolution

Understanding the processes that give rise to networks gives us a better grasp of why we see the networks we do, where we might expect to find them, and how we might expect them to change over time. One way to achieve this is to generate networks of our own. Such simulated networks allow us to build networks based on detailed principles. We can then ask how networks derived from these principles behave. This will give us insight into how our observed networks may be generated by similar principles.

For example, we might hypothesize that the structure of an observed network is based on some assortative property. To evaluate this hypothesis, we can create networks with and without this property and then measure to what extent these networks resemble the observed network. More generally, if we believe a network grows based on a certain principle, we can simulate networks that grow with or without that principle and see how they are like or not like the one we observe.

It is useful to have an understanding of multiple generative network processes by which we can compare the networks we study. Such a toolbox will not only allow us to test our assumptions, but will help us recognize some of the many dimensions along which networks vary.

This chapter focuses on mechanisms for generating networks that are agnostic with respect to the behavioral system under study. They therefore provide useful abstractions which we can use to understand a range of real-world phenomena.

To rapidly develop a conceptual understanding without getting bogged down in unnecessary details, the descriptions here focus on the generative processes necessary to give rise to simple undirected and unweighted networks. Directed and weighted versions of these processes can be derived from this basic framework. I provide some details about how to go about this at the end of the chapter.

3.1 Erdös–Rényi Random Graphs

Suppose your social network was made up of a random sample from all the people in the world. Such a network would be unlike anything you currently experience. Your mother might be from Canada and your father from Jamaica, but they wouldn't know each other and they wouldn't know your grandparents either. We might imagine a dystopian society where everything happens online and all relationships are independently and randomly assigned by the state. This, of course, doesn't make any sense for the typical social network. Most of our new relationships depend on the people we already know. In the real world, networks are like onions. New layers of edges depend on old layers of edges – recall that a friend of my friend is likely to be my friend as well. As a result, edges tend

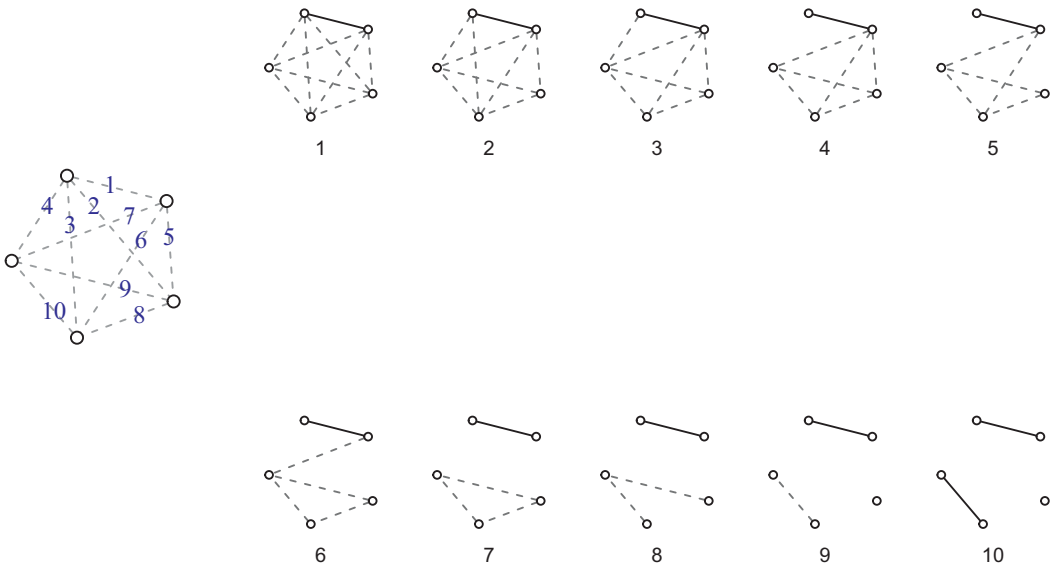

Figure 3.1 The construction of an ER random graph. Starting with the empty graph on the left, we evaluate each possible edge (indicated with dotted lines) and with probability p assign it a value of 1 (solid), or otherwise 0. In this network with five nodes, there are ten edges to evaluate. Each is evaluated in turn in the panels labeled 1 through 10, creating an edge with probability 0.2. Network 10 is one possible output of this random graph generation.

to depend on one another. What would happen if this weren't the case – if all edges were formed independently?

The **Erdös–Rényi (ER) random graph** is an effort to answer this question. First introduced by Erdös and Rényi (1959), it is easy to generate and study and is commonly used as a point of comparison for real-world networks. It is therefore a kind of fruit-fly of the network world. Most network software and the online code for this book will generate these and other network algorithms without complaint, but it is useful to understand the underlying process.

Given a set of nodes, N, an ER random graph places an edge between each pair of nodes with probability p. Each time we generate an ER network, we can get a different answer. But if we do it repeatedly, we will get a distribution of answers that reflect how likely different network structures are if all edges were independently created. See Figure 3.1 for an example.

Generally speaking, comparisons with observed networks involve sampling many hundreds of simulated random graphs and evaluating the observed network with respect to the distribution of simulated networks. Figure 3.2 provides some intuition for the variation in metrics produced by an ER random graph with a fixed edge probability.

A second way of generating ER random graphs uses a fixed number of edges and distributes them randomly. For an undirected network, this can be done by sampling E cells from the adjacency matrix without replacement and adding an edge to each cell sampled. A fixed edge number always produces the same density ($\rho = \frac{2E}{N(N-1)}$). Its distribution of other metrics are also lower in variance because in an ER random graph with probabilistic edges the total number of edges can vary.

