

DETC2016 – 60036

A DEGREE-BASED DECISION-CENTRIC MODEL FOR COMPLEX NETWORKED SYSTEMS

Zhenghui Sha
Department of Mechanical Engineering
Northwestern University
Evanston IL 60208, USA

Jitesh H. Panchal
School of Mechanical Engineering
Purdue University
West Lafayette IN 47907 USA

ABSTRACT

The system-level structure and performance of complex networked systems (e.g., the Internet) are emergent outcomes resulting from the interactions among individual entities (e.g., the autonomous systems in the Internet). Thus, the systems evolve in a bottom-up manner. In our previous studies, we have proposed a framework towards laying complex systems engineering on decision-centric foundations. In this paper, we apply that framework on modeling and analyzing the structure and performance of complex networked systems through the integration of random utility theory, continuum theory and percolation theory. Specifically, we propose a degree-based decision-centric (DBDC) network model based on random utility theory. We analyze the degree distribution and robustness of networks generated by the DBDC model using continuum theory and percolation theory, respectively. The results show that by controlling node-level preferences, the model is capable of generating a variety of network topologies. Further, the robustness of networks is observed to be highly sensitive to the nodes' preferences to degree. The proposed decision-centric approach has two advantages: 1) it provides a more general model for modeling networked systems by considering node-level preferences, and 2) the model can be extended by including non-structural attributes of nodes. With the proposed approach, systems that are evolved in a bottom-up manner can be modeled to verify hypothesized evolution mechanisms. This helps in understanding the underlying principles governing systems evolution, which is crucial to the development of design and engineering strategies for complex networked systems.

NOMENCLATURE

AS – autonomous system
ATS – Air Transportation System
BA – Barabasi-Albert
CCD – Complementary Cumulative Distribution
DBDC – Degree-Based Decision-Centric

DCA – Discrete Choice Analysis
FAA – Federal Aviation Administration
GPA – Generalized Preferential Attachment
LCC – Largest Connected Component
 d – Degree
 $\langle d \rangle$ – Average degree
 f_c – Critical fraction of nodes
 m – Number of edges added at each step
 \mathbf{s} – Decision-maker-specific variables vector
 t – Time
 \mathbf{x} – Alternative-specific variables vector
 N – Number of nodes in a network
 P – Probability
 U – Utility
 V – Observed utility
 $\boldsymbol{\beta}$ – Preferences parameter vector of utility function
 β_0 – Intercept of the utility function
 β_1 – Preference parameter of degree
 ϵ – Uncertainty
 Φ – Gaussian density

1. INTRODUCTION TO COMPLEX NETWORKED SYSTEMS

1.1. Unique features of complex networked systems

Research in complex systems has been increasingly focused on one class of systems in which the system components are highly connected as network and the system structures are not directly controlled by the designers, but evolve as a result of decisions of self-directed components/entities [1]. These systems are called complex networked systems. An example of such a system is the air transportation system (ATS), which contains different types of decision-makers (stakeholders) at different levels, such as passengers, airline companies and Federal Aviation Administration (FAA). Passengers decide which flight to take

depending on the fare, time of transit, in-plane service, etc. The passengers' decisions contribute to the demand between cities, which drives the airlines' decisions on whether to add/delete a route or no (i.e., route planning) and what type of aircraft to be served on a specific route (i.e., fleet planning). FAA at a higher level will make policy and regulations in response to decisions made by airlines to ensure the overall system runs safely and efficiently. Therefore, the stakeholders' decisions drive system trends. Another example is the Internet in which the autonomous systems (ASes) make decisions about which target ASes to link with in order to route data. These local decisions affect the global structure of the Internet, which in turn affects the Internet performance in terms of its robustness and resilience to node failure. Other examples include smart grid [2], customer-product systems [3] and smart vehicle networks [4].

The distinguishing feature of complex networked systems is that they evolve endogenously in a bottom-up manner as a result of decision-making interactions among individual entities. These individual entities (e.g., humans) are heterogeneous in nature and each of them has its own goals. Rational/irrational decisions are then made according to those goals. Moreover, individual entities may be at different levels in a complex system hierarchy, such as FAA vs. passengers in the context of ATS. Accordingly, interactions also exist between levels, resulting in recursive levels of integration. Therefore, the causal relationship between the node-level properties and the system-level performance is complex and hard to predict. These unique features make the engineering and design of complex networked systems a challenge task.

