Network Theory: 80/20 Rule and Small Worlds Theory

Introduction

Starting with isolated research in the early twentieth century, and following with significant gaps in research progress, network theory has recently formed into a broad theoretical discipline. As an emerging field of research, network theory studies information networks (the World Wide Web), technical networks (the Internet, railways, airline routes), biological networks (the human genome), and social networks (human relationships) (Newman, 2003). Networks are systems of nodes and links and network theory is the study of the interconnections found in networks. As such, network theory is an empirical discipline; it studies ‘real-world’ networks in natural settings.

Network theory fits within a broader theoretical discipline known as complexity theory (the study of complex systems). Like network theory, complexity theory is interdisciplinary and studies complexity on multiple levels. Examples of complex systems include weather patterns, food webs, and traffic flow. Network Theory is a subset of complexity theory that specifically studies complex networks. These theories overlap and influence each other and their boundaries are fuzzy.

Small worlds theory is the dominant network theory and social network analysis is the dominant methodology. Small worlds theory has its foundations in the mathematical discipline of graph theory, and is the study of emergent ‘small worlds’ phenomena within networks; social network theory focuses specifically on the unique characteristics of social networks.

Network theory is relevant to the communication and use of information because such communications implies networks: computer networks, networks of libraries and library resources, the social networks of information scientists, librarians, and patrons. Network theory helps us understand the structure and behavior of networks, creates models to help us understand the meaning of network properties, and helps us predict future behavior of networks.

One of the fundamental conclusions of complexity and network theory contradicts a popular scientific misconception that we live in an entirely random universe. Network theory suggests there is a degree of order in the universe and there are common patterns.

All networks have common properties and are made up of nodes (also know as vertices) and links (also known as connectors, branches, edges). Nodes are points in a network where a message can be created, received, or repeated. Hubs are a specific type of node that has many links flowing out from it; authorities are another type of node that has many links flowing into it. Links transmit messages and connect nodes. Local links connect relatively close nodes; long distance links connect nodes that are far away from each other.

Networks can be random or scale-free; random networks are hypothetical (they tend to be made up by mathematicians), they are evenly distributed (like a bell-curve, ex. most WebPages would have the same number of links), and they are static (no growth or preferential attachment between nodes). They also tend to have fewer nodes and links than ‘real-world’ networks.

Scale-free (or aristocratic) networks are ‘real-world’ networks; they tend to be clustered, and they are dynamic (display growth and preferential attachment). Because of the dynamic nature of ‘real-world’ networks, early nodes have more time than latecomers to acquire links, hence the preferential attachment. Also, scale-free networks tend to have millions of nodes and links (Barabási, 2002, p. 87).

Networks can be centralized, in which all links emerge from a central core, such as in mainframe computer architecture, with all processing going on at the core; decentralized, in which links emerge from clusters of nodes such as in a client/server architecture with processing shared among all nodes; or distributed, in which all nodes are connected to all others, such as the physical network of the Internet.

Originally designed as a military communications network designed to withstand a nuclear attack, the Internet’s distributed architecture has developed into the world’s largest example of a ‘real-world’ network.

Empirical observation tells us that networks are ubiquitous. As an ‘interconnected’ universe, we inhabit multiple networks simultaneously; social networks, information networks, technical, and biological networks. It is network theory that best offers the opportunity of understanding such networks.

80/20 Rule

The 80/20 rule is a “rule of thumb” concerning economic inequalities found in many phenomena. It has been applied to: business management, citation analysis, criminology, and Web analysis. For example, business managers have described the 80/20 rule as the ‘Murphy’s law of management.’

Examples of the Murphy's law of management include: 80% of profits are the result of 20% of employees, 80% of problems are caused by 20% of customers, 80% of decisions are made in 20% of meeting time, 80% of efforts are wasted and 20% are productive (Barabási, 2002, p. 66).

Citation analysts have found that 80% of citations cite 20% of scientists; criminologists have found that 80% of crime is committed by 20% of criminals; Hollywood film actors have found that 80% of films are made by 20% of actors; and Web analysts have found that 80% of Web links are linked to 20% of WebPages. Even libraries have applied the 80/20 rule, with 80% of resources rarely used and 20% used often.