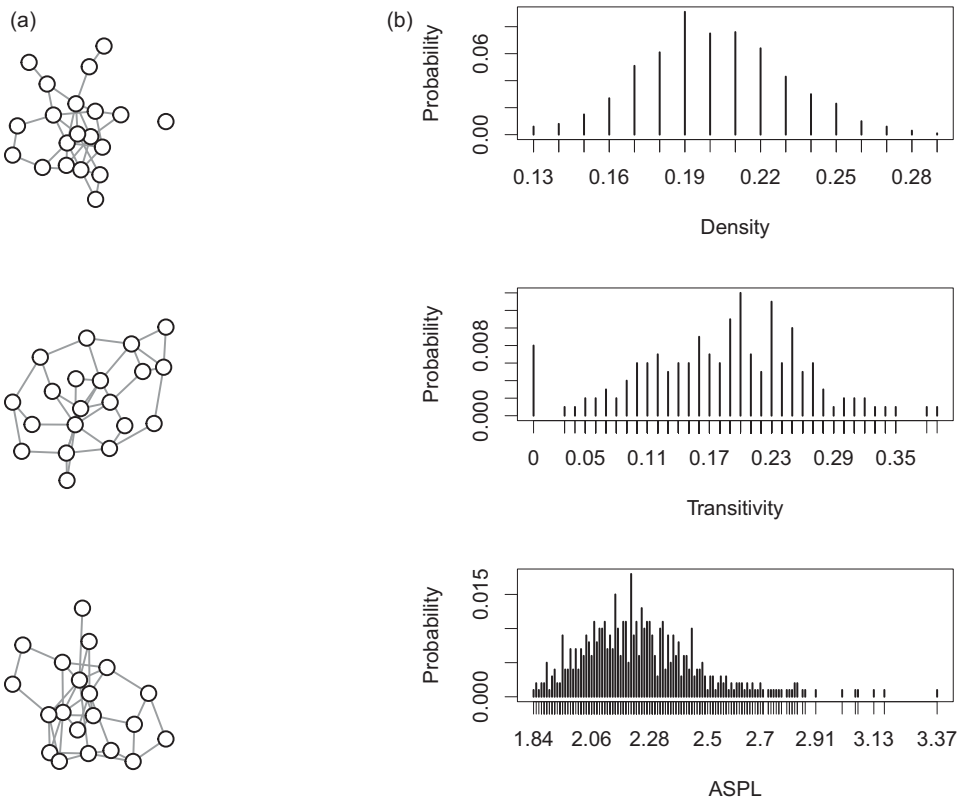

Figure 3.2 Erdös–Rényi random graphs with $N = 20$ and $p = 0.2$. (a) Three networks showing random examples. (b) Histograms of the density, transitivity, and average shortest path length (ASPL) observed for 1,000 random graphs.

3.2 Small-World Networks

If you want to get away from it all, you might travel to one of the world's most remote places, like the Amazonian city of Iquitos in Peru. There are no roads leading to it. If you don't get there by plane from Lima, you'll require a multiday journey by river. If you're in Iquitos, make sure to say "hi" to my old friend Josh, and tell him I sent you.

The term "small world" captures two ideas. The first is that most of the people we know also know each other; we live in local communities of well-connected individuals. The second is that when we meet someone new – even far from home – they often tend to know people we know. In a small world, path lengths between random pairs of nodes are surprisingly short.

To formally capture this notion, Watts & Strogatz (1998) introduced small-world networks. **Small-world networks** have local clustering coefficients that are higher than one would expect from an ER random graph of the same size and density, yet the average path length is relatively similar. In other words, they offer both local clustering and distant connectivity.

There are a number of ways to measure small-world properties, discussed in Chapter 7. Here I will focus on how small-world networks are constructed.

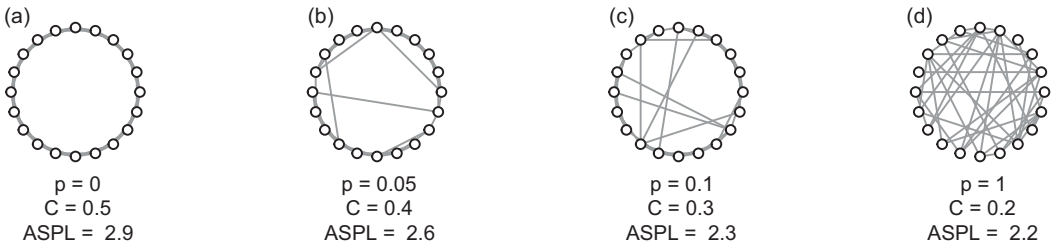

Figure 3.3 The construction of small-world networks. The starting lattice (a) has twenty nodes and each node is connected to its four closest neighbors. Moving from (a) to (b), the probability of a rewiring, p, increases. The clustering coefficient and average shortest path length are shown for the resulting network.

The method described by Watts and Strogatz (1998) involves starting with a regular lattice – all nodes of the same degree – and then rewiring edges with some probability. This process does not alter the network density. However, Watts and Strogatz (1998) showed that the random rewiring reduces the average shortest path length more rapidly than it reduces the clustering coefficient.

Figure 3.3 shows the transition across networks as rewiring probabilities increase. With $p = 0$ the ring maintains its regularity, each node has the same degree, and each is equidistant from all other nodes. With $p = 1$ the network becomes fully disordered. In between, networks are more clustered than a random equivalent even though their path lengths stay relatively short.