To address the challenge, a perspective in engineering that is both *decision-centric* and *bottom-up* is desired. Rather than directly controlling the system structure, the behaviors of the local interacting entities must be influenced in a way that leads to desired systems structure. Such influence requires improved coordination mechanisms and global policies, which can be realized through the provision of incentives, imposition of penalties or taxes and definition of rules or protocols. Especially, incentive structures should be carefully designed so that while individual entities are pursuing their private goals, desired system structures can be simultaneously achieved without compromising the system-level performance.

1.2. Initial efforts on complex system engineering and design

In our previous study [5, 6], we have proposed a framework towards laying complex systems engineering on decision-centric foundations. Such framework helps in characterizing the unique features of complex networked systems, as shown in Fig. 1. The local entities represented by nodes are self-directed and make decisions based on their preferences (Level 1). The preferences lead to different linking behaviors (Level 2), which affect the network structure (Level 3). Different network structures result in different network properties (Level 4) and the resulting network performance (Level 5). In the context of the ATS

example, airlines make decisions on route planning with the intent of maximizing their profitability. These decisions directly affect the structure of air transportation network, such as the degree distribution, which in turn impact the whole system's performance such as propagation of delays, robustness of network to service disruptions, and the network's traffic flow capacity.

The framework also helps in identifying the studies needed to be performed for facilitating the complex systems engineering and design. As shown in the Fig. 1, we summarize four aspects of studies with such framework, including 1) modeling the relationship between node-level preferences and behaviors; 2) analyzing the effect of node-level preferences on the system-level; 3) estimating the node-level preferences from system structures; and 4) design incentive structures to influence individuals for desired system performance. Among the four aspects, the modeling part is a critical step. This is because the estimation and analysis of node-level preferences rely on the behavioral model established, which in turn enables designers to predict the system dynamics, to test alternative design strategies and ultimately to explore possible incentive structures.

Our previous work [7, 8, 9] has integrated the Discrete Choice Analysis (DCA) based on the Random Utility Theory [10] with the network theory to model the nodes' linking behaviors in a complex system. In DCA, individuals' preferences are modeled as a utility function and they make decisions by maximizing the utility (see detailed discussion in Section 3.1). On the other hand, network model is very useful in modeling large-scale complex systems because of its domain-independent nature and appropriate abstract of systems architecture and inter relationships among components in a system. Therefore, with the integration of DCA into network theory, the proposed network model helps in providing an explanatory framework for the evolution of complex networked systems.

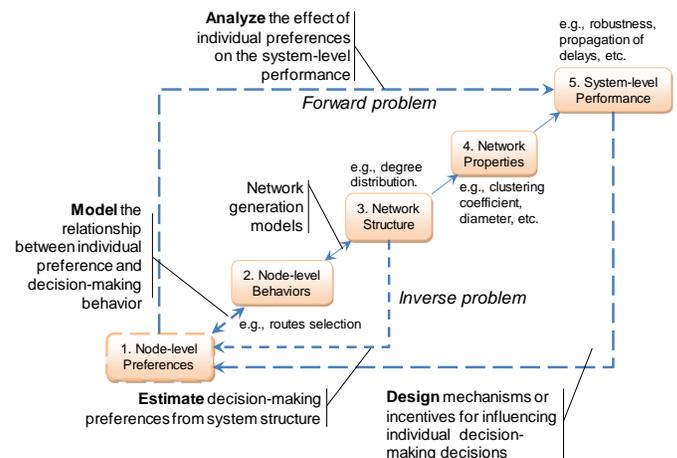


Fig. 1 Five-level decomposition and the associated mappings in complex networked systems

The focus of our previous work is on estimating node-level preferences in complex systems from network structural data. The estimation is an *inverse* problem, as shown in Fig. 2. However, our **research focus** in this paper is on analyzing the effect of node-level preferences on system structures and performance, which is a *forward* problem. To this end, we first propose a network generation model by taking the node-level preference into consideration. Then we integrate the continuum theory [1] and percolation theory [11] to analyze the impacts of nodes' preferences on network degree distribution and the network robustness. The research outcome is a direct mapping from node-level preferences to system-level performance, which enables the traction and prediction system performance.

The remaining of the paper is organized as follows. In Section 2, we present the literature review on existing studies of using network generation models for complex networked systems. In Section 3, DCA and the proposed degree-based decision-centric (DBDC) model are introduced. In Section 4 and 5, the theoretical and experimental analysis of the DBDC model are performed, and the impacts of node-level preferences on the network structures and robustness are evaluated. To highlight unique features of the DBDC model, a comparative study to the Generalized Preferential Attachment (GPA) network model is performed in Section 6. Finally, closing thoughts are discussed in Section 7.

