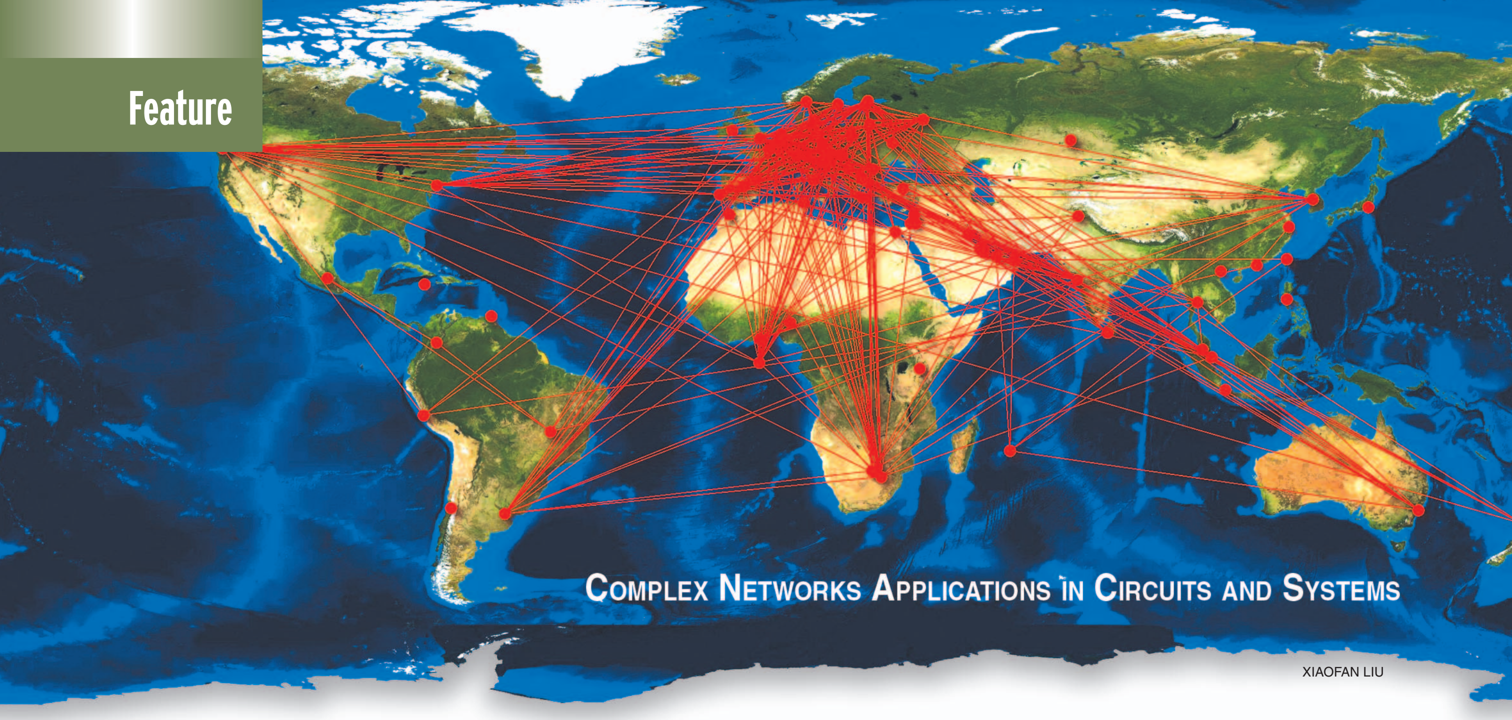
Feature



LiYing Cui, Soundar Kumara,
and Réka Albert

Complex Networks: An Engineering View

Abstract

The last decade has seen an explosion of research in network science, a science whose initial work and subsequent developments are grounded in statistical physics applied to natural systems. In recent years researchers in the engineering discipline have also taken a keen interest in complex networks which has resulted in several application areas being investigated in engineering. In this paper, we focus on network optimization and network dynamics in the context of engineered (human made) networks.

Digital Object Identifier 10.1109/MCAS.2010.937883

I. Introduction

The number of publications in the area of large scale networks in the last decade is almost equal to the total number appeared before 1999 (17,300 publications). This is an indication of the explosion of network science and its establishment as a strong interdisciplinary research field in recent years. It is increasingly recognized that network science is highly relevant to engineering as the sheer size of engineered systems poses unique challenges in their design and analysis. Information Technology provides for rich connectivity and thus makes the world highly interconnected. This is an opportunity and a challenge as the networks tend to be of millions of nodes (e.g. You Tube/Facebook users, mobile phone owners) and heterogeneous (the nodes include devices and people). Due to low cost transmission technology energy networks have grown continent wide. Transportation networks have also increased globally leading to higher connectivity and richer dynamics. Sensor networks have grown at a tremendous pace in the last decade, integrating humans and sensor devices seamlessly. The proliferation of mobile devices is influencing the way society is evolving as a networked one. During the last few years product networks and economic networks have been explored to study the evolution of economics of different countries. The correlation among suppliers, products, and enterprises is being studied in the past five to six years to make the supply chain system more robust. Fast paced development in transportation systems makes tourism becoming very

 1531-636X/10/$26.00©2010 IEEE

Network representation is critical to the success of modeling engineered networks.

profitable, however globalization and tourism are also the leading causes of disease spreading. Integrated modeling of all these systems is becoming increasingly important. Tools and techniques developed in the past are applicable to networks of tens or hundreds or in extreme cases thousands of nodes. The growth and complexity of the systems described above necessitate the development of the network science grounded principles regarding the representation and analysis of engineered networks.

Network optimization is a well-established field of operations research; algorithms have been developed for improving the performance of real world networks and reduce waste or cost. In this paper, we focus on optimization in various engineered systems. We hope that future large scale engineered networks can be designed by taking the concepts of robustness and optimization into account. In order to optimize complex networks, i) various network representations have to be studied since different representations serves different purposes., ii) network properties have to be measured and quantified so that we can formulate the utility functions for the optimization problem, and iii) network models need to be introduced and analyzed.

This paper is organized as follows: First, we briefly review network science and its relationship to engineered systems. We then describe some applications and follow with a summary and discussion of future research directions.

II. Network Representation

Engineered networks differ from natural networks in their requirement of real time analysis and inference. Therefore, network representation is critical to the success of modeling engineered networks. Moreover, representation influences the computational efficiency of search, clustering, and network optimization. Several network representation techniques are used in the literature [Gross and Yellen, 2004]: node-arc incidence matrix, node-node adjacency matrix, adjacency list and forward star are some commonly used representations.

In a *node-arc incidence matrix,* rows correspond to nodes, and columns correspond to arcs; the column for arc (i, j) has exactly two non-zero entries: $+1$ in row i and -1 in row j. Though this is an intuitive scheme, it wastes memory space, and in general leads to a sparse matrix.

In a *node-node adjacency matrix*, rows and columns both correspond to nodes. In position (i, j), a 1 indicates an arc going from node i to node j. A node-node adjacency matrix is space and time efficient only if the network is dense. The *Laplacian matrix* is a modified version of node-node matrix, $L = [l_{ij}]$, where

$$l_{ij} = \begin{cases} -1, & i \neq j, \quad i \text{ is adjcent to } j; \\ \text{degree(node } i), & i = j \\ 0, & \text{otherwise} \end{cases}$$

This is widely used in graph connectivity analysis because in undirected graphs the Eigen values of this matrix tend to be real and hence interpretable.

In *adjacency lists*, the arc adjacency list $A(i)$ is stored as a linked list for each node i. Each record in the linked list corresponds to an arc (i, j) and stores the following information: (a) The head of the arc j; (b) The cost c_{ij}; (c) The capacity lower and upper bounds l_{ij} and u_{ij}; (d) The pointer to the next record in the linked list. An array of pointers store a pointer to the first record of each linked list. This scheme ensures storage efficiency.

A *forward star representation* stores the same information as an adjacency list does, but using arrays instead of linked lists. This method stores arc information with four elements: tail, head, cost and capacity. Each array has length m, which denotes the number of arcs. The arcs are sorted in increasing order of tail nodes. An array of *pointers* with n rows (n is the number of nodes) stores the information on where to look for arcs emanating from each node. The pointer (i) is the index of the first arc emanating from node i, otherwise point (i) equals point $(i + 1)$ if node i has no arcs out of it. The forward star representation requires less memory and is easy to implement. The disadvantage of this method is that it is difficult to add or delete arcs and the time taken for adding and removing operation is proportional to m, while adding or deleting arcs with adjacency lists is instantaneous [Ahuja et al., 1993].

Table I compares the four representation schemes with respect to network storage, arc and node addition and removal. Many engineered networks are best represented

LiYing Cui (phone: 814-4413913, email: luc5@psu.edu) and Soundar Kumara (phone: 814-8632359, email: skumara@psu.edu) are with the Department of Industrial Engineering at Pennsylvania State University, University Park, PA 16802. Réka Albert is with the Departments of Physics and Biology at Pennsylvania State University, University Park, PA 16802 (phone: 814-8656123, email: ralbert@phys.psu.edu).