The 80/20 rule was originally known as Pareto's law. Vilfredo Pareto was an influential Italian economist at the turn of the century who attempted to turn economics into an exact science describable by laws. Through empirical observation in his garden, Pareto discovered that 80% of peas were produced by 20% of peapods. Combining economic formulas with his empirical observations, he then went on to discover that 80% of Italy's land was owned by 20% of population. Pareto's application of the 80/20 rule grew from there.

So what does the 80/20 rule have to do with network theory? The answer is that it displays a property that plays a key role in understanding complex networks: the power law. A power law distribution is a characteristic of the 80/20 rule and scale-free networks. It is represented by a continuous decreasing curve and is indicative of the distribution of 'real-world' networks such as the Internet, the neural networks of the brain, and the human genome. Let's take the Web for example; a bell curve distribution would suggest most WebPages would be equally popular. Instead we find that relatively few WebPages are popular and most are not. This indicates a power law distribution in which many small events coexist with a few large events (The decay of distribution in a power law is known as the degree exponent).

The power law is a mathematical expression that indicates that the interconnected universe (as network theorists like to call it), and complex networks within the universe, are not entirely random. In other words, there is a degree of order in the universe. Networks feature common patterns and rules of behavior. Network theory attempts to divine these patterns and rules of behavior.

Small Worlds Theory

Small worlds theory is a popular network theory that was preceded by a branch of mathematics called graph theory: the mathematical study of how groups of things can be connected together (Buchanan, 2002, p.34). The Swiss born mathematician Leonhard Euler (1707-1783), pioneered graph theory as a branch of mathematics; his focus was on studying the properties of various graphs. Two centuries later, Paul Erdõs (1913-1996), a Hungarian mathematician, revolutionized graph theory by exploring how graphs, now referred to as networks, are formed. He wrote over 1500 mathematical papers in his lifetime with a seminal paper in 1959 that introduced random network theory. He was famous for saying, “a mathematician is a machine that turns coffee into theorems.” Erdõs discovered that no matter how many nodes there may be in a network, a small percentage of randomly placed links is always enough to tie the network together into a more or less completely connected whole. More surprisingly, the percentage required dwindles as the network gets larger. For example, for a network of 300 nodes, there are almost 50,000 possible ways to link them. But if no more than 2% of those links are in place, the network will be completely connected.

Erdõs, along with fellow mathematician Alfréd Rényi, developed random graph theory in an attempt to understand the rules of interconnected networks. No matter what type of network, Erdõs and Rényi noticed that they could be simplified into nodes and links; these nodes and links form a graph. The simplest mathematical way of describing the interconnections of these graphs was to connect the nodes randomly. In other words, each node on the graph features an equally distributed number of links; this forms a random network. Thus, Erdõs and Rényi, viewed graphs and the networked world they represent as fundamentally random; it was left to small worlds theory to discover the misconception of this conclusion. Network theorists found that ‘real-world’ networks were not random at all, but followed predictable patterns of order and growth.

While graph theory exclusively studies random networks, small worlds theory studies the scale-free networks found in the ‘real-world’, thus, it is the study of the interconnections that form ‘real-world’ networks.

Perhaps the single most influential piece of research in small worlds theory is Mark Granovetter’s “The Strength of Weak Ties” (1973). Granovetter found that the links found in social networks could be divided into strong ties (between family members, friends, coworkers, colleagues) and weak ties (between casual or rare acquaintances). Strong ties tend to form clusters and have little effect on the overall connectivity of the network. It is the weak ties that are most important to the formation of real world networks. Granovetter called these weak links ‘bridges’; they act as crucial ties that bind the social network together. As a real world example, Granovetter performed a research study that found that only 16 percent of people he interviewed got their job through a strong contact, 84 percent were through contacts they saw occasionally or rarely.

Small worlds theory has expanded our understanding of ‘real-world’ networks of all kinds by discovering fundamental properties at work in such networks. Yet, there is much yet to be discovered; perhaps the most groundbreaking research is being conducted specifically on social networks.

