

Introduction to the Special Issue on AI and Networks

Marie desJardins, Matthew E. Gaston,
and Dragomir Radev

■ This introduction to AI Magazine's special issue on networks and AI summarizes the seven articles in the special issue by characterizing the nature of the networks that are the focus of each of the articles. A short tutorial on graph theory and network structures is included for those less familiar with the topic.

Most people ... would agree that a fundamental property of complex systems is that they are composed of a large number of components or "agents," interacting in some way such that their collective behavior is not a simple combination of their individual behaviors.

– Mark Newman

The importance of networks permeates the world today. From biology to social systems, from the brain to the Internet, networks play an important and central role in the way the world works. In the last 10 years, due in part to large increases in computational power, large-scale, real-world networks have received much attention from a variety of fields of study.

Within the artificial intelligence community, networks appear in some form in nearly every subdiscipline: knowledge representation, inference, learning, natural language processing, multiagent systems, analogical reasoning, and many others. The goals of this special issue are to provide a sampling of research efforts focused on how networks can be used in AI systems and to facilitate cross-communication among subdisciplines that are studying networks from different perspectives.

The seven papers we include here cover a broad range of network-inspired AI research—in natural language processing, data mining, the semantic web, peer-to-peer networks, multiagent systems, analog networks, and the modern social network

of the "blogosphere." Each article represents a snapshot of the area it describes; for example, the collective classification problem surveyed by Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Gallagher, and Tina Eliassi-Rad is just one of many problems within the emerging research area of link mining. Moreover, networks are influential in many other areas of AI that are not represented here, including Bayesian networks and graphical models, sensor networks, swarm systems and cellular automata, graphical games, trust and reputation systems, and computational organizational design.

Table 1 summarizes the articles in this collection by characterizing the nature of the networks that are the focus of each of the seven papers.

Basic Graph Theory

In reading the articles presented here, some basics of graph and network theory may be useful for the reader who is not familiar with these terms. We start with some basic terminology.

A *graph* G is defined to be a pair (V, E) , where V is a vertex set and E is an edge set (see following). The terms *graph* and *network* are often used interchangeably.

The *finite vertex set* V is a set of descriptors for the vertices in the graph. Each vertex may just have an identifier, or it may have an arbitrarily complex set of attributes. The terms *vertex* and *node* are often used interchangeably. Depending on the

Authors	Topic	Nodes	Edges	Tasks
Radev and Mihalcea	Natural language processing	Words, word senses, sentences, documents	Cooccurrences, collocations, syntactic structure, lexical similarity	Analyze syntax, identify lexical semantics, retrieve and summarize text, extract keywords
Berners-Lee and Kagal	Semantic web	Agents, terms, ontologies	Connections between communities, subtask relationships, ontological relationships	Disseminate knowledge, construct and share ontologies, provide and request services, create new communities
Menczer, Wu, and Akavipat	Peer-to-peer networks	Agents	“Social” connections along which queries flow	Locate relevant knowledge sources, learn which peers can answer queries
Pearce, Tambe, and Maheswaran	Cooperative multiagent systems	Agents	Interactions, joint reward structures	Multiagent plan coordination, meeting scheduling, teamwork (such as RoboCup soccer)
Mattiussi, Marbach, Dürr, and Floreano	Analog networks	Dynamic devices	Signal flows with varying strength	Synthesize and reverse-engineer analog networks (for example, gene regulatory networks and analog electronic circuits)
Finin, Joshi, Kolari, Java, Kale, and Karandikar	Blogosphere	Web pages, blog postings, bloggers, blog sites	Social networks, comments, trackbacks, advertisements, tags, RDF data, metadata	Recognize spam blogs (splogs), find opinions on topics, identify communities of interest, derive trust relationships, detect influential bloggers
Sen, Namata, Bilgic, Getoor, Gallagher, and Eliassi-Rad	Social and natural networks	Entities (such as scientific articles)	Relationships among the entities (for example, citations or cocitations)	Perform collective classification, construct features for relational classification

Table 1. An Overview of the Articles in This Special Issue.

application, nodes may also be referred to as *agents* or *entities*.

The *finite edge set* E specifies the relationships between the vertices in the graph. Each edge $e \in E$ is a pair of vertices, which are called the *end points* of the edge. Edges may be *ordered* or *unordered* and also *weighted* or *unweighted*. A *hyperedge* may connect more than two vertices. Edges are often used to represent *relations*.