Table 1.
Representation of networks.

Representation	Storage	Add or Remove Arcs and Nodes
Node-arc incidence	Not efficient	Easy to operate
Node-node adjacency	Better than Node-arc incidence method	Easy to operate
Adjacency list	Efficient	Not easy to delete a node
Forward star	Efficient	Hard for all the operations

by directed graphs or bipartite graphs. Recently, tripartite hyper graphs have been employed in certain applications. A notable example is folksonomy in which there are users, resources and tags. Flickr and CiteULike are two well known applications in this regard.

III. Characterization of Networks

The structure of networks conveys rich information useful for inference. The last decade has seen a proliferation of topological metrics. We discuss some of the important ones here. The *order* of a network is the total number of nodes (also called vertices), and its *size* is the total number of links (also called edges) in a network. The *degree* of a node is the number of links connecting the node to its neighbors. The incoming (*in-degree*) and outgoing (*out-degree*) sum up to the degree of a node. The *degree distribution* is a 2 dimensional graph showing the frequency of nodes with different degrees in the network. The network *density* is the ratio between network size m and the maximum possible number of links. One of the most important measures that has been explored is the *distance*. The *distance* between two nodes is the length of the shortest path between them, i.e. the minimum number of links that one needs to follow when going from one node to the other. The shortest path can be found through Dijkstra's algorithm. The average path length of a network is the average value of distance between any pair of nodes in the network. The *diameter* of the network is the longest distance between any pair of nodes in a network. The *clustering coefficient of a node* is the number of triangles centered at the node divided by the number of triples centered at the node. The *clustering coefficient of a network* is the arithmetic mean of the clustering coefficients of all the nodes. The *betweenness* centrality quantifies how much a node is between other pairs of nodes. Let us define the ratio between the clustering coefficient and the average path length as the *CP ratio*. The *proximity ratio* of a network is the CP ratio between this network and a random network. This property captures the extent of a network's small-worldness. The *efficiency* of a network is the communication effectiveness of a networked system (global) or of a single node (local). The mixing coefficient is the Pearson correlation between the degrees of neighboring nodes. The *modularity index* measures the topological similarity in the local patterns of linking. Table II summarizes the most commonly used metrics and their equations. [Baggio et al., 2010] show a similar table.

The current network metrics characterize the network at three levels: micro, global, and mesoscopic levels. The properties of individual nodes or edges such as degree, centrality and node rank functions are at the micro level; the properties of the entire network, such as degree distribution, diameter and clustering coefficient are at the global level; the detection of cohesive groups, e.g. communities is at the mesoscopic level. The recently introduced concept of "stochastic block structures" [Reichardt, 2010] is also at the mesoscopic level and it is shown that communities are a special case of this concept in a static network. Stochastic block structures can be applied in studying the evolution of complex networks.

IV. Network Models

Before the computerization of data acquisition and sharing allowed the mapping of real-world networks, several abstract prototypical networks were studied.

The mathematicians Erdös and Rényi analyzed *Random networks* based on a set of nodes (ER model). In this network, the links are added into a network consisting of n nodes and with no links among them. The link between any pair of nodes is placed with a probability of p. The degree distribution of the network follows Poisson distribution with the average node degree of $<k> = np$, thus $P(k) \approx (<k>^k e^{-<k>}/k!) = ((np)^k e^{-(np)}/k!)$. For a large node set, there is a giant connected component if the average number of links per node is larger than 1. The average path length l in this giant connected component scales as $\ln(n)$, for example $l \propto \ln(n)$. Further, $l = \ln(n - <k>)$. In this resulting network, most of the nodes will have a degree of $<k>$ When $p < 1/n$, almost all the vertices in the network belong to isolated trees. Cycles of all orders in the random network appear at $p \approx 1/n$.

Regular networks include rings, lattices, trees, stars and complete graphs. A ring is a connected graph in which a node is linked exactly with two other nodes. A *lattice* is a graph in which the nodes are placed on a grid and the neighbors are connected by an edge. A one dimensional lattice is like a chain. A *tree* is a connected graph containing no circles. A *star* graph is a tree in which every node is connected to the root. In a *full* (*complete*) graph there is an edge between all pairs of nodes.

Table 2.
Metrics for networks.

Metric	Math Equation (in undirected graph)
Order	The number of nodes n
Size	$m = \sum_i \sum_j a_{ij}$, where $a_{ij} = \begin{cases} 1, & \text{node } i \text{ and } j \text{ are linked} \\ 0, & \text{otherwise} \end{cases}$
Node degree	$d = \sum_i a_{ij}$, where $a_{ij} = \begin{cases} 1, & \text{node } i \text{ and } j \text{ are linked} \\ 0, & \text{otherwise} \end{cases}$
Density	$\delta = \frac{2m}{n(n-1)}$, where m is the number of arcs in the network and n is the number of nodes in the network.
Average path length	$l = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}$, where d_{ij} is the distance between node i and node j.
Diameter	$D = \max\{d_{ij}\}$, where d_{ij} is the distance between node i and node j, and n is the number of nodes in the network.
Clustering coefficient of a node	$c_i = \frac{2t_i}{k_i(k_i - 1)}$, where t_i is the number of triangles centered at node i, and k_i is the number of degrees of node i.
Clustering coefficient network	$c = \frac{1}{n} \sum_i c_i$, where c_i is the node degree of node i.
Betweeness	$B_k = \sum_{k \in \nu(i,j)} \frac{1}{n_{ij}}$, where n_{ij} is the number of paths from node i to node j.
Proximity ratio	$\mu = \frac{\frac{C}{l}}{\frac{C_{\text{rana}}}{l_{\text{rana}}}}$, where C is the clustering coefficient of the network, and l is the average path length of the network; C_{rana} is the clustering coefficient of a random network and l_{rana} is the average path length of a random graph.
Efficiency	$E_{\text{global}} = \frac{1}{n(n-1) \sum_{i \neq j} \frac{1}{d_{ij}}}$, where d_{ij} is the distance between node i and node j
Mixing coefficient	$r = \frac{\sum_i (dg_i - \overline{dg})(dn_i - \overline{dn})}{\sqrt{\sum_i (dg_i - \overline{dg})^2 (dn_i - \overline{dn})^2}}$, where dg_i is the degree of node i, and dn_i is the first order mean degree of the neighbors of node i. The mixing coefficient is the standard deviation.
Modularity index	$Q = \sum_i (e_{ii} - a_i)^2$, where e_{ii} is the fraction of edges in subgraph i; a_i is the fraction of edges between this subgroup and all other subgraph.

Simple characterization of engineered networks is not sufficient, optimizing an objective is important.

The *Graph Atlas* [Read and Wilson, 1998] contains all undirected graphs with up to seven nodes. Regular networks are often used in materials science and in parallel computing in computer science.

One of the most important network models is the *small-world network*. The *small-world* concept describes the fact in a network that regardless of size there is a relatively short path between any two nodes. The small-world phenomenon was first mentioned by the writer Frigyes Karinthy, in Hungary, in 1929. 30 years later, it became a research problem named "contact and influence" [Kochen and Pool, 1978] posed the question "What is the probability that two strangers will have a mutual friend?", "When there is no mutual friend, how long would the chain of intermediaries be?" In the seminal work of Watts and Strogatz, the small world concept was expanded to mean networks with low average shortest path length and a high clustering coefficient, which therefore are in a sense situated between regular and random networks. The [Watts and Strogatz, 1998] model (WS model) starts from a ring lattice with n vertices and k edges per vertex, and rewires each edge at random with probability p. This construction tunes the graph from regular graph ($p = 0$) to a random graph ($p = 1$). Small word networks ($0 < p < 1$) have high clustering coefficient and a low average path length. The average path length of this network scales in $\log_{<k>} n$. Here $<k>$ is the average out-degree and n is the number of nodes. When pairs are selected uniformly at random, they are connected by a short path with high probability. Many real world networks exhibit this property. Random networks also have small average distances, however they are not clustered. Small-world networks usually appear in social sciences.

Many systems in real world are dynamic and the order of the networks (number of nodes) grows over time. The concept of a scale-free network was introduced by Barabási and Albert (BA model) in 1999. The term scale-free represents the property of a function that changing the variable in a scale results in changing the function value in scale as well, *e.g.,* $f(ax) = bf(x)$. The power-law form $f(x) = \alpha x^{-\beta}$ is the only form that satisfies the scale-free property. In a scale-free network, the number of nodes n is expected to change over time. In BA model, the network dynamics is described by introducing new nodes into an existing network. When a vertex is linked in, it tends to link with higher probability to a vertex that already has a large number of edges, which is the so called *preferential attachment*. The probability for a new node to be attached to an existing node is proportional to how many links the existing node already has. $\prod(k_i) = (k_i/\sum_j k_j)$. The resulting network has a degree distribution that follows $p(k) \propto k^{-\alpha}$. Many networks in real world have the scale-free property, especially in social and biological sciences. A comparison of random, small-world and scale free networks is shown in Fig.1.