3.3 Preferential Attachment

According to the economist Thomas Picketty, author of *Capital in the Twenty-First Century* (2014), rising inequality is an inevitable outcome of capitalism. The rich get richer when the return on capital exceeds economic growth, and they will continue to until an equal and opposite force prevents them. *The Bible* put it more simply in the Gospel of Matthew: "to them that have shall be given."

Social systems with rich-get-richer dynamics tend to be strongly skewed. Some individuals enjoy popularity or wealth that far outstrips that of their nearest neighbor, where most people enjoy little or none of it. Barabási and Albert's (1999) **preferential attachment model** captures this for networks: as new nodes are added to the network, they preferentially attach to nodes that are already well connected. This leads the degree of some nodes to skyrocket, whereas most others fizzle at the smaller end of the scale.

To achieve this, the preferential attachment starts with one node. Then it adds nodes, one at a time. Each new node forms edges with nodes that are already part of the network. The probability of attaching to an existing node in a network of size N is proportional to the degree, k, of the existing node in the following way:

$$P(i) = \frac{k_i^{\alpha} + a}{\sum_{j \in N}^{N} k_j^{\alpha} + a}$$

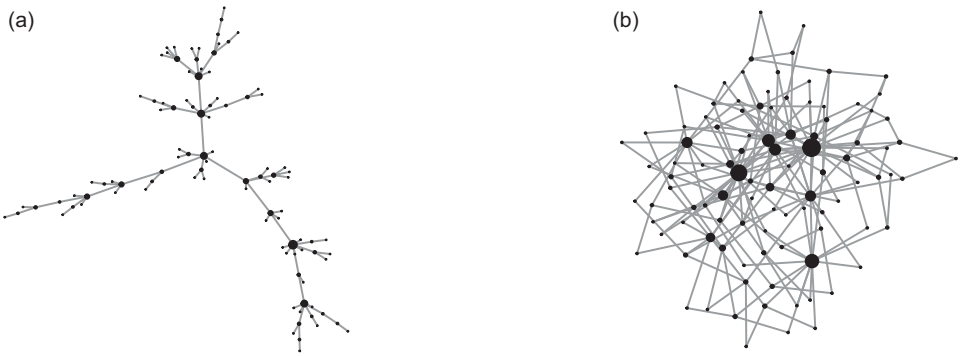

Figure 3.4 Preferential attachment networks. The network adds nodes as it grows, with new nodes attaching to m existing nodes, with a preference for higher-degree nodes. The networks have $m = 1$ (a) and $m = 2$ (b). Node size is proportional to the degree of the node. Both networks have 100 nodes.

To get a feel for this equation, note that if $a = 0$ and $\alpha = 1$, then the probability of forming an edge to node i is directly proportional to its degree, k_i. Higher values of α produce increasingly preferential attachment. At the most extreme, this leads to a star network where one node receives all the edges. Values of α approaching 0 become increasingly insensitive to degree and more random as a result. Figure 3.4 shows two examples of preferential attachment networks generated with $a = 1$ and $\alpha = 1$.

The a is sometimes called the *attractiveness* parameter. It determines the inherent attractiveness of added nodes when their degree is zero (for directed networks) or small (for undirected networks). Tuning the relative values of α and a determines the balance between preferential attachment and inherent attractiveness. The denominator normalizes the probabilities so that across all nodes they sum to 1.

One intuitive instantiation of the preferential attachment model is called the **bag model** (see Figure 3.5). It has $\alpha = 1$ and $a = 1$. The bag model is equivalent to sampling existing nodes from a bag that contains $k_i + 1$ samples for each node, where k_i is the degree of node i. A new node forms an edge to the node drawn from the bag. If multiple edges are formed, multiple nodes are drawn from the bag.

The preferential attachment model has several interesting properties: (1) Early nodes tend to become hubs of increasingly high degree. First movers into the network have a strong advantage over late comers; (2) The degree distribution across nodes is highly skewed. Many nodes will have low degree, whereas a few will have high degree.

This latter property is sometimes referred to as a scale-free distribution (see Chapter 4). Though preferential attachment is sometimes hailed as the quintessential process for generating scale-free networks, there are many generative processes that give rise to scale-free network distributions. For example, forest fire models, preferential acquisition, duplication and divergence, and fitness models can also generate scale-free distributions (all described in the next section).

Another property of preferential attachment models is that if the number of edges that form for each new node is 1, then the graph will not contain any **cycles**. That is, it will not contain paths that are able to end where they originate without crossing over the