The degree of a node, k_i , is the number of edges that are connected to node i . In directed graphs, degree can be broken down into “in-degree” (number of edges coming into the node) and “out-degree” (number of edges pointing out of the node).

Network Properties

A number of properties prove to be useful in graph theory and social network theory for analyzing and

understanding the behavior of graph structures.

The *path length* between two nodes is the minimum number of edges that must be traversed to move from one node to the other in the graph. The *average path length* is an average across all pairs of nodes in the graph.

Real-world graphs often *exhibit short average path lengths*, meaning that the average path length is less than would be expected in a random graph. This “small-world effect” was first recognized by Stanley Milgram (1967) in analyzing the number of hops it took for human subjects to send a piece of postal mail to a predefined destination by following only links to people whom they knew on a first name basis. This phenomenon is sometimes called “six degrees of separation,” based on the hypothesis that any two people in the world can be connected by at most a six-link “chain” of acquaintances. A game created in the mid-1990s

called Six Degrees of Kevin Bacon (find a short path connecting any given movie actor or actress to Kevin Bacon) in fact initiated some of the research work that led to the current boom in interest in network studies.

Several other properties are related to path length: The *betweenness* of a node i is the number of other pairs of nodes (j, k) whose shortest paths pass through i . The *closeness* of a node is the average shortest path to all other nodes in the graph. The *diameter* of a graph is the length of the longest of all shortest paths (that is, it is the maximal distance between any pair of nodes (i, j)).

Clustering measures are used to characterize the frequency of transitive relationships in networks (Newman 2003, Albert and Barabási 2002, Watts and Strogatz 1998). The *clustering coefficient* of a network is the ratio of triangles in a network (sets of three nodes that are all connected to each other) to the number of connected triples (sets of three nodes in which at least one node is connected to the other two).

Real-world networks often exhibit *excess clustering*, in the sense that they have a much higher (often two orders of magnitude or more) clustering coefficient than would be expected in a random graph of the same size (Newman 2003). This is because in many processes that generate networks, two nodes that are connected to a common neighbor are more likely to become connected.

The *degree* of a node is also sometimes called its *degree centrality*, since the number of edges that are connected to a node give an indication of how “central” it is to the network. The *degree distribution* of a network is the frequency of occurrence of nodes with each degree. A useful summary property is the network’s *average degree*, which can be thought of as the *density* of the network. The *normalized standard deviation of the degrees* of the nodes can be used to characterize how much variability there is in the network density. The *degree correlation* of adjacent nodes in a network indicates whether neighboring nodes are likely to have similar degree.

Many real-world networks have highly skewed degree distributions, with high normalized standard deviation. In particular, the degree distribution in real-world networks often follows a *power law*, where the probability of a node in the network having degree k is proportional to $k^{-\gamma}$ for some parameter γ (typically γ is between -2 and -3) (Newman 2003). Such networks have a hub-and-spoke structure, with some nodes having very large degree (Albert and Barabási 2002).

Network Models

A variety of network models have been proposed to represent various types of network formation

processes and graph behaviors. Several of the most common models are described in the following paragraphs.

Regular graphs have a homogeneous connectivity pattern for all of the nodes in the graph. In these graphs, the degree distribution is trivial: all nodes have the same degree. Examples of regular graphs include lattices, hypercubes, and fully connected networks (in which all nodes are connected to all other nodes).

The coordination number (Watts 1999) of a lattice graph determines the number of connections that each node has with its spatial “nearest neighbors” in each dimension. An example of a one-dimensional lattice with $K = 2$ is shown in figure 1(a).

Random graphs were first introduced by Paul Erdős and Alfréd Rényi (1959). A random graph $G_{n,p}$ consists of n nodes where p denotes the probability of an edge existing between each pair of vertices. Random graph models have been widely studied, in part because their properties can be computed analytically. For instance, the expected number of undirected edges in $G_{n,p}$ is $n(n-1)p/2$, and the average degree of a vertex is $k = p(n-1)$.

A *random geometric graph* is a special case of a random graph that is generated by randomly placing N nodes in the unit square, then connecting pairs of nodes if they are within some specified distance of each other (Dall and Christensen 2002). More specifically, two nodes, i and j , are connected in a random geometric graph if $d(i, j) < \phi$, where ϕ is a threshold parameter of the model. Figure 2 shows an instance of a random geometric graph with $\phi = 0.09$.