V. Network Optimization

The focus of network science has been to characterize networks and understand the dynamics of their nodes and edges. However engineered systems are built with an objective in mind. Simple characterization of the interconnections alone is not sufficient; instead optimizing the objective becomes critical. We describe some of the most popular network optimization [Yildirim, 2008] models here.

A. Maximum Flow

The maximum flow problem is to find the maximum amount of flow passing from the source node to the sink node. This is equivalent to finding the minimum capacity that needs to be removed from the network so that no flow can pass from the source to the sink. The problem can be formulated as a linear program (LP):

$$\max \sum_{(sj)\in A} x_{sj}$$

subject to

$$\sum_{\{j\mid(i,j)\in A\}} x_{ij} - \sum_{\{j\mid(j,i)\in A\}} x_{ji} = 0, \quad \text{for all} \quad i \in N, \text{ad } i \neq t$$

(flow conservation constraint)

$$0 \leq x_{ij} \leq u_{ij}, \text{ for all } (i,j) \in A,$$

where u_{ij} is the upper bound of the flow on arc (i,j). s and t are the source and target respectively. The maximum flow problem can also be solved by augmenting a path algorithm or by solving the dual of the minimum cut problem. Maximum flow and minimum cut problems are very useful in Internet networks.

B. Minimum Cut

Minimum cut is to find a 2-partition of nodes in a network that minimizes the number of edges or the flow on the edges between two sets. The minimum cut

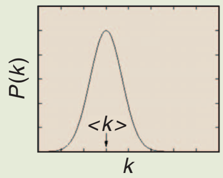
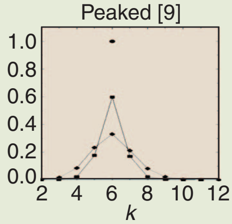
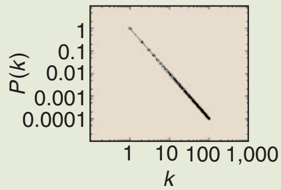
Degree distribution	Poisson 	Peaked [9] 	Power Law 
Characteristic path length	Scales as log(N)	Scales linearly with N for small P. And for higher P scales as log(N).	Scales as log(N)/log(log N)
Clustering coefficient	P (the connection probability)	High, but as $P \to$ l behaves like a random graph.	((m–1)/2)* (log(N)/N) where m is number of edges with which a node enters
Robustness to failures	Similar responses to both random and targeted attacks.	Similar response as random networks. This is due the fact that it has similar degree distribution as random networks.	Highly resilient to random failures while being very sensitive to targeted attackes.

Figure 1. Comparison of Random network, small-world and scale-free networks (H.P. Thadakamalla et al., 2004).

problem is the dual of maximum flow problem. It can be formulated as the dual form of the above linear programming problem:

$$\min \sum_{(i,j)\in A} u_{ij} d_{ij}$$

subject to

$$\begin{aligned} d_{ij} - p_i + p_j &\geq 0,\ (i,j) \in A \\ p_s &= 1 \\ p_t &= 0 \\ p_i &\geq 0,\ i \in V \\ d_{ij} &\geq 0,\ (i,j) \in E. \end{aligned}$$

Minimum cut problem is widely used in reliability engineering.

C. Minimum Cost Flow

A minimum cost flow problem is to find the optimal flow allocation so that the mass can be transferred from the source to the sink. This problem can be formulated as a LP:

$$\min \sum_{(i,j)\in A} c_{ij} x_{ij}$$

subject to

$$\sum_{\{j\mid(i,j)\in A\}} x_{ij} - \sum_{\{j\mid(j,i)\in A\}} x_{ji} = b(i), \text{ for all } i \in N$$

(flow conservation constraint)

$$l_{ij} \leq x_{ij} \leq u_{ij}, \text{ for all } (i,j) \in A,$$

where c_{ij} is the unit cost on arc (i,j); $b(i)$ is the consumption at node i, and l_{ij} and u_{ij} are the lower bound and upper bound of the flow on arc (i,j).

Minimum cost flow problem is important in transportation networks. If the supply and demand at each node is exactly 1, and the number of supply nodes equals to the number of demand nodes, the minimum cost flow problem becomes an assignment problem. The assignment problem can be solved by the Hungarian algorithm.

D. Shortest Path

The shortest path problem is to find a path for the minimum cost flow problem of sending one unit of flow from the source node s to every other node. The LP formulation is:

$$\min \sum_{(i,j)\in A} c_{ij} x_{ij}$$

subject to

$$\sum_{\{j|(i,j)\in A\}} x_{ij} - \sum_{\{j|(j,i)\in A\}} x_{ji} = (n-1), \text{ for } i = s$$

(flow conservation constraint for the source)

$$\sum_{\{j|(i,j)\in A\}} x_{ij} - \sum_{\{j|(j,i)\in A\}} x_{ji} = -1, \text{ for } i \neq s$$

(flow conservation constraint for nodes other than source)

$$x_{ij} \geq 0, \text{ for } (i,j) \in A.$$

The shortest path problem can be solved by Dijkstra's algorithm, which is a labeling method. Finding the shortest path is probably the most popular problem in engineered systems and it occurs in any routing context. Many clustering algorithms are built upon the shortest path problem.

E. Minimum Spanning Tree

In an undirected and connected graph, a spanning tree is a subgraph that connects all the nodes. Given each arc has a cost c_{ij}, the length of the spanning tree is $\sum_{(i,j)\in T} c_{ij}$. The minimum spanning tree problem is to find the spanning tree with minimum length. The LP formulation of this problem is:

$$\min \sum_{(i,j)\in A} c_{ij} x_{ij}$$

subject to

$$\sum_{(i,j)\in A} x_{ij} = (n-1)$$

(to guarantee no cycles in the network)

$$\sum_{(i,j)\in A(S)} x_{ij} = |S| - 1$$

(to guarantee no cycles in the subgraph)

$$x_{ij} \in \{0, 1\}, \text{ for } (i,j) \in A.$$

The minimum spanning tree problem can be solved by Kruskal's algorithm [Kruskal, 1956]. This problem is widely used in transportation network design.

F. Minimum Cover

Given an undirected network, the *minimum* cover problem is to find the minimum number of nodes that can touch all edges. This problem is as hard as the unique game problem (Khot and Regev, 2003). An unique game problem can be introduced as an unique label cover problem, which is to find a color assignment for each node so that as many as the edges satisfy the predefined constraints. The minimum cover problem is usually applicable in setting minimum number of policies in transportation problems and map generating.

G. Hierarchical Tree Generation

In many engineered systems, a hierarchical framework needs to be extracted to understand the essential properties of the system. However, a hierarchical framework is rarely exhibited in general networks other than governments and military organizations. In complex real world networks, the hierarchical tree detection is equivalent to sorting the nodes and edges in the network. When we store the network using adjacency matrix, this problem is transformed into finding the lower triangular matrix. An optimization formulation can be written based on the in and out degrees of individual nodes. The solution is similar to Google's page-rank method [Crofts, 2010].

[Grady et al., 2010] introduced the concept of shortest-path tree in computing the networks' tomogram, e.g. backbone, and applied it to dollar bill tracking of www.whereisgeorge.com and the world wide airline network.

H. Maximize Communication Efficiency

In communication or transportation networks, designers wish to maximize the efficiency of the networks. [Krioukov, 2008] found a connection between heterogeneous topology and the hyperbolic geometry of the metric space of the networks. If a heterogeneous network has strong clustering on top of the space, the metric space is a hyperbolic space, and vice versa. This tells us that hyperbolic metric space controls network heterogeneity. This work shows that the efficiency of communication or transportation networks can reach the maximum only if the metric space is hyperbolic.

I. Control

We consider network control and design problem as an approximate optimization problem, since the purpose of control and design is to optimize the systems.

Network motifs are the building blocks in complex networks, whether engineered or natural. We may consider motifs as an important network metric in addition to the metrics introduced in Section III. Different motifs have different functions. Most of the simple motifs occur many times in a network, which makes the network function in a certain pattern. The basic motifs include positive auto-regulation, negative auto-regulation, feed-forward loop (FFL). Each network motif has different specific information processing functions [Alon, 2007]. These functions were analyzed using mathematical models and tested with dynamic experiments in living cells. Our thesis is that if we can identify the basic motifs in general classes of engineered networks, we will be able to study the dynamics of the networks. That is the fundamental question relating network's structural changes over time to functionalities may be possible. We introduce some of the basic motifs here.

In biological networks, positive *auto-regulation* (Fig. 2 (3)) makes a transcription factor to enhance its own rate of production. In living cell experiments, this motif slows the response time compared to simple regulation (Fig. 2 (1)), and increases cell-cell variation in these experiments. In engineered systems positive auto-regulation may lead to a better response to stochastic environments.