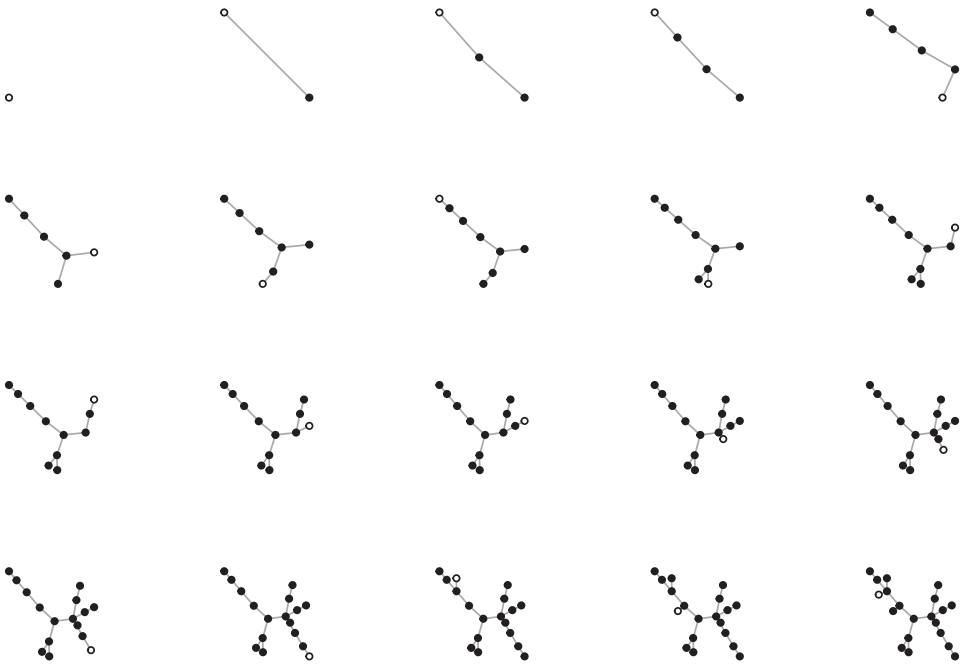

Figure 3.5 The evolution of a preferential attachment "bag" model with $N = 20$. Each frame indicates the addition of the white node to the pre-existing network of black nodes.

same edge twice. Such **acyclic graphs** are peculiar and unlike many social or behavioral systems. However, when they do arise, they may signal certain kinds of behavior such as clandestine networks – where individuals are attempting to minimize contact but maintain connection with the giant component – or minimum spanning trees that may reduce, for example, the cost of maintaining transport networks.

3.4 Fitness Models

In preferential attachment models, early nodes have an advantage. Their degree gains ground early on and that inequality only increases over time. But in some real-world systems, nodes have an advantage because of other qualities unrelated to their degree. For example, they might have an advantage based on attractiveness, and this can happen whenever they enter the network. This dynamic is captured by fitness models.

Fitness models form new edges to existing nodes in proportion to their fitness. The fitness values, η, are randomly assigned to nodes as they are created. This can be combined with Erdös–Rényi random graphs or preferential attachment models.

For Erdös–Rényi random fitness models, edges are formed between nodes chosen at random in proportion to their fitness values. The probability of choosing a node for an edge is:

$$P(i) = \frac{\eta_i}{\sum_{j \in N}^{N} \eta_j}$$

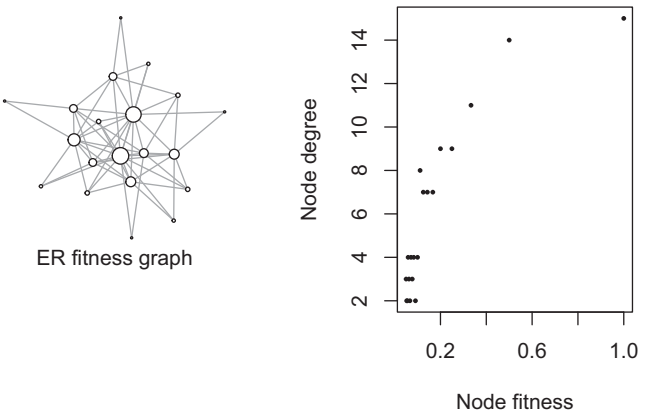

Figure 3.6 Erdös–Rényi fitness-based models. The ER fitness model is based on the scale-free model presented in Goh et al. (2001) with $\alpha = 1$, $N = 20$, and $E = 60$. Node size is proportional to degree.

Two nodes are chosen based on the above probabilities, then an edge is formed between the two nodes. This is repeated for E edges. Figure 3.6 provides an example of an ER fitness network.

Like the preferential attachment network described earlier these ER random graphs can be made to have long-tailed degree distributions by assigning fitness to the nodes in an appropriate way. In Chapter 12, a variation of this approach is used to simulate the structure of conceptual relationships we experience in daily life.

Fitness models can be combined with preferential attachment (e.g., Bianconi & Barabási, 2001), such that the probability of an edge is

$$P(i) = \frac{\eta_i k_i}{\sum_{j \in N}^{N} \eta_j k_j}.$$

The evolution of a fitness-based preferential attachment model is shown in Figure 3.7. Unlike the bag model, nodes in this network can arrive late and still develop high degree.

Fitness networks could have a variety of other implementations, limited largely by one's creativity. For example, attractiveness, a, in the earlier preferential attachment implementation could be a form of fitness and applied independently for each node. Fitness could also be combined with other structural properties besides degree. For example, if fitness were a result of clustering coefficient this could represent a form of group popularity – community fitness.

3.5 Configuration Models

Sometimes we are interested in how the network we observe might be different if only some particular aspect of the world were different. For example, suppose we all had the same number of friends we currently have, but they were replaced with random friends. Hubs are still hubs and hermits are still hermits, but something has changed. What is it?