2. LITERATURE REVIEW OF NETWORK GENERATION MODELS

We would like to first highlight that the study presented in this paper is for understanding how the system structure (represented as a network) emerges from the local interactions of individual entities (or agents, represented as nodes). There is literature, largely within the multi-agent systems, also analyzing such local interactions, but considering the learning effects. For example, Colby and Tumer [12] study the reinforcement learning mechanism and its resulting dynamics on the coordination of agents in a sensor network. In such study, the network size is typically small and the network structure is defined. The interest exists in understanding the coordination dynamics on a known network resulted from the learning of agents. However, the networks considered in this paper are typically large-scale, and may contain thousands of nodes. Therefore, the interest exists in first understanding how such a large-scale network comes to the structure we observe, and what are the effects of nodes' characteristics on the network generation, with a given local decision-making rules by assuming no learning effect among nodes. Being anchored on this perspective, in what follows, we review existing studies on network generation models.

Existing network generation models have been focused on developing networks that explicitly fit certain characteristics of complex networked systems, such as degree distribution. Typical examples of such models are the Watts-Strogatz model that

generates local clusters and triadic closure, Barabasi-Albert (BA) model that results in scale-free networks [13]. For modeling real-world networks, models for generating scale-free networks with tunable clustering coefficient are proposed by Herrera [14], Holme [15], and Klemm and Eguiluz [16]. The primary limitation of these models is that they mainly focus on matching the observed network structure and properties (Levels 3 and 4) using hypothesized behaviors of the nodes (Level 2). These models are not suitable for explaining why individual nodes behave in that way. Hence, these models do not capture the relationship between Levels 1 and 2 in Fig. 1.

While majority of the research is focused on modelling the network formation in which behaviors of nodes are independent of each other, the theory of network formation games is concerned with modeling networks that arise from strategic choices of nodes (players) on link formation [17, 18]. The node-level behaviors are generally defined in terms of the minimization of cost incurred in building a network, but can result in maximal system payoff. Models such as the local connection game [19] the global connection game [20], facility location game [21] and dynamic network formation game [22] provide insights on the basic principles of network formation. Although network formation games provide an explanatory framework for node-level decisions, the main limitation is that the effects of node-level preferences and behaviors on the network structure and performance have not been clearly understood.

Snijders [23] presents actor-oriented models to model the evolution of social networks as a result of independent actors making linking decisions to maximize their own utilities. The models result in the probability of link addition/deletion based on the multinomial logit model. The author uses various structural measures of the network, including degree, number of reciprocated relations, and centrality, to estimate the nodes' linking decisions. While the actor oriented models use a decision-making idea, the author does not analyze the effects of the node-level preferences on the network performance.

In systems engineering, design drivers such as cyber security and robustness should be built into the solution at the very beginning. Thus, we need to have a good understanding on the impact of local attributes and interaction on system-level performance in advance. However, existing network generation models for complex networked systems place too much emphasis on the outcomes of decision-making (node-level behavior) instead of the reasons for those decisions (decision-making preferences). There is a lack of theoretical foundations that explain how their proposed nodes' linking behavioral models are established. In other words, the relationship between node-level preference (Level 1) and the resulting behaviors (Level 2) is unexplored. So existing studies are limited in establishing relationships between the nodes' preferences and the system performance. These limitation impedes our understanding on the evolution of complex networked systems, and ways to direct the evolution of

these systems through incentive schemes. Therefore, this paper aims 1) to establish a decision-centric model for complex networked systems by coupling the node-level preferences and behaviors (see Section 3), and 2) to analyze the effect of node-level preferences on network structure and performance (see Section 4 and 5).

3. A DECISION-CENTRIC MODEL FOR COMPLEX NETWORKED SYSTEM

3.1. Theoretical foundation

DCA [24] has been widely used to model and forecast product demand by modeling individuals' choice behavior in many fields ranging from transportation systems [25] to engineering systems design [26]. DCA is conducted based on the assumption that individuals seek to maximize their own utility while making decisions. Mathematically, the utility of decision maker n acquired by choosing alternative i , denoted as U_{ni} , can be represented as:

$$U_{ni} = V_{ni} + \epsilon_{ni} \quad (1)$$

where V_{ni} is the observed utility and ϵ_{ni} describes the uncertainty of the researchers in understanding decision makers' utility. Here, V_{ni} can be modeled as a function of x_{ni} and s_n , given parameter vector $\beta_n = (\beta_{n1}, \dots, \beta_{nk})'$, where k is the total number of variables identified. So $V_{ni} = V(x_{ni}, s_n | \beta_n)$, where x_{ni} is alternative-specific variables that affect decision makers' utility, such as choices' attributes; and s_n is decision-maker-specific variables, such as decision-makers' age, income, etc. Following utility-maximization, the probability of decision maker n chooses alternative i is the probability that the utility acquired by choosing i is greater than any other choices,

$$P_{ni} = P(U_{ni} > U_{nj}) = P(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj}), \forall i \neq j \quad (2)$$