In biological networks, *negative auto-regulation* (Fig. 2(2)) makes a transcription factor to decrease its own rate of production. It accelerates the response time compared with simple regulation and decreases the cell-cell variation in biological living cell experiments. In engineered systems, the behavior of negative auto-regulation needs to be tested. System noise could be reduced and the system could be synchronized by negative auto-regulation.

The *feed-forward loop (FFL)* consists of three regulations: X regulates both Y and Z, and Y regulates Z. Since all the three regulations can represent either activation or repression, there are $2^3 = 8$ different types of FFL as shown in Fig. 3. Since both X and Y regulate Z, there are several different cases, including an "AND" gate (if both X and Y are needed to regulate Z), an "OR" gate (if either X or Y is sufficient to regulate Z), or a SUM function. In the experiments of living cells, two types of FFL occur more frequently than other types. They are coherent type 1 and incoherent type 1.

Coherent type 1-FFL is a sign sensitive delay motif. Both X and Y accelerates Z. When Z has an "AND" gate with X and Y, it shows a delay in stimulation stage and no delay in stop stage. When Z has an "OR" gate with X and Y, it shows no delay in stimulation stage but a delay in stop stage.

Incoherent type 1-FFL is a pulse generator and response accelerator. In this motif, X activates Z, but also represses Z by activating Y. This causes a pulse like dynamics. After stimulation, production of Z starts, and then soon after Y starts to be produced, Z is repressed. If the initial inputs are strong, Z may reach a non-zero steady state.

In addition to the simple motifs above, there are several more complicated motifs which need further investigation and are: Feed-back loops (FBL), multi-output FFL, single-input modules and dense overlapping regulons. The *feed-back loop (FBL)* is a motif that X and Y regulates Z, and Z regulates X. *Multi-output FFL* is a motif in which the regulators X and Y may regulate many Zs at the same time. However, the strength that these regulations to Zs may be different and this results in a temporal order in the outputs. The *single-input motif* is the case that a single regulator regulates a set of genes simultaneously. *Dense overlapping regulons* occur when a set of regulators combinatorially control a set of genes. This motif is a complex regulation pattern with dense links.

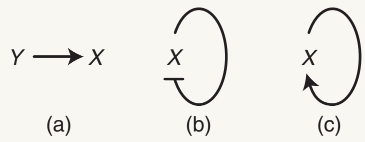
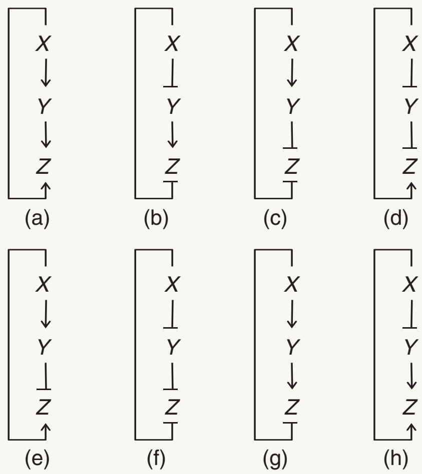


Figure 2. Simple, negative auto and positive auto regulation.



Figuer 3. Eight feed-forward loops.

These motifs were observed from living cells, and they are widely present in engineered systems such as circuits, power grids and internet [Milo et al., 2002]. Circuits and gates are well studied in engineering, and these efforts have and will continue to bring guidance to the motif study in network science. From this perspective, engineered systems brought new ideas and direction to the development of network science. Further work is required to relate the characteristics of motifs to temporal models. This may help in generating some structural inferences on engineered networks similar to the ones carried out in linguistics [Sinatra, 2010].

Other than network study and design via motifs, there is an alternative approach that seeks the conditions that need be imposed on the local network structure or dynamics to generate a desired behavior of the networks. For example, network failure may be rescued by adding or removing targeted nodes or edges. This issue is now known as synthetic rescue or cascading control and network stabilization [Motter, 2010].

In summary, the minimum cost flow networks are frequently used to find an optimal flow allocation through the system at a lower cost in transportation and supply chain networks. The minimum spanning tree problem usually occurs in transportation design. Minimum cut and maximum flow problems are used in reliability engineering to analyze failures. Shortest path problem is common in many routing problems in engineering field, and the clustering algorithms were usually developed based on it. The hierarchical tree or shortest path tree detection is an important research area nowadays. We consider network control and design problem as a quasi-optimization problem. In complex networks, control based on motifs is an emerging research field. Many engineered network optimization problems may not be able to be transformed into one of the above problems. A more generalized utility function may need to be calculated. Some problems involve temporal network representations, which we discuss in the following section.

VI. Temporal Model

We have described the most popular network models in Section IV. What happens when we introduce time into these network models is an important question. The definition of static network $G(V, E)$, where V is the a set of nodes and E is a set of edges can be extended for constructing temporal networks. A temporal network is an undirected graph $G(V, E)$ in which each edge has a time label $\lambda(e)$ representing the time when the two end nodes of the edge come into contact to communicate. In general, each edge can have multiple paths P in G in time respecting if the labels on the edges of the path are non-decreasing.

A temporal network is constructed as follows: A static network is one in which we have single label per edge. Every edge in a temporal network has a set of labels with upper and lower bounds of each labeled interval as $\lambda_{\min(e)}$ and $\lambda_{\max(e)}$. At each discrete time step t between $\lambda_{\min(e)}$ and $\lambda_{\max(e)}$, a node $v(t)$ is joined to a copy of node $v(t+1)$ by a directed arc. Now, finding a critical path from $v(t)$ to $v(t+1)$ is to find the path in this time-expanded network between these two nodes. This definition transforms a dynamic network into a static network. This representation is applicable to the case when the edges have multiple time labels.

Temporal models are useful in epidemic modeling. When a person is infected by a communicable disease at time t and then becomes a non-contagious carrier at time $t + x$, this temporal model can be applied. This model can also be used in network routing with time-dependent edge delay functions. It allows us to add edges in the temporal networks on the definition of static network. Then we can use the tools that we have in static networks to analyze temporal networks. For periodical systems, some results are easily found.

Temporal models become more and more important in resilience, cascading and synchronization analysis. Many other temporal models can be built implicitly with the concepts mentioned above. [Motter and Lai, 2002] have studied cascading failures of overload. Their model demonstrates how a small number of highly overloaded nodes trigger global cascading failures. [Motter, 2004] introduced a strategy of defense to selectively further remove nodes or edges right after the initial attacks and failure so that the propagation of cascading can be prevented. Temporal modeling is an open area worthwhile to study. However, many real temporal models have no analytical solution, and their computational time-complexity is high. Whitney studied cascades of random dynamic networks with a threshold rule [Whitney, 2009]. In this work, a dynamic Markov model is formulated for a finite network with arbitrary average node degree. The problem gives a threshold rule, e.g., a node flips its state only when a proportion of its neighbors has changed their states.

VII. Engineered Networks

Engineered networks are the consequences of proliferation of information technology and globalization. Due to IT connectivity and reach, the world is becoming a networked one. How can we use the knowledge that is gained so far in network science in these new application areas of engineered systems? We consider a few prototypical examples such as supply chains, energy and power grids, and services. The last decade has seen many firsts in these areas with respect to applying network science principles. In this section we discuss ten engineered network applications and hope to stir the readers' interest in pursuing these lines of research.

A. Internet Networks

Information sharing and retrieval drives the day-to-day business across the world. This need has propelled research interest on technological networks such as the Internet and the WWW. The map of the Internet is considered at two different scales. At the level of Autonomous Systems (an AS is an organizational unit of a particular domain or service provider), the edges represent physical communications that connect sub-networks or devices across these ASs. The exchange of information between devices or sub-networks is done using routers. The map of the Internet at this scale consists of nodes as routers and their communications within and across ASs as edges.

[Faloutsos et al., 1999] have analyzed the Internet AS network individually on data collected from 1997 to 1999. During these years, the number of nodes and edges increased from 3112 to 10,100, and the average degree remained constant. This was also the case with average path lengths, which was found to be approximately 3.8 for all the three years. Furthermore, the path length distribution is peaked around the average value and the shape remains essentially unchanged over the three years. On the other hand, one property of the network that did change over the years was the clustering coefficient. It increased from 0.18 in 1997 to 0.24 in 1999. The distribution of the clustering coefficients is a power-law decaying function of the degree on the nodes. All properties except the distribution of the clustering coefficient are similar for the Internet network at the router level. The clustering coefficient is independent on the degree of the nodes at the router level. One can safely state that the research in this realm is ongoing and the focus is on survivable internet, which can realize the objective of navigation under node and link failures.

[Lee at el., 2007] worked on the concept of maximum relatedness sub network, which captures the most essential relation for the entities on World Wide Web, including the hidden correlation between any two objects on the web.