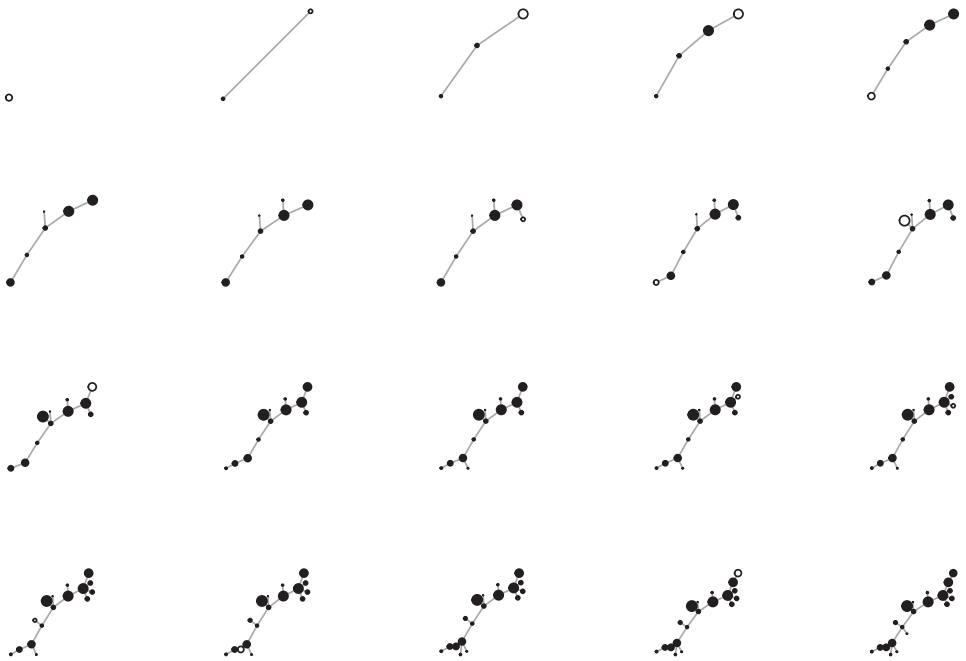

Figure 3.7 The evolution of a fitness-based preferential attachment (Bianconi–Barabási) model with $N = 20$. Fitness is assigned to each node when it originates from a uniform distribution between 0 and 1. The size of the node is proportional to its fitness. Notice that the node with an early headstart in degree development is lower fitness than the nodes that later connect to it. This leads still later nodes to attach preferentially to these higher fitness nodes, not to the early hub.

Configuration models help us understand this by generating networks that preserve features of our original network while randomizing other properties. The random networks will have, for example, the same degree distribution as the observed networks but lack additional structural properties that are not a result of the degree distribution alone (e.g., Milo et al., 2002). Configuration models can be used to evaluate the role of this additional structure because they effectively randomize this structure while preserving degree. It is therefore a useful tool for establishing that something more than degree is responsible for the structure we see.

Configuration models begin with a known network and then cut all edges, turning them into *stubs*. Pairs of stubs are then chosen at random and reattached. This is repeated until all stubs are re-attached (Van Der Hofstad, 2009). See Figure 3.8 for an example.

There are many variations on the configuration model theme. These variations constrain networks to maintain certain network motifs (or subgraph structures), and then generate distributions of networks that preserve these statistical properties. This allows researchers to statistically measure dependencies in the network while controlling for other network features, such as reciprocity, transitivity, and assortativity. For example, to see if reciprocity is greater than we might expect by chance, we can compare the observed network with different randomized variations of itself. This is closely associated with **exponential random graph models**, which are a useful approach for statistically

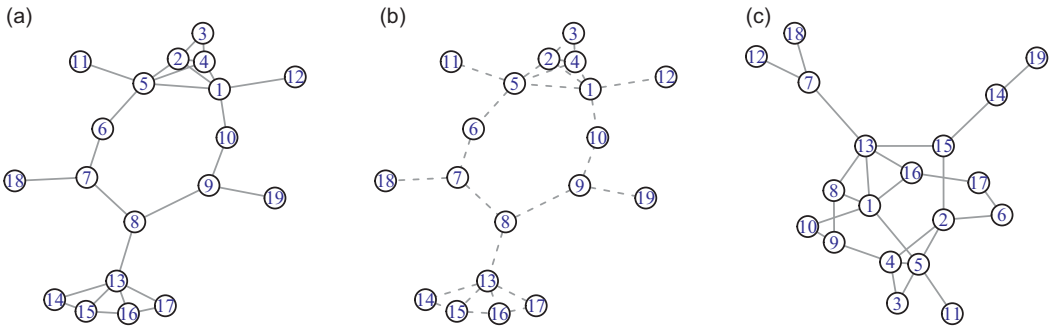

Figure 3.8 Configuration model. Starting with the network in panel (a), all edges are cut and turned into stubs (shown in the middle network). Pairs of stubs are then chosen at random and connected until all pairs are reconnected, producing the network in panel (c). This network has the same number of nodes, density, and degree distribution as the original network.

evaluating network properties relative to controlled variations of your observed network structure (Robins et al., 2007; Snijders, 2001).

3.6 Community Formation

At any local community gathering one can watch as new relationships are formed. These new relationships are rarely made at random. Similar people tend to form edges with one another. Whether by gender, age, or ethnicity, we seem to be drawn to those most like us. How might we capture this process in a network?

Girvan and Newman (2002) – who gave us the Girvan–Newman method for community detection – also present a model for community formation that achieves the desired comingling. Their model for **community formation** starts by assigning each of N nodes to different community memberships. The network is then constructed by forming edges between nodes conditional on whether they belong to the same community: edges are formed with probability p_d if they are from different communities and p_s if they are from the same community, with $p_d < p_s$.

When $p_d = 0$ and $p_s = 1$, communities are completely isolated into cliques of community members. When $p_d = p_s$, there is no differentiation between connections within and between communities. In this latter case, we have an Erdös–Rényi random graph with edges created with probability $p = P_d = P_s$. Figure 3.9 provides an example.