The *small-world network* model of Duncan Watts and Steven Strogatz (1998) is an attempt to produce networks that exhibit the real-world properties of excess clustering and short average path length. Small-world networks have properties that lie between those of regular (lattice) networks and random graphs. Small-world networks are constructed by randomly “rewiring” each edge in a lattice network with some probability ρ . This process results in shortcut connections across the network, as seen in figure 1. (When edges are replaced with random shortcuts with probability $\rho = 1$, the resulting graph is a random graph.)

The *scale-free graph* model is motivated by the empirically measured degree distributions of the Internet and the World Wide Web (Albert and Barabási 2002; Albert, Jeong, and Barabási 1999). The model is a highly intuitive model based on the way that many networks are believed to evolve and grow in the real world. The generation of scale-free graphs has two simple rules: (1) *growth*—at each time step, a new node is added to the graph, and (2) *preferential attachment*—when a new node is added to the graph, it attaches preferentially to existing nodes with high degree. Figure 3 shows an

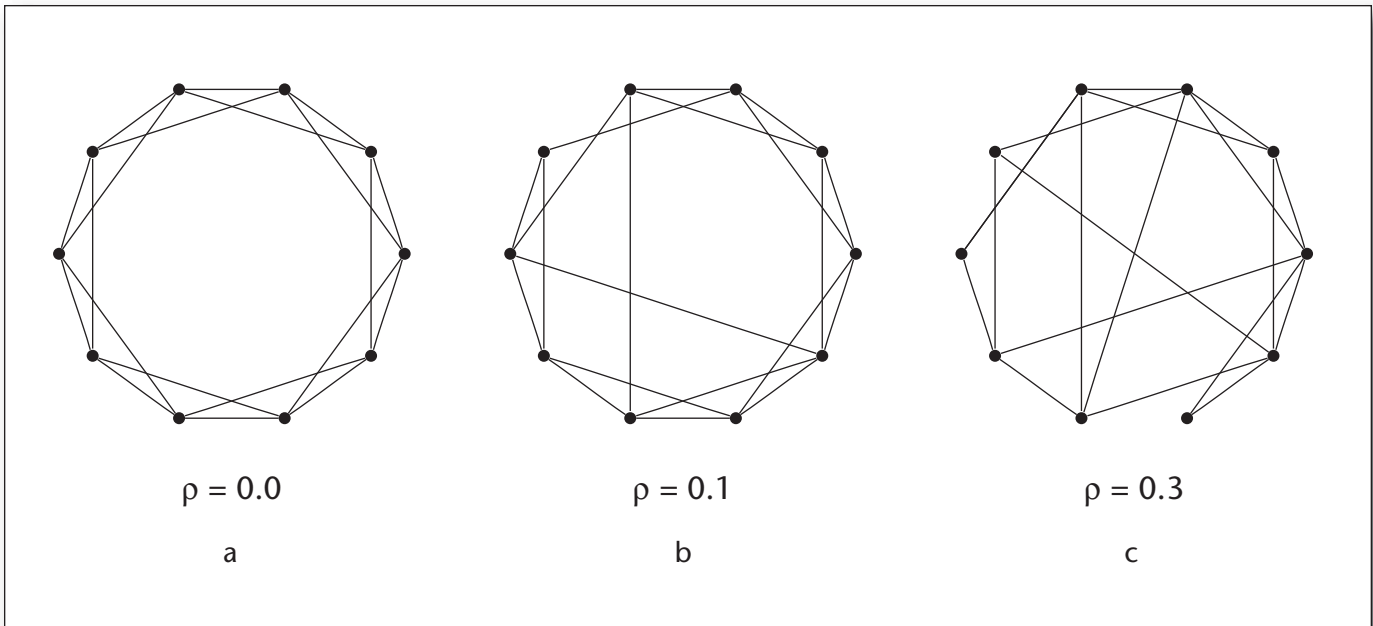


Figure 1. Three Increasingly Random Small-World Networks.

(a) A small world with no shortcut links. (b) The same small world with a few shortcuts. (c) A small world with many shortcuts, which begins to resemble a random graph. All three of the networks are constructed from a one-dimensional lattice where nodes are connected to $K = 2$ other nodes in each direction, based on physical proximity. This particular choice of initial layout is deliberate in that it ensures a high initial clustering coefficient.