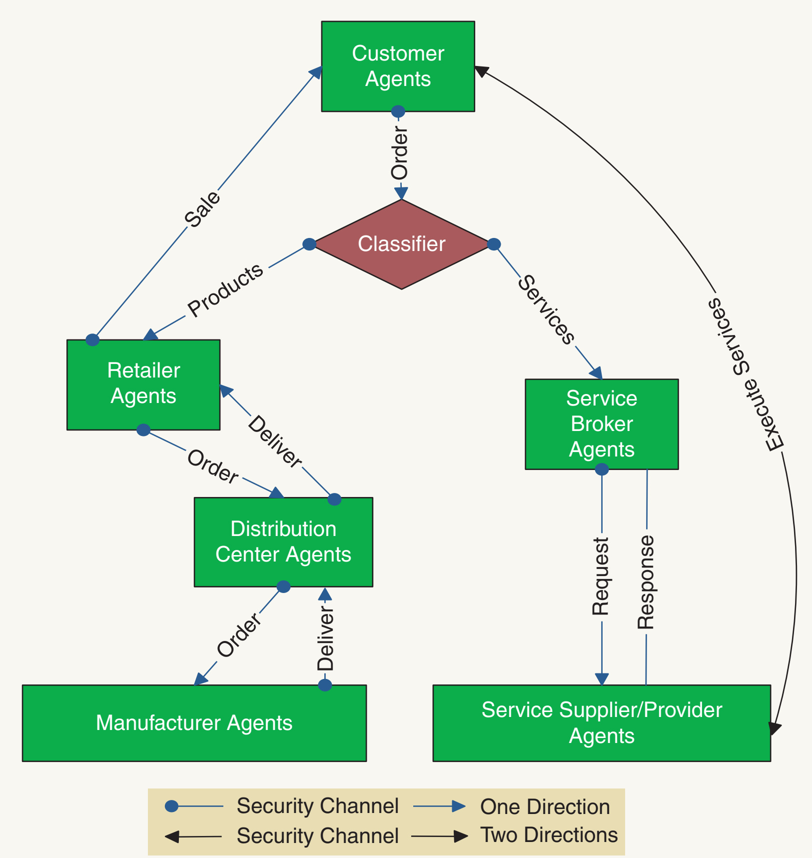


Figure 4. Agent based supply chain networks.

B. Supply Chain Networks

Each multi-product, multi-stage, multi-supplier, and multi-customer supply chain is a system of systems. [Chen et al., 2004] studied the supply chain problem with multiple incommensurable goals for a multi-echelon supply chain network with uncertain market demands and product prices.

From the customers' (manufacturers') point of view, many supply chain problems can be modeled as minimum cost flow problems. Either mathematical programming or heuristic methods can be used to solve them. From the suppliers' point of view, it can be a maximum cover problem. Fig. 4 shows an agent based architecture of supply chain networks. In these networks, customers, retailers, distribution centers, manufactures and even service providers and brokers are involved. The network is complex, and the scale of transportation network usually is national or world-wide. The work reported in [Thadakamalla et al., 2004] studies supply chains from a survivability point of view and lays the foundation for using network science in this field.

C. Service Networks

The number of organizations contemplating the integration of services into their strategic plans and daily operations is continuously increasing. Companies and experts tend to offer their expertise to other companies, organizations and individuals so that the knowledge and resources can be fully utilized and well allocated among the whole society Numerous initiatives on service outsourcing have provided incentives for organizations to become more benign. One popular area that continues to gain importance is one that focuses on the external relationships among organizations. Service composition plays an important role in enterprise integration. From the customers' point of view, it is an optimization problem with reliability, speed, quality and cost of composition as the objectives, which can be transformed into a maximum flow problem. From the service providers' point of view, it will be a maximum cover problem. [Cui, et al., 2009] proposed a schedule of service composition based on network structure. Two services are connected if they share common inputs or outputs. Fig. 5 shows a service network with betweenness centrality analysis. The nodes with larger diameter have large betweenness, indicating the popularity of these services. [Engelen, 2005] studied the web service repository of XMethod, and claimed that web service network is scale free.

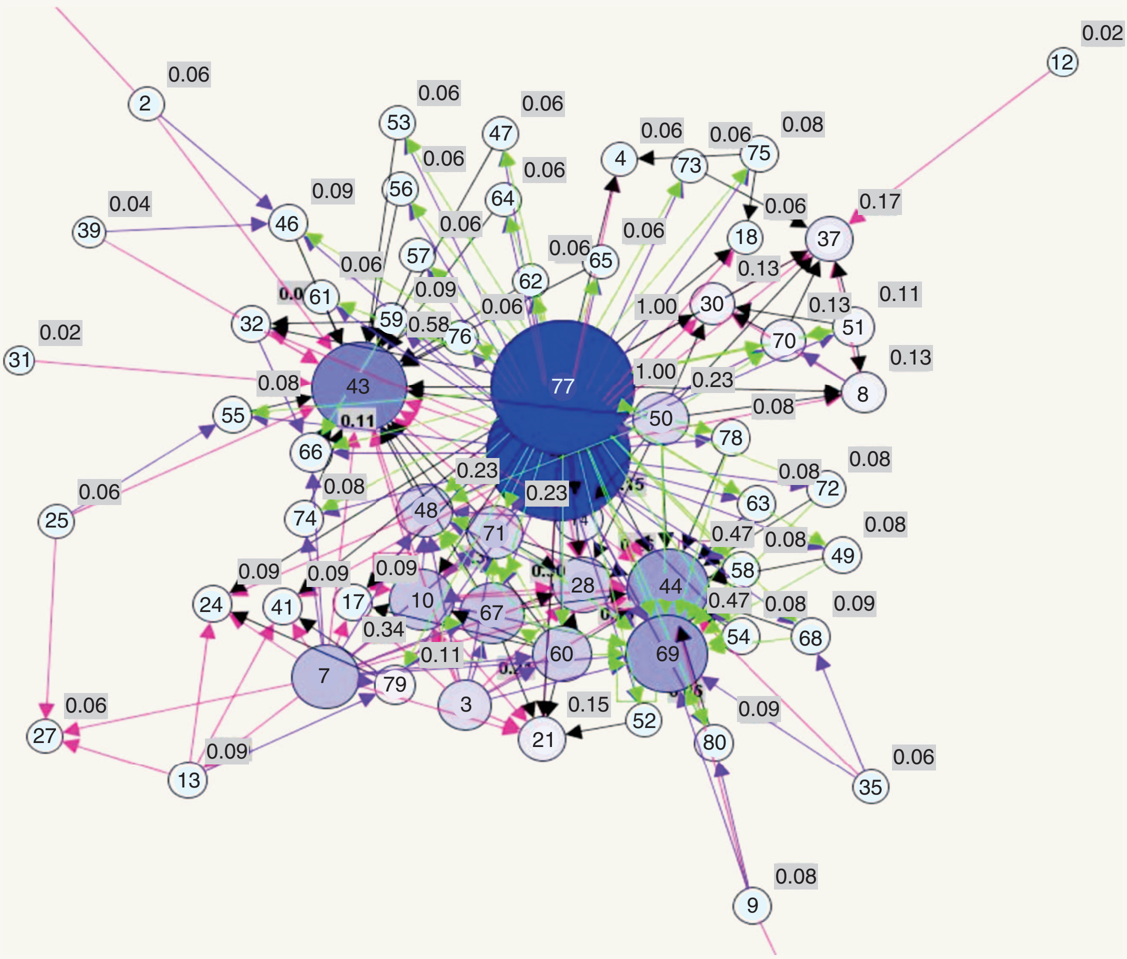


Figure 5. Web service centrality analysis: Nodes are the web services, and there is a link between two services if the input of one service can be used as the output of another service.

D. Energy Networks and Power Grid

Network science tools are used in analyzing the failure of power grids. [Albert et al., 2004] studied the structural vulnerability of the North American power grid. They demonstrated that the complexity and interconnectivity of the electric infrastructure determines the vulnerability to perturbations in the power grid.

In energy networks, the Minimum-Energy Broadcast problem is to assign a transmission range to every station in the energy distribution and consumption network so that a given source station is allowed to perform broadcast operations, and the overall energy consumption for the range assignment is minimized. [Calamoneri et al., 2008] proved a tight asymptotical bound on the optimal cost for the Minimum-Energy Broadcast problem on square grids. It is difficult to achieve optimal assignment for this problem on regular and large-scale instances.

In sensor networks [Slijepcevic et al., 2001], the maximum cover problem is to deploy the sensor nodes in clusters so that maximum location spots be fully guaranteed to be monitored. Under stochastic conditions, the number of nodes that must be deployed in order to completely cover the whole monitored area is often higher than the ones needed if a deterministic procedure were to be used. In a sensor network, with stochastically placed nodes, activating only the necessary number of sensor nodes at any particular moment can save energy. Network science is useful in the risk analysis of power grids and energy networks [Albert et al., 2004]. An approach similar to [Thadakamalla et al., 2004] will help in the survivability analysis of power grid networks.