3.7 Forest Fire Models

When we go "down the wormhole" what we are doing in network terms is following a path of relations further and further away from where we started. The network structure left over in our mind tends to look like a forest fire. We started off at one node, and then this led us to activate another, which ignited all the nodes nearby, and so on. This

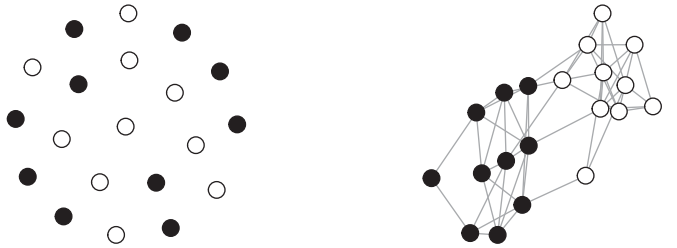

Figure 3.9 Girvan–Newman community formation is achieved in two steps. First, the algorithm assigns nodes to separate communities. Second, edges are formed between nodes, with the probability of an edge dependent on whether the nodes are (p_s) or are not (p_d) in the same community. Here we have $p_s = 0.5$ and $p_d = 0.2$ with $N = 20$ nodes and 2 communities (black and white).

forest fire process mimics many human behaviors. For example, a new relationship leads to introductions to others, and then still others. When we follow citations to scientific research, we create little wildfires of intellectual exploration.[1]

Forest fire models capture this expanding connectivity by allowing a new edge formed to one node to increase the likelihood that edges form to its neighbors as well. Nodes form edges not only to randomly selected nodes, but also with some probability to their neighbors, and then their neighbor's neighbors, and so on.

Numerous models based on forest fire dynamics have been introduced (Vázquez, 2001). These models make different assumptions but roughly follow the variation described here:

Given a pre-existing network of one or more nodes, new nodes are added sequentially. As each node is added, edges are formed to the pre-existing network nodes as follows:

- Step 1. *m ambassador* nodes are chosen at random from the pre-existing network and edges are formed to each of them.
- Step 2. For each ambassador, edges are formed to its neighbors with probability p. Newly attached neighbors become new ambassadors. Step 2 repeats until there are no new ambassadors.

This process allows new nodes to form connections with nodes that are nearby their "entry point" into the network. Occasionally, as Leskovec et al. (2007) describe, this process produces large "conflagrations," in which many new edges are formed by a newly entering node.

In the model shown in Figure 3.10, an approach similar to Vázquez (2001) is used, in which new edges are formed to each ambassador's neighbors, each with probability p.

[1] Many networks might grow like this. Leskovec et al. (2007) found that many real-world networks show an increase in density when the number of nodes added to the network increases. This is a marker of a forest fire process.

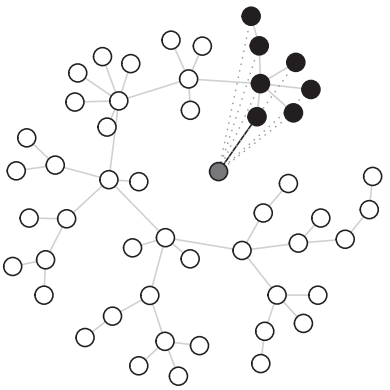

Figure 3.10 A forest fire model based on Vázquez's (2001) recursive search model in which a node attaches to a random node in a network and then follows each edge with probability p, forming a new edge to the adjoining neighbors or otherwise stopping. When all paths are exhausted, the network adds the next node which attaches to a new random target node. Here the new node (gray) is added to a pre-existing preferential attachment network with 50 nodes. The probability of moving down an edge to the next node is $p = 0.94$. The additional new edges that are reached via neighboring nodes are shown as dotted lines.

3.8 Duplication and Divergence

Van Gogh became the artist we now know by making himself the disciple of many of the artists who came before him, copying their works even as he developed his own style. Many novel things arise by a similar process. New genes often start out as copies of existing genes, generated by gene duplication. A protein produced by the duplicated gene maintains protein interactions (edges) held by the gene it duplicated. Over time, further mutations lead to divergence, removing some of the edges and forming new ones as well. New species arise by the same process – the descendants of two siblings slowly diverging over time until one hairy ground-dweller stares up into the feathered face of a distant cousin.

Duplication and divergence models approximate this evolutionary process in network formation (Slanina & Kotrla, 1999; Steyvers & Tenenbaum, 2005; Vázquez, 2003). Starting with a pre-existing network of one or more nodes, the growth process follows these rules:

- Step 1. Identify a node to duplicate. This can be uniformly chosen among existing nodes or chosen proportional to some feature (e.g., degree).
- Step 2. Duplicate the node, giving it the same edges as its parent. This new node is initially **structurally equivalent** to its predecessor as it has the same neighbors.
- Step 3. Modify the edges of the new node to reflect its divergence over time from the original.

An example of this process is shown in Figure 3.11.