With different distributions of the unobserved utility ϵ_{ni} , different types of discrete choice models (DCMs) can be obtained. In this paper, a multinomial logit model [27] is adopted under the assumption that ϵ_{ni} is identical independent distributed type I extreme value [24]. The reason of using this model is that a) closed form of the resulting choice probability can be obtained and b) the existing evidences have showed that logit model has good performance on predicting the decision-making preferences in different fields [8, 23-26]. If the preferences to the variables x_{ni} and s_n are homogeneous for all the decision-makers, then $\beta_n = \beta$. With Equation (2), the logit form of the choice probability is:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^N e^{V_{nj}}} = \frac{e^{V(x_{ni}, s_n | \beta)}}{\sum_{j=1}^N e^{V(x_{nj}, s_n | \beta)}} \quad (3)$$

3.2. Proposed model for network evolution

Following the framework of DCM in [25], a node i within a network is viewed as a decision maker, and the node's

preferences are modeled in a utility function. The node i makes decisions about which target node j to connect to (or disconnect from). The explanatory variables for those decisions can be categorized based on whether they are specific to alternatives or not, and whether they are related to network structure or not. Examples of network structure attributes include degree (the number of connectivity), clustering coefficient, and centrality. Non-network-centric attributes are social, economic, ecological and/or political factors such as income, cost, gender, etc. In this paper, using the DCM in Equation (3), we develop a degree-based network generation model as follows:

- *Growth mechanism:* During each time step, one new node enters the network, and will connect with m existing nodes with one at a time.
- *Utility function:* A commonly used linear utility function is adopted. In addition, only one alternative-specific network attribute, target node's degree (d), is adopted as the explanatory variable, i.e., $x_{1i} = d_j$.

$$V_i = \beta_1 d_i + \beta_0 \quad (4)$$

- *Choice probability:* The decision-making nodes are assumed to have homogeneous preferences to the degree. Hence, the resulting choice probability that the i^{th} node is selected is

$$P_i = \frac{e^{\beta_1 d_i + \beta_0}}{\sum_{j=1}^N e^{\beta_1 d_j + \beta_0}} = \frac{e^{\beta_1 d_i}}{\sum_{j=1}^N e^{\beta_1 d_j}} \quad (5)$$

Equation (5) defines a simple network formation mechanism which indicates that the probability linking a node only depends on the alternative nodes' degree and the decision-making nodes' preference to such degree. The degree is chosen as the variable for two reasons: a) In network science, a nodes' degree has been regarded as one of the most important factors affecting the network topology [28]. b) By choosing degree, it enables a direct comparison with the existing network generating models, such as the BA model. Other assumptions in the model are:

- 1) The network is undirected, and only one edge exists between a pair of nodes.
- 2) The network is initialized with a random network with 5 nodes, and the probability of each link is 0.5.
- 3) The parameter β_1 is time independent.

The model is called degree-based decision-centric (DBDC) model in which β_1 captures the preferences of the nodes to degree. For example, in a friendship network, β_1 simply implies how a person is sensitive to the number of friends of another person. It reveals a person's intrinsic characteristic of whether prefer to playing with people who have a lot of friends or people with a few friends. The resulting choice probability is derived from the utility-maximization behavior based on such preference. The random utility theory provides an explanatory framework for establishing the relationship between node-level

preferences and behaviors¹. In the following section, theoretical analysis of structure of the networks generated with the DBDC model is performed.

4. INTEGRATION OF THEORIES TO ANALYZE THE DBDC MODEL

4.1. Solution of degree distribution

Degree is an effective indicator of whether a node in a network is critical or not. Therefore, degree distribution is the one of the most effective ways to quantitatively represent network structure. In the context of real system, such as the ATS, the degree of nodes represents the number routes connected to an airport. Higher degree indicates an airport is more likely to be a hub. Understanding how the number of routes of each airport is distributed in the ATS is crucial for predicting travel demand in the future, which is important for airlines to make decision on route and fleet planning.