E. Transportation Networks

In transportation networks, minimum cost flow problem deals with assigning the mass through the traffic from the origin to destination so that the total cost can be minimized. The costs and origin-destination requirements are modeled as constraints.

In the transportation network design problem, the maximum edge coverage problem is useful ito decide the placement of sensors or monitors. Braess' paradox needs to be considered in design. Network science can contribute to risk analysis of the transportation system. An effective way of modeling will be to combine the transportation problem with the supply chain problem. The flows can be modeled as weighted arcs with heterogeneous nodes. Weighted search [Thadakamalla et al., 2007] will be an important methodology that needs to be explored in this realm. Another important question to ask is "are there any basic motifs in the network that can be related to the temporal demand dynamics?"

[Braunstein, 2010] studied network congestion via the defined nodal degree-degree coefficient e.g. Pearson function, or mixing coefficient. It was found that the congestion of a dissortative network (In dissortative networks, nodes tend to connect with dissimilar nodes) is lower than that in uncorrelated networks; and the congestion of an assortative network (In assortative networks, nodes tend to connect with similar nodes) is higher than that in uncorrelated networks, which gives a new insight into the road network design.

F. Tourism Networks

Tourism networks have been relatively neglected as an area of academic research. [Morrison et al., 2004] present research findings focusing on international tourism networks. Tourism networks in relation to learning, exchange, business activity and community have been studied [Gaggio et al., 2010]. Network science can help in finding significant success factors and consequent management implications, and community learning to tourism destinations. The bipartite network between the scenery or sports and human communities can be used to predict the dynamics of tourism thus help estimate the profitability for some countries that depend on tourism. Another important aspect of this problem is to study tourism networks in the context of infectious disease propagation.

G. Communication/Mobile Networks

Monitoring infrastructure of communication among people through sensors leads to the formation of mobile networks which can be dynamically changing. These changes may include the change in the location of the nodes and new nodes entering the network. Due to the volume of traffic and devices, signature collection, storing and analyzing the data becomes a difficult proposition. How to establish connections, in the event of disruptions, how to find alternate routes and most importantly are there predictable patterns in the evolution of these networks as well as in the networks themselves. Song and Barabási [Song et al., 2010] recently analyzed cell phone data and reported that human mobility patterns are highly predictable. The commute pattern can be predicted accurately for 97% of the population studied in the experiment. It is possible that such a result may be viable for general communication networks, which need to be studied further.

H. Crowd Sourcing

In the past few years many design activities are being crowd sourced. Given the magnitude of the crowds it becomes very difficult to discover experts and gather communities of common interests. Crowd sourcing can be used in the areas of marketing activities, expert discovery, problem solving, project planning, logistics, knowledge management *etc*. Collective innovation is produced both as an aggregated byproduct of everyday information consumption. [Kozinets et al., 2001] discussed the collective consumer innovation on transforming the nature of consumption, with the diffusion of networking technologies. Four types of online creative consumer communities—Crowds, Hives, Mobs, and Swarms are studied in the work.

We posit that given a crowd sourcing platform, we can construct networks of relationships among people, for example by defining edges as common interest or expertise. Analysis of such Crowd Networks may yield insights into communities. Modularity measures and community metrics will be very useful to draw conclusions on the networks. One of the important issues will be the development of fast community detection algorithms [Raghavan, 2007] due to the large size of the crowds.

I. Malware Heredity and Provenance Detection

The concepts of evolutionary biology and genetics are applicable to analyze digital artifacts in cyber engineering. [Cui, et al., 2010] proposed an original approach to establish the provenance and heredity of digital artifacts in general, and malware in particular. In biology, heredity is the genetic transmission of characteristics from parent to offspring.

Defining the equivalent of genes in binary code as CyGenes, a variant (descendant) will contain one or more of the original Cygenes as well as one or more altered, mutated Cygenes that are present in the parent code. Thus the individual genes that are present in a binary artifact will testify to the provenance of the code and will reveal previously untraceable information, such as the identity of the programmer, characteristics of the development environment, compiler features, traits of libraries, and details of the execution environment. Fig. 6 shows a malware network constructed from Cygenes. Green nodes

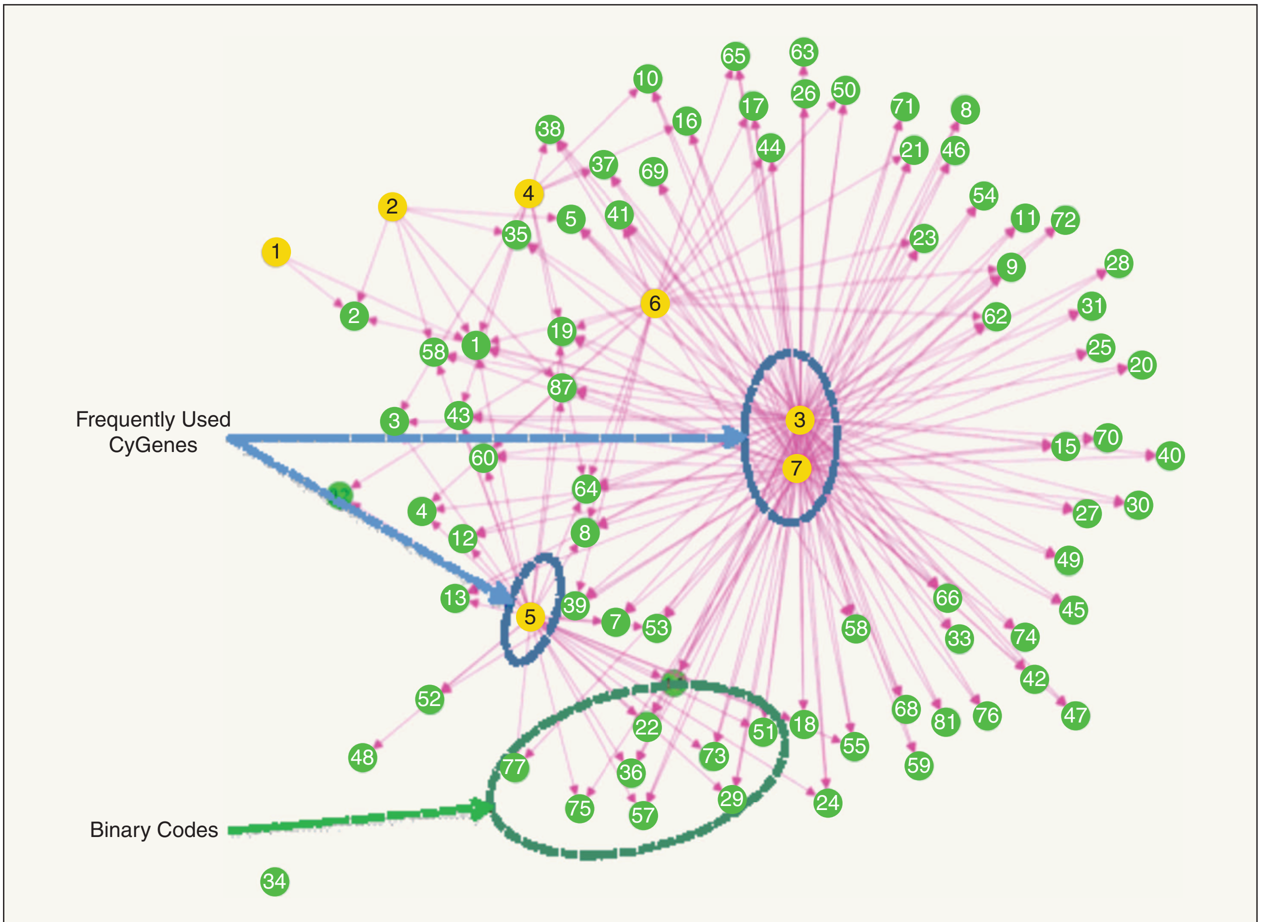


Figure 6. Malware networks: Green nodes represent malware in binary form, and the yellow nodes represent Cygenes.

are malware, and the yellow nodes are Cygenes. If a Cygene presents in a piece of malware, there is a link from the Cygene towards the malware. Our experiments on the malwares before 2002 show that many computer viruses repeatedly have some common pieces of codes shared among them and there is always a giant connected component in the network.

J. Product Networks

Retail business involves a large number of entities including internal products, external products from competitors, customers and point of sales. It is intuitive that the interrelations among these entities need to be considered in smart decision making, which leads network science to be put into use. For example, when two products in a product portfolio of a company have similar function, customers who buy one of them most likely will not buy the other one, thus these two products are negatively correlated in sales. On the other hand, two products with a similar function may positively be correlated when they serve people of different age or gender in a family. These products can be represented as nodes in a network, and the links among the nodes represent the correlation among the products. Several network science techniques and instances are useful to study the mutual influence of the products. As a result, the advertising and promotion frequency, amounts and methods can be developed to improve profitability.