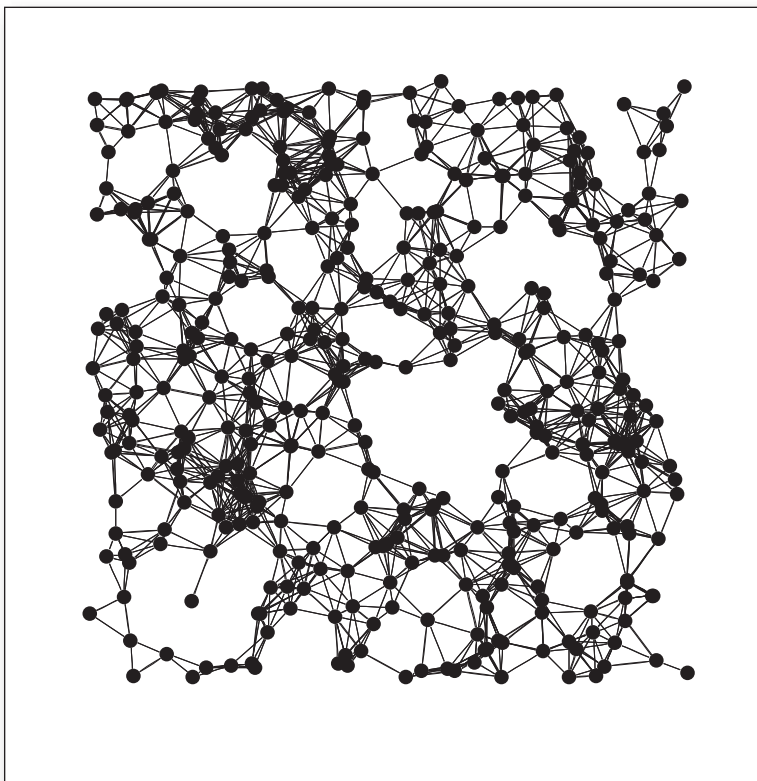


Figure 2. An Instance of a Random Geometric Graph on the Unit Square with 400 Nodes and $\phi = 0.09$.

example of a scale-free network structure and the power-law degree distribution that it exhibits.

Summary

Networks have been studied in artificial intelligence in a variety of contexts since AI's inception over 50 years ago. However, new insights from the study of very large real-world networks and from the connective power of the Internet call for an updated perspective on the importance and role of networks in AI. This special issue provides a sampling of the most innovative recent research in networks and AI.

References

Albert, R., and Barabási, A.-L. 2002. Statistical Mechanics of Complex Networks. *Review of Modern Physics* 99(3) (May): 7314–7316.

Albert, R.; Jeong, H.; and Barabási, A.-L. 1999. The Diameter of the World-Wide Web. *Nature* 401: 130–131, (10.1038/43601).

Dall, J., and Christensen, M. 2002. Random Geometric Graphs. *Physical Review E*, 66(016121): 1–9, July. (10.1103/PhysRevE.66.016121).

Erdős, P., and Rényi, A. 1959. On Random Graphs. *Publicationes Mathematicae* 6(Debrecen): 290–297.

Milgram, S. 1967. The Small World Problem. *Psychology Today* 22 (May): 61–67.