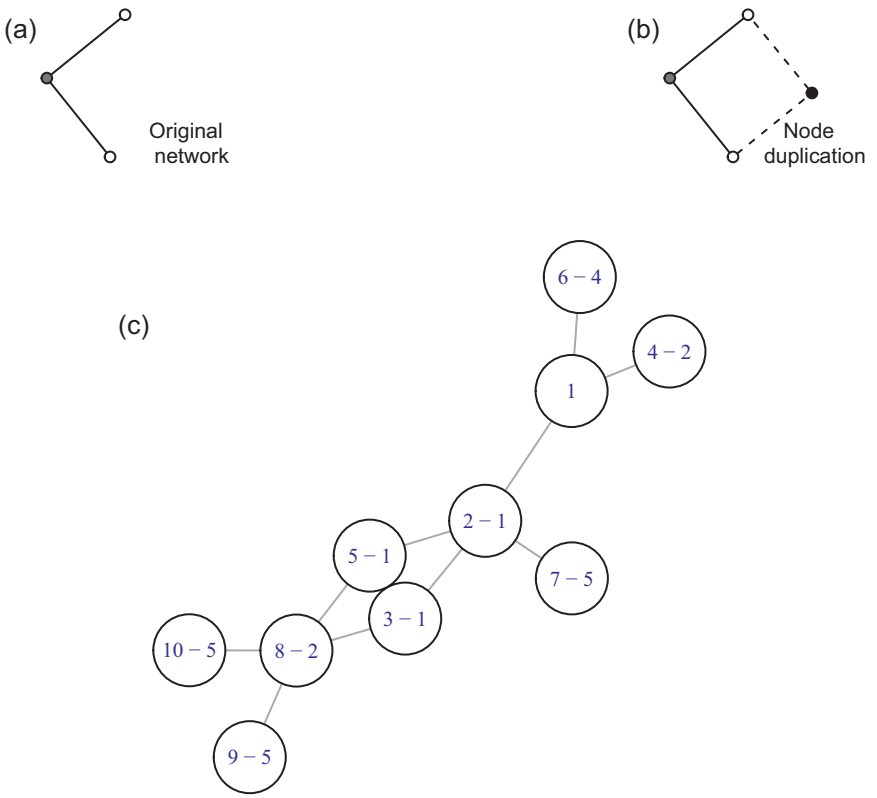

Figure 3.11 Duplication and divergence model. The two networks on the top show the duplication of the gray node (a), which becomes the black node (b). The two dotted edges indicate the duplication of the gray node's two edges. The network (c) shows a duplication and divergence model with $N = 10$ nodes and $p = 0.8$. Each node after node 1 is labeled with "X–Y" where the first number is the order of the node addition – from 1 to 10 – and the second number is the node that was duplicated. For example, the node labeled "10–5" is the 10th node and it is duplicating node 5.

3.9 Preferential Acquisition

The models given earlier all assume the network is generated from thin air. There is no structural information in the environment to inform the network's evolution. This is unlikely for most networks. Social networks exist prior to the formation of online social networks. Spatial networks exist prior to the formation of internet hubs and airport connections. The structure of the world and other forms of information exist prior to our assigning language to it.

Thus, while many of the above networks might be useful for theorizing about the development of networks when other information is absent, we can also consider how networks might develop in response to existing structure.

One network model that does this is called **preferential acquisition** (Hills et al., 2009). It adds nodes to the new developing network with a preference for nodes with certain structural properties in the pre-existing structure.

For example, we might ask which individuals are most likely to join a new social media site. If individuals with high degree in their existing social networks are the first to sign up for the new social media site, the network's development will show a signature pattern of growth that would be different from a network joined first by individuals with lower degree.

This can be formalized by allowing nodes to be acquired in the new network with probability proportional to some developmentally appropriate structural property, d:

$$P(i) = \frac{d_i^\beta}{\sum_j^N d_j^\beta}$$

Here d represents the preferential feature and β indicates the sensitivity to that feature. When $\beta = 0$ nodes are added at random. For very large β, the network adds the nodes in order of their preferred characteristic. For our social network example, d might be the degree of the individual in their existing social networks prior to joining the social media site.

Figure 3.12 shows preferential acquisition at work. Nodes are selected from a larger pre-existing network (shown in panel (d)) and added to a growing network. The development is shown in three phases, with a preference for attaching higher degree nodes first. As discussed in Chapter 5, this model can be applied to a number of developmental networks to ask which properties best predict network growth.

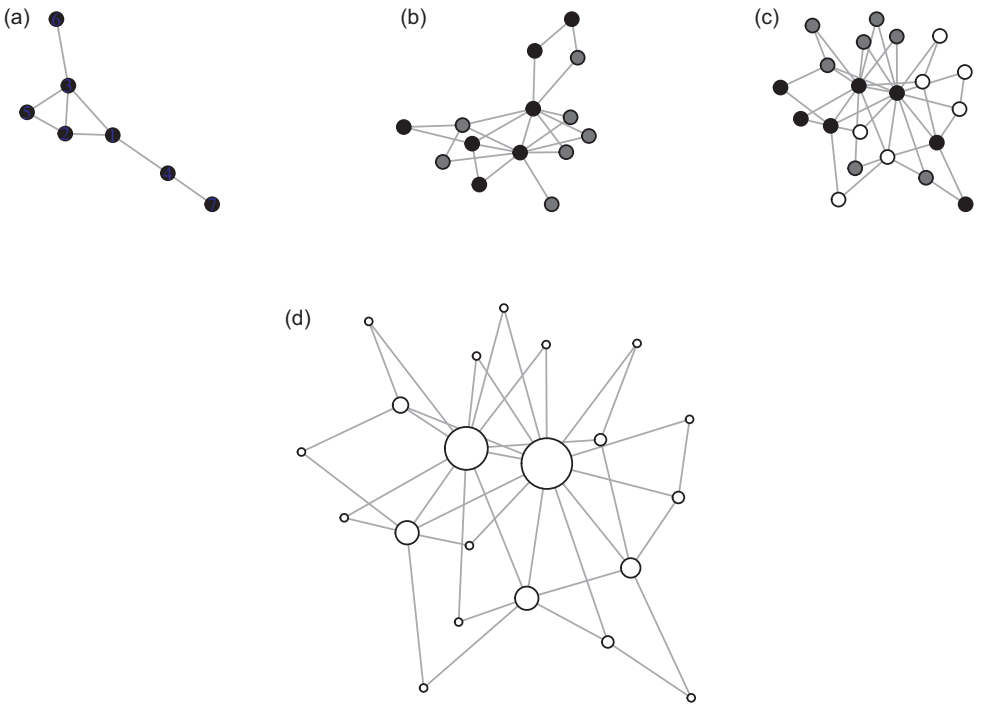

Figure 3.12 Preferential acquisition. The network develops from (a) to (c) along the top. The first nodes to develop in the new network are shown in black. These are followed by the addition of the gray nodes. The final additions are the white nodes shown in panel (c). The network in panel (d) shows the network from which the above networks are preferentially acquired, with node size indicating node degree. Nodes with higher degree are preferentially acquired earlier in development.