Fig. 7 shows a product network constructed from a family of consumer goods. When a product's profit influences another product's profit positively, there will be a link between them. This figure shows the positive influence of two categories of consumer goods [Cui and Kumara, 2010]. The experiments on product portfolio management for a consumer goods company shows that profit could be improved on the average by 40% by considering the interactions among products and POSs via network science.

Aral and Walker worked on examining the peer effects by randomized trials of peer-to-peer products [Aral and Walker, 2010]. The peer influence was created by both passive and active broadcast through Facebook with 9687 users.

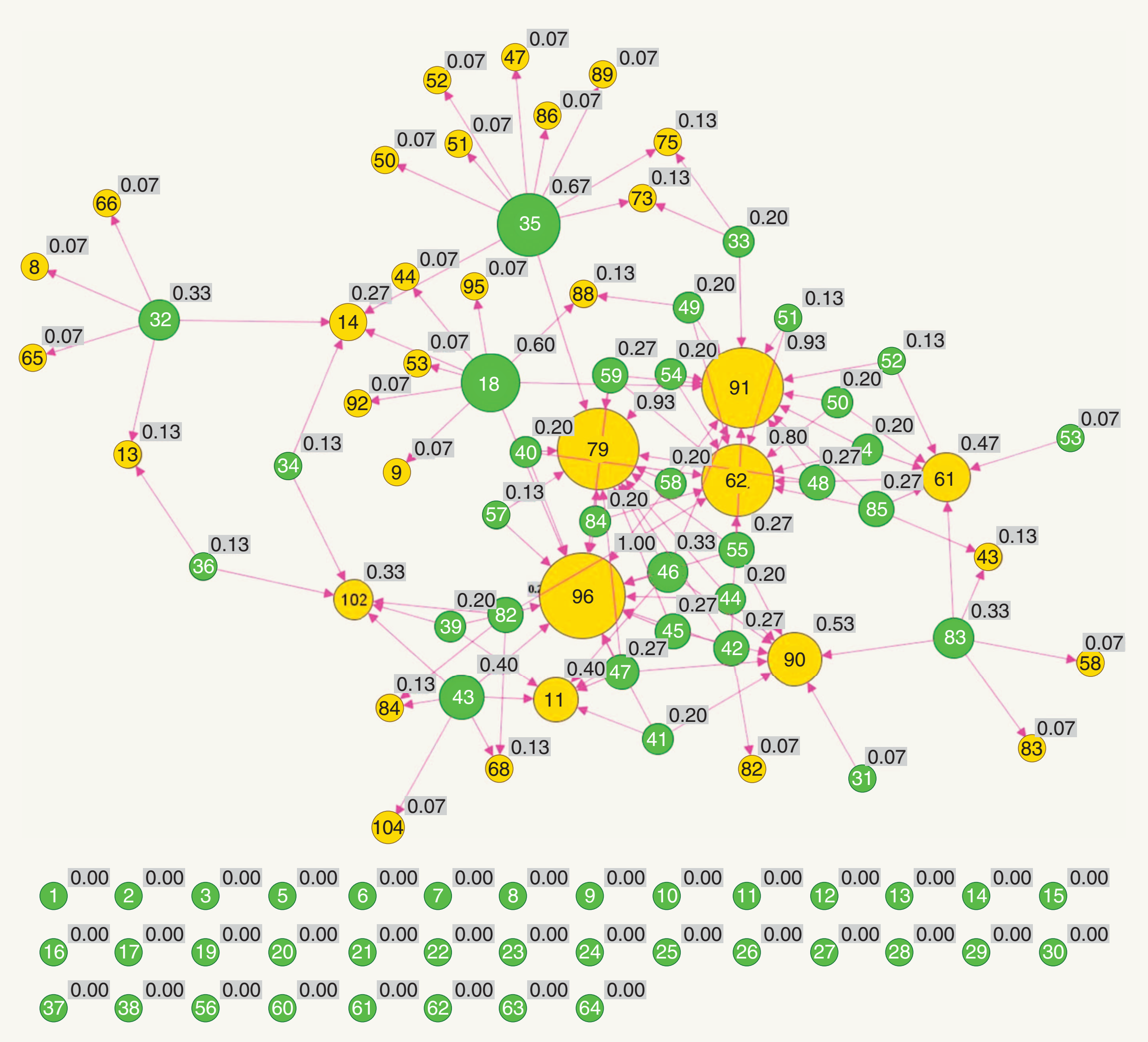


Figure 7. Product network: Green nodes: products of one category; Yellow nodes: products of another category.

Hidalgo et al studied the world economy as networks consisting of products, countries and companies. They found that the diversity of the products of the countries correlates with their economy level. High diversity indicates high economy level and vice versa [Hidalgo et al, 2009].

A summary of engineered networks is shown in Table III. The scale of the networks, the model, optimization problem and the contribution of network science mechanism to the area are shown in this table.

VIII. Conclusions and Future Work

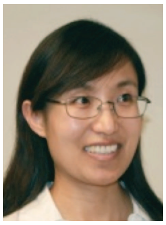
Our motivation in this paper is to introduce new researchers into the exciting area of network science and expose them to several interesting and promising applications with respect to engineered networks. The last decade has seen an increased interest in network science from all disciplines—biology to engineering to sociology. In this paper, we have explored several domains of engineered systems and pointed out some of the important research issues. In engineered networks, the topics of network representation, finding motifs, combining optimization techniques with network science, and addressing dynamics are critical. Based upon our own study in the last decade on engineered networks and our familiarity with the literature, we conclude with the following:

(1) In engineered systems, we are not simply trying to understand the properties and topological structures of networks, but we would like to get feasible and optimal solutions for desired objectives. Due to

Table 3.
Applications of networks.

Engineered System	Scale	Network Model	Optimization Problem	Network Science Contribution
Supply chain	Large, often global	Not known	Minimum cost flow; Minimum cover; Game theory	Cascading analysis
Service	Large, often global	Scale free	Minimum cost flow; Minimum cover; Game theory	Cascading analysis, Searching improvement
Energy	Large, mostly country wide	Scale free	Minimum cost flow; Minimum cover; Game theory	Cascading failure analysis
Transportation	Large, global	Not known	Minimum cost flow; Minimum cover; Game theory	Routing, Risk analysis
Tourism	Large, global	Not known	Minimum cost flow; Minimum cover; Game theory	Cascading analysis, Pattern identification
Mobile	Large, global	Not known	Minimum cost flow; Minimum cover; Game theory	Cascading analysis, Pattern identification
Crowd sourcing	Large, global	Not known	Minimum cost flow; Minimum cover; Game theory	Pattern identification, Search improvement
Products	Small, mediate, or global	Not known	General utility function	Risk analysis, Pattern identification
Internet	Large scale, global	Scale free	Minimum cost flow; Minimum cover, and other general utility function	Pattern identification, Search improvement
Malware detection	Global	Not known	Hierarchical tree generation	Pattern identification, community detection and cascading analysis

this reason, decomposition of large engineered networks into sub graphs and solving them using math programming techniques is very important. (2) Efficient network representation of engineered systems is crucial as we will be looking at real-time or near real-time analysis and inference. (3) The notion of motifs will be beneficial in deriving general principles of control in engineered networks. (4) Advanced search algorithms are critical in the study of networked systems. (5) Network dynamics from both granularity and time scales need to be investigated and effectively combined in large-scale engineered systems. Such enquiries may help advance network science in the near future.



LiYing Cui is a Ph.D. candidate in the department of Industrial Engineering at Pennsylvania State University. She received her B.S. degree in Mathematics and M.S. degree in Operations Research and Control from Tianjin University, China. She is a recipient of Bronze Medal in 1995 at the National Math Competition "Hope Cup" during her 10th grade in China. Her research interests are in Service Computing, Retail Intelligence, Optimization and Complex Networks.

Soundar Kumara is the Allen, E., and Allen, M., Pearce Professor of Industrial Engineering at Penn State. He also holds a joint appointment with the Department of Computer Science. He holds an Adjunct Professor position with C.R. Rao Institute of Advanced Mathematics, Statistics and Computer Science, University of Hyderabad, India. His research interests are in engineered large-scale networks, sensor networks, and web services. He is a Fellow of Institute of Industrial Engineers and Fellow of the International Academy of Production Engineering (CIRP).

Réka Albert is a Professor in the Department of Physics and the Department of Biology at the Pennsylvania State University, and is affiliated with the Huck Institute of the Life Sciences, Center for infectious disease dynamics, and College of Information Science and Technology at the Pennsylvania State University. Her main research interest is in modeling the organization and dynamics of complex networks. She received her PhD in Physics from the University of Notre Dame. She is a Fellow of the American Physical Society and a member of the Society for Mathematical Biology.

References

[1] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[2] A.-L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: The topology of the World Wide Web," *Physica A*, vol. 281, pp. 69–77, 2000.