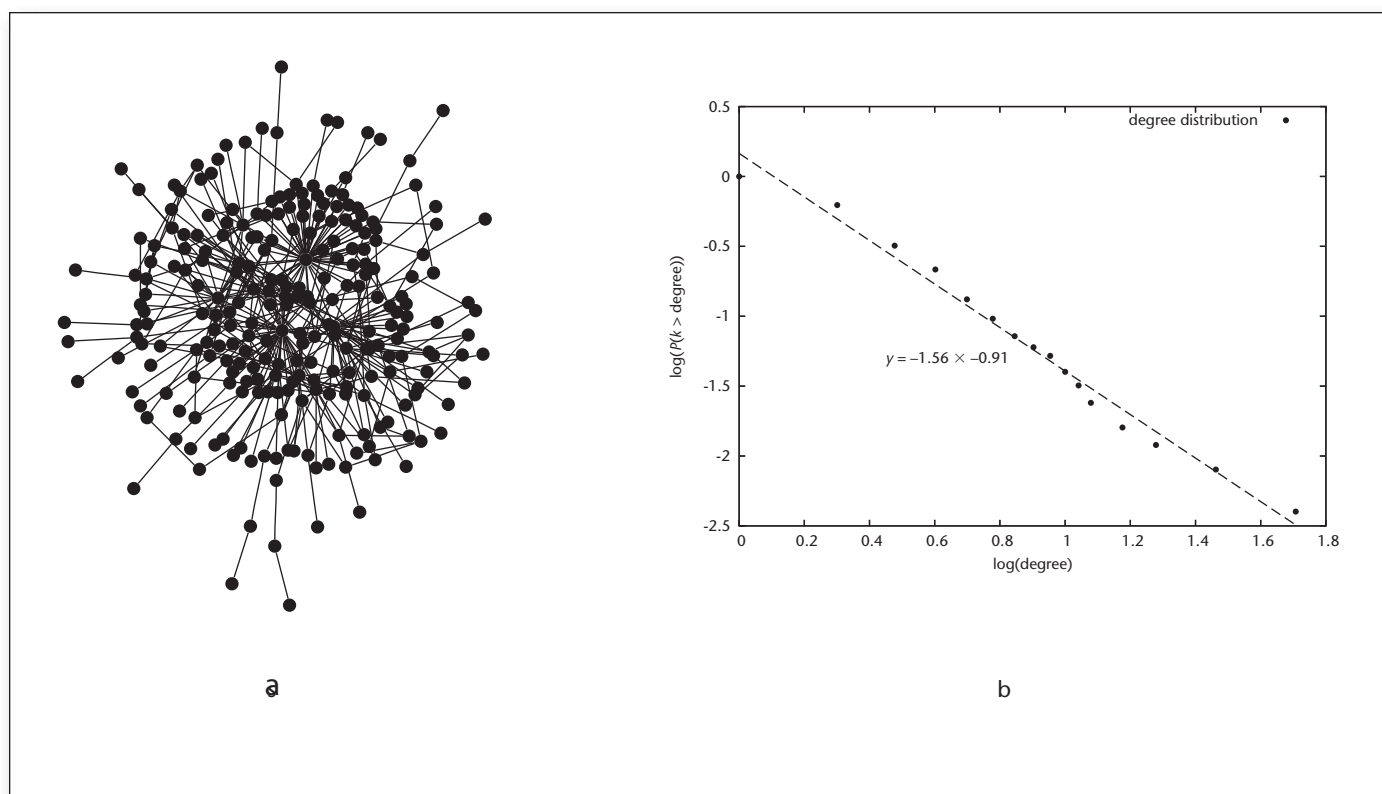


Figure 3. An Example of a Scale-Free Network Structure with 250 Nodes.

(a) A rendering of the network that clearly shows the hub-and-spoke structure. (b) A log-log plot of the cumulative degree distribution of the network shown in (a). Note that a linear curve in a log-log plot implies a power-law behavior of the underlying system.

Newman, M. 2003. The Structure and Function of Complex Networks. *SIAM Review* 45(2):167–256.

Watts, D. 1999. *Small Worlds*. Princeton, NJ: Princeton University Press.

Watts, D., and Strogatz, S. 1998. Collective Dynamics of “Small-World” Networks. *Nature* 393(6684): 440–442.



Marie desJardins (mariedj@cs.umbc.edu) is an associate professor in the Department of Computer Science and Electrical Engineering at the University of Maryland Baltimore County. Her research focuses on machine learning, planning, and multiagent systems, with an emphasis on interactive and integrated approaches to AI.



Dragomir R. Radev is an associate professor at the School of Information and the Department of Electrical Engineering and Computer Science at the University of Michigan. Radev received a Ph.D. in computer science from Columbia University in 1999. He is currently secretary of the Associa-

tion for Computational Linguistics and program chair of the North American Computational Linguistics Olympiad and also coach of the US national linguistics teams. His areas of research include natural language processing, information retrieval, and network theory.



Matthew E. Gaston is a senior research scientist at Viz, a division of General Dynamics C4 Systems, where he leads research and development activities focused on creating intelligent user interfaces for information visualization and collaboration. He received his Ph.D. in computer science from the University of Maryland Baltimore County, where his research focused on networks and multiagent learning. He can be contacted at mgasto1@cs.umbc.edu.