3.10 Directed and Weighted Graphs

Each of these models can be generated as directed or weighted networks. The basic logic is a follows: (1) sample from possible directed edges or replace undirected edges with directed edges, and (2) sample weighted edges from predefined distributions. Some possible variations of these approaches are described here, but these are not exhaustive:

- *ER random graph*: Directed edges are randomly assigned between pairs of nodes with the specified probability or edge count. Weighted edges can be drawn from a prespecified distribution, for example, from the distribution of an observed network.
- *Small-world networks*: Directed and/or weighted edges can be specified at the lattice phase.
- *Preferential attachment*: In the original Barabási–Albert preferential attachment model, edges are formed directed toward the preexisting node and indegree is used to evaluate preferential edge formation.
- *Fitness models*: Nodes can have indegree and outdegree fitness. To form an edge, one node is sampled proportional to its indegree fitness across all nodes, and one node is sampled proportional to its outdegree fitness. Weighted edges can be sampled from a predefined distribution or as a function of a fitness value.
- *Configuration models*: As these models recreate an existing structure, they preserve the (un)directed and (un)weighted structure of the original network. Weighted configuration models require additional assumptions to coordinate weights across stubs.
- *Community formation*: As edges are probabilistically evaluated for all pairs of nodes, the probabilities can be defined to produce unreciprocated or reciprocated directed edges. Weighted edges can be sampled from different distributions depending on whether edges are in the same or different communities.
- *Forest fire models*: Edge formation can be directed toward nodes in the preexisting network, emulating a citation process. Weighted edges can be sampled from a predefined distribution or sampled in proportion to some structural property during edge formation, such as the number of new neighbors that an individual node provides access to.
- *Duplication and divergence*: Directed and/or weighted edges can be assigned at random and replicated in the duplication process. Divergence could alter directions or weights.
- *Preferential acquisition*: As this process provides various development pathways for arriving at a pre-existing structure, the directed or weighted values of edges is predefined in the existing structure. These values can then be used to derive the node sampling procedure.

Metrics computed on directed and weighted networks can sometimes be difficult to interpret with respect to a specific problem. For example, there are multiple methods for computing weighted clustering coefficients (e.g., Barrat et al., 2004; Onnela et al., 2005; Opsahl & Panzarasa, 2009). Determining which is most appropriate for a given system and interpreting what it means in relation to the behavioral processes can be a valuable but challenging exercise.

3.11 A Taxonomy of Basic Network Architectures

Basic network architectures provide useful intuition pumps. If we imagine how a process might work on a simplified network, it can help us to generalize to more complicated networks with many more nodes and edges. Figure 3.13 provides a few example networks for comparison. Which of these networks is best insulated from the spread of infection or opinion? Which of these networks is best for spreading good news or most resilient to losing nodes or edges? Which nodes should be vaccinated first against disease or misinformation? Which networks will have the most behavioral diversity or innovation? Which networks would be most coordinated? We will tackle these and many other questions in the chapters to come.

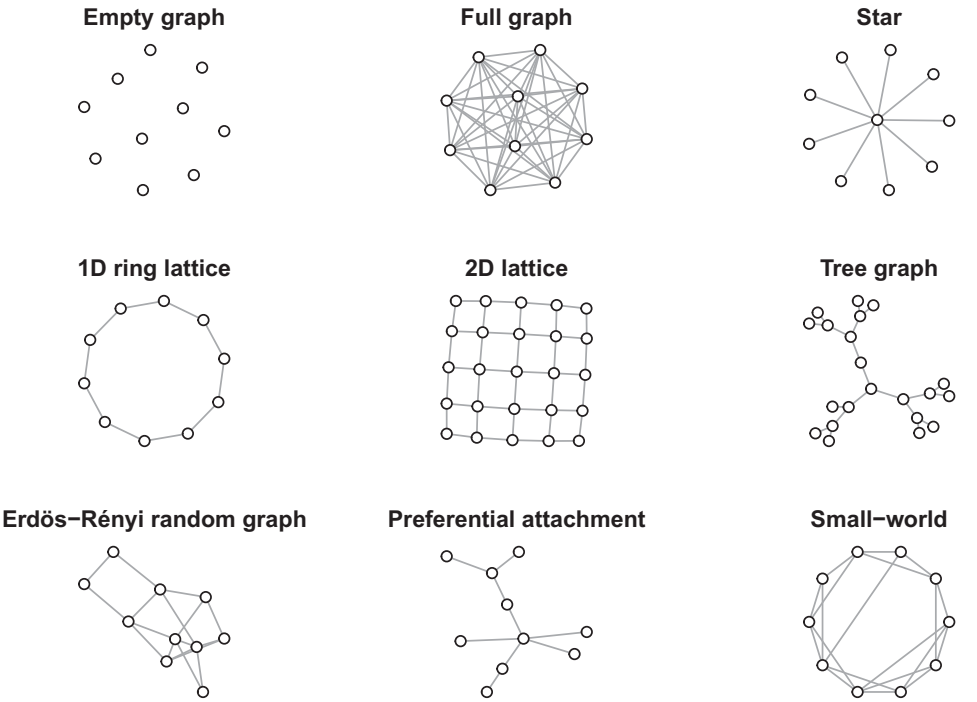

Figure 3.13 A menagerie of some basic network types.