[3] A.-L. Barabási, "Scale-free networks: A decade and beyond," *Science*, pp. 325–412, 2009.

[4] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the World Wide Web," *Nature*, vol. 401, pp. 130–131, 1999.

[5] R. Albert and A.-L. Barabási, "Dynamics of complex systems: Scaling laws for the period of Boolean Networks," *Phys. Rev. Lett.*, vol. 84, pp. 5660–5663, 2000. works: Local events and universality," *Phys. Rev. Lett.*, vol. 85, p. 5234, 2000.

[7] R. Albert, A.-L. Barabási, H. Jeong, and G. Bianconi, "Power-law distribution of the World Wide Web," *Science*, vol. 287, p. 2115a, 2000.

[8] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance in complex networks," *Nature*, vol. 406, p. 378, 2000.

[9] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, 2002.

[10] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Phys. Rev. E*, vol. 69, p. 025103(R), 2004.

[11] N. Gnanasambandam, S. Lee, N. Gautam, S. R. T. Kumara, W. Peng, V. Manikonda, M. Brinn, and M. Greaves, "Reliable MAS performance evaluation using queueing models," in *Proc. IEEE Multi-Agent Security and Survivability Symp.*, Aug. 2004.

[12] Y. Hong, N. Gautam, S. R. T. Kumara, A. Surana, H. Gupta, S. Lee, V. Narayanan, and H. Thadakamalla, "Survivability of complex system—Support vector machine based approach," in *Proc. Int. Conf. Artificial Neural Networks in Engineering (ANNIE'02)*, 2002, pp. 153–158.

[13] H. Jeong, Z. Neda, and A.-L. Barabási, "Measuring preferential attachment for evolving networks," *Europhys. Lett.*, vol. 61, pp. 567–572, 2003.

[14] J. Kleinberg, "The small-world phenomenon: An algorithmic perspective," in *Proc. 32nd ACM Symp. Theory of Computing*, 2000.

[15] M. E. J. Newman, "Models of small world," *J. Statist. Phys.*, vol. 101, pp. 819–841, 2000.

[16] H. P. Thadakamalla, U. N. Raghavan, S. R. T. Kumara, and R. Albert, "Survivability of multiagent based supply networks: A topological perspective," *IEEE Intell. Syst.*, vol. 19, no. 5, pp. 24–31, 2004.

[17] D. J. Watts, P. S. Dodds, and M. E. J. Newman, "Identity and search in social networks," *Science*, vol. 296, May 2002.

[18] D. J. Watts, "The 'new' science of networks," *Annu. Rev. Sociol.*, vol. 30, pp. 243–270, 2004.

[19] R. V. Engelen. (2005, June). *Are web services scale free?* [Online]. Available: http://www.cs.fsu.edu/~engelen/powerlaw.html

[20] J. L. Gross and J. Yellen, *Handbook of Graph Theory*. Boca Raton, FL: CRC, 2004.

[21] U. Alon, "Network motifs: Theory and experimental approaches," *Nature Rev. Genet.*, vol. 8, pp. 450–461, 2007.

[22] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, Oct. 2002.

[23] R. Baggio, N. Scott, and C. Cooper. (2010, Feb.). *Network science: A review focused on tourism* [Online]. Available: http://cdsweb.cern.ch/record/1245639

[24] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, pp. 269–271, 1959.

[25] S. Khot and O. Regev, "Vertex cover might be hard to approximate to within 2-∈," in *Proc. 18th Annu. Conf. Computational Complexity (CCC'03), IEEE Computer Society*, 2003, pp. 379–386.

[26] T. Calamoneri, A. Clementi, M. Lauria, A. Monti, and R. Silvestri, "Minimum-energy broadcast and disk cover in grid wireless networks," *Theor. Comput. Sci.*, vol. 399, no. 1–2, pp. 38–53, June 2008.

[27] S. Slijepcevic and M. Potkonjak, "Power efficient organization of wireless sensor networks," in *Proc. IEEE Int. Conf. Communications*, 2001, vol. 2, pp. 472–476.

[28] A. Morrison, P. Lynch, and N. Johns, "International tourism networks," *Int. J. Contemp. Hosp. Manage.*, vol. 16, no. 3, pp. 197–202, 2004.

[29] R. V. Kozinets, A. Hemetsberger, and H. J. Schau, "The wisdom of consumer crowds: Collective innovation in the age of networked marketing," *J. Macromarket.*, vol. 28, no. 4, pp. 339–354, 2008.

[30] A. Percus, G. Istrate, and C. Moore. *Computational Complexity and Statistical Physics*. New York: Oxford, 2006.

[31] L. Y. Cui and S. R. T. Kumara, "Binary code heredity and provenance detection based on networks," in *Proc. NetSci Conf.*, Boston, MA, May 2010.

[32] L. Y. Cui, S. R. T. Kumara, and T. Yao, "Service composition based on networks and mathematical programming," in *Proc. INFORMS Annu. Conf.*, San Diego, CA, Oct. 2009.

[33] L. Y. Cui, S. R. T. Kumara, J. Yoo, and F. Cavdur, "Large-scale network decomposition and mathematical programming based web service composition," in *Proc. IEEE Conf. Commerce and Enterprise Computing*, 2009, pp. 511–514.

[34] S. K. Moon, T. W. Simpson, L. Y. Cui, and S. R. T. Kumara, "A service based platform design method for customized products," in *Proc. CIRP Integrated Production and Service Systems-II*, Sweden, Apr. 2010.

[35] H. P. Thadakamalla, S. R. T. Kumara, and R. Albert, "Complexity and large-scale networks," in *Operations Research and Management Science Handbook*, A. Ravindran, Ed. Boca Raton, FL: CRC, 2008, ch. 11.

[36] M. B. Yikldirim, "Network optimization," in *Operations Research and Management Science Handbook*, A. Ravindran, Ed. Boca Raton, FL: CRC, 2008, ch. 4.

[37] A. E. Motter and Y. C. Lai, "Cascade-based attacks on complex networks," *Phys. Rev. E*, vol. 66, p. 065102(R), 2002.

[38] A. E. Motter, "Cascade control and defense in complex networks," *Phys. Rev. Lett.*, vol. 93, p. 098701, 2004.

[39] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *Proc. SIGCOMM*, 1999, pp. 251–262.

[40] C. L. Chen, "Multi-objective optimization of multi-echelon supply chain," *Comput. Chem. Eng.*, vol. 28, no. 6–7, pp. 1131–1144, 2004.

[41] L.Y. Cui, S. R. T. Kumara, et al. "Network based Marketing Intelligence," unpublished.

[42] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[43] U. N. Raghavan, R. Albert, and S. R. T. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Phys. Rev.* vol. 76, no. 3, p. 036106, 2007.

[44] S. Aral and D. Walker, "Viral product design: Testing social contagion and peer influence effects in networks using randomized trials," *Working Paper*, 2010.

[45] C. Hidalgo, "The dynamics of economic complexity and the product space over a 42 year period," *Working Paper*, 2009.

[46] S. H. Lee, P. J. Kim, Y. Y. Ahn, and H. Jeong, "Googling social interactions: Web search engine based social network construction," 2007.

[47] A. E. Motter, "Improved network performance via antagonism: From synthetic rescues to multi-drug combinations," *BioEssays*, vol. 32, pp. 236–245, 2010.

[48] J. Reichardt, "Mesoscopic structure in complex networks," in *Proc. NetSci Conf.*, Boston, MA, May 2010.

[49] J. Crofts, "Discovering hierarchical and asymmetric network structure, with applications in neuroscience," in *Proc. NetSci Conf.*, Boston, MA, May 2010.

[50] D. Grady, C. Thiemann, and D. Brockmann, "The tomography of human mobility—What do shortest-path trees reveal?" presented at the *APS Meeting*, Mar. 2010.

[51] D. E. Whitney, "A dynamic model of cascades on random networks with a threshold rule," Arxiv preprint, arxiv: 0911.4499, 2009.

[52] L. Braunstein, "Effects of degree correlations to the pressure congestion on complex networks," in *Proc. NetSci Conf.*, Boston, MA, May 2010.

[53] R. Sinatra, "Networks of motifs from sequences of symbols," in *Proc. NetSci Conf.*, Boston, MA, May 2010.

[54] D. Krioukov, F. Papadopoulos, M. Boguna, and A. Vahdat, Arxiv preprint, arxiv: 0805.1266, 2008.

[55] M. Faloutaos, P. Faloutsos, and C. Faloutsos, "On power-law relationship of the Internet topology," *Comput. Commun. Rev.*, vol. 29, no. 4, pp. 251–262, 1999.

[56] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. Amer. Math. Soc.*, vol. 7, no. 1, pp. 48–50, 1956.

[57] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 409–410, 1998.

[58] M. Kochen and S. Pool, "Contacts and influences," *Social Netw.*, vol. 1, no. 1, pp. 5–51, 1978.