Research progress

80/20 Rule

As an example of early research on network theory, Pareto’s Law, better known as the 80/20 rule, is of significant interest. Pareto and his followers thoroughly explored the 80/20 rule at the turn of the century and it has been integrated into current network theory research. Network theorists have recently moved beyond the 80/20 rule as a topic of significant research and are now exploring broader concerns.

Small Worlds

It was almost 30 years later that Cornell mathematicians Duncan Watts and Steve Strogatz attempted to bridge the research of Granovetter and Milgram in “The Collective Dynamics of ‘Small-World’ Networks” (1998). They found that ‘real world’ networks have a degree of order and randomness and all networks share common patterns that result in degrees of separation. Starting with a completely ordered network, Watts and Strogatz then added 2% random links and found that in an almost completely ordered network, the random links contributed to six degrees of separation; even a tiny fraction of random links has an immense influence on the degrees of separation in a network. At the same time, the random links have little noticeable impact on the degree of local clustering found in networks.

Watts’ and Strogatz’ research has had a strong impact on small worlds research and continues to be influential as an apparent explanation for the small worlds phenomenon in ‘real world’ networks. Recently Watts and Strogatz have applied their theories to the US power grid with similar results.

Current research in small worlds theory includes the work of physicists such as Albert Barabási (2002), and M.E. Newman (2003) who are exploring small worlds theory as a way to explain the apparent order and connectivity of the information networks of the World Wide Web and the technical networks of the Internet. Biologists are also exploring small worlds theory in an attempt to understand the information networks within the human genome.

Conclusions

Network theory helps us understand: the structure and behavior of networks, the meaning of network properties, and helps us predict the future behavior of networks. As an emerging discipline, there are significant research opportunities (Newman, 2003). While network structure and network properties are currently being explored, there are future opportunities for research in the prediction of network behavior.

The ubiquity of networks is an empirically observable phenomenon. The networks that form the world we live in is a fundamental property of the ‘interconnected’ universe. By attempting to understand networks, we understand ourselves. Thus, the conclusions of network theorists can have an immediate and profound impact on the fundamental awareness of our selves and the universe.

Network theory is relevant because the communication and use of information implies networks: computer networks, networks of libraries and library resources, the social networks of information scientists, librarians, and patrons. And technologically mediated communications depends on communications networks such as the Internet, satellite, radio, and television. The more we understand about networks, the more we can understand communications.

It is important to emphasize the interdisciplinary nature of network theory. Mathematicians, physicists, and sociologists are just a few of the professionals that are currently exploring network theory. Notably, Watts (2003) observes:

> The small-world problem provides the perfect example of how the different disciplines can help each other build the new science of networks. Back in the 1950s, Kochen (a mathematician) and Pool (a political scientist) were the first to think about it but couldn’t find a solution without computers. Milgram (a psychologist), aided by White (a physicist-sociologist) and followed by Bernard (an anthropologist) and Killworth (an oceanographer), then attacked the problem

> empirically but couldn't explain how it actually worked. Thirty years later, Steve [Strogatz] and I (mathematicians) turned the problem into one about networks generally but failed to see its algorithmic component, leaving that door for Jon [Kleinberg] (a computer scientist) to open. Jon, in turn left the door open for Mark (a physicist), Peter (a mathematician), and me (now a sociologist of sorts) to walk through. (p. 160).

Interdisciplinary by nature, Information scientists would do well to pursue the many opportunities for research, explanation, and understanding that network theory offers. As a discipline inundated with networks, the future of information science may depend on it.

References

Barabási, A.L. (2002). *Linked: The new science of networks.* Cambridge, MA: Perseus Publishing.

Buchanan, M. (2002). *Nexus: Small worlds and the groundbreaking science of networks.* NY: W.W. Norton & Co.

Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology, 78(6),* 1360-1380.

Milgram, S. (1967). The small world problem. *Physiology Today, 2,* 60-67.

Newman, M.E.J. (2003). *The structure and function of complex networks*. arXiv: cond-mat/0303516v1.

Watts, D.J. (2003). *Six degrees: The science of a connected age.* NY: W.W. Norton & Company, Inc.

Watts, D.J. & Strogatz, S.H. (1998). Collective dynamics of ‘small-world’ networks. *Nature, 393,* 440-442.