To analyze the degree distribution of networks generated by DBDC model, we employ the continuum theory presented by Albert and Barabasi [1]. The degree, d_i , of node i is a continuous, time-dependent variable. It increases every time a new edge is added to the network. According to the growth mechanism, m edges are added at each time step. The rate at which d_i changes is proportional to P_i in Equation (5). Consequently, d_i satisfies the differential equation:

$$\frac{\partial d_i}{\partial t} = mP_i = m \frac{e^{\beta_1 d_i}}{\sum_{j=1}^N e^{\beta_1 d_j}} \quad (6)$$

Algorithm: Numerically solving Equation (6)

Results: $d_i[t]$: Node i 's degree as a function of time t

Initialization: set the initial condition for the degree of node 1, $d_{init1}[t_1] = m$

for $i = 1$ to N **do**

Construct i differential equations using Equation (6);
 Define the time span for solving the differential questions, $Timespan = 1$;
 Solve the system of i differential equations and get solution for each node i 's degree $d_i[t]$;

for $j = 1$ to $N - 1$ **do**

Update the value of $d_{initj}[t_{i+1}]$ based on the solution at time t_{i+1} ;

end

Set the initial condition for the incoming node, $d_{initj}[t_{i+1}] = m$;

end

This differential equation is hard to solve analytically because of the exponential term. Hence, we solve numerically using the

algorithm shown above. Note that at each step, a new node enters the network. At time t_i , it is a system of i differential equations, where $i = 1, \dots, N$. The initial condition for i^{th} differential equation at time t_i is $d(t_i) = m$. The initial conditions for the remaining $(i - 1)$ differential equations at time t_i are determined by solving the differential equations at t_{i-1} .

With the proposed algorithm, $d_i(t)$ is obtained for each node i . At a certain time t_i , each node's degree is determined and the resulting degree distribution is evaluated by discretizing the continuous results. This approach is used to evaluate the degree distributions for different β_1 values. Fig. 2 shows how β_1 influences the networks' complementary cumulative distribution (CCD) of degree for $m = 3$ and $N = 2000$. Based on the results, it is observed that the degree distributions are qualitatively different for ranges: (a) $\beta_1 < -1$, (b) $-1 \leq \beta_1 \leq 0.04$, and (c) $\beta_1 > 0.04$, as plotted separately. The cutoff values for these ranges have been chosen through observation.

For $\beta_1 < -1$ and $\beta_1 > 0.04$, the degree distributions are found to be qualitatively similar. However, the degree distributions are highly sensitive to β_1 within $-1 \leq \beta_1 \leq 0.04$. By fitting different distributions to the data, it is observed that the best fitting functions are: a) exponential for $\beta_1 = 0$: $P(d) = 2.5097e^{-0.325d}$ with $R^2 = 0.999$; b) power law for $\beta_1 = 0.04$: $P(d) = 62.919d^{-2.956}$ with $R^2 = 0.981$; and c) truncated Gaussian [29] for $\beta_1 = -1$: $P(d) = \Phi(5.5, 1.11)$ truncated between [3, 8] with $R^2 = 0.929$, where $\Phi(\cdot)$ is the density of Gaussian distribution. So, as β_1 increases from -1 to 0.04, the resulting network transitions between different types of topologies as illustrated in Fig. 2(d) using probability density. The emergence of different types of degree distributions can be explained using the choice probability in Equation (5). First, when $\beta_1 = 0$, the choice probability P_i is $\frac{1}{N}$ indicating that all the nodes have the same linking probability, resulting in a random network with exponential degree distribution. This is also verified by Equation (6): with $P_i = 1/N$, as $t \rightarrow \infty$, the distribution decays exponentially following $p(d) = \frac{e}{m} e^{-\frac{d}{m}}$.

Second, when β_1 is positive, the higher degree nodes are more likely to get a connection. This process results in scale-free networks with power-law degree distribution. As β_1 increases beyond 0.04, giant hubs emerge that occupy most of the edges in a network. This is reflected by the degree distribution shown in Fig. 2(c). After β_1 increases to a certain value (the solution indicates that for $\beta_1 \geq 1.0$, the network structure remains unchanged even if β_1 is increased further. Referring to Equation (5) for $\beta_1 \geq 1.0$, the choice probability of the first node is close to 1, and the other nodes' choice probabilities are

¹ Note that the proposed model only applies to one organization-level network, e.g., airlines make decision on whether to connect an airport to another airport. To address the hierarchical relationship in a complex system with

multiple organizational levels, e.g., the FAA vs. the airlines, we need to create the decision-making models of stakeholders at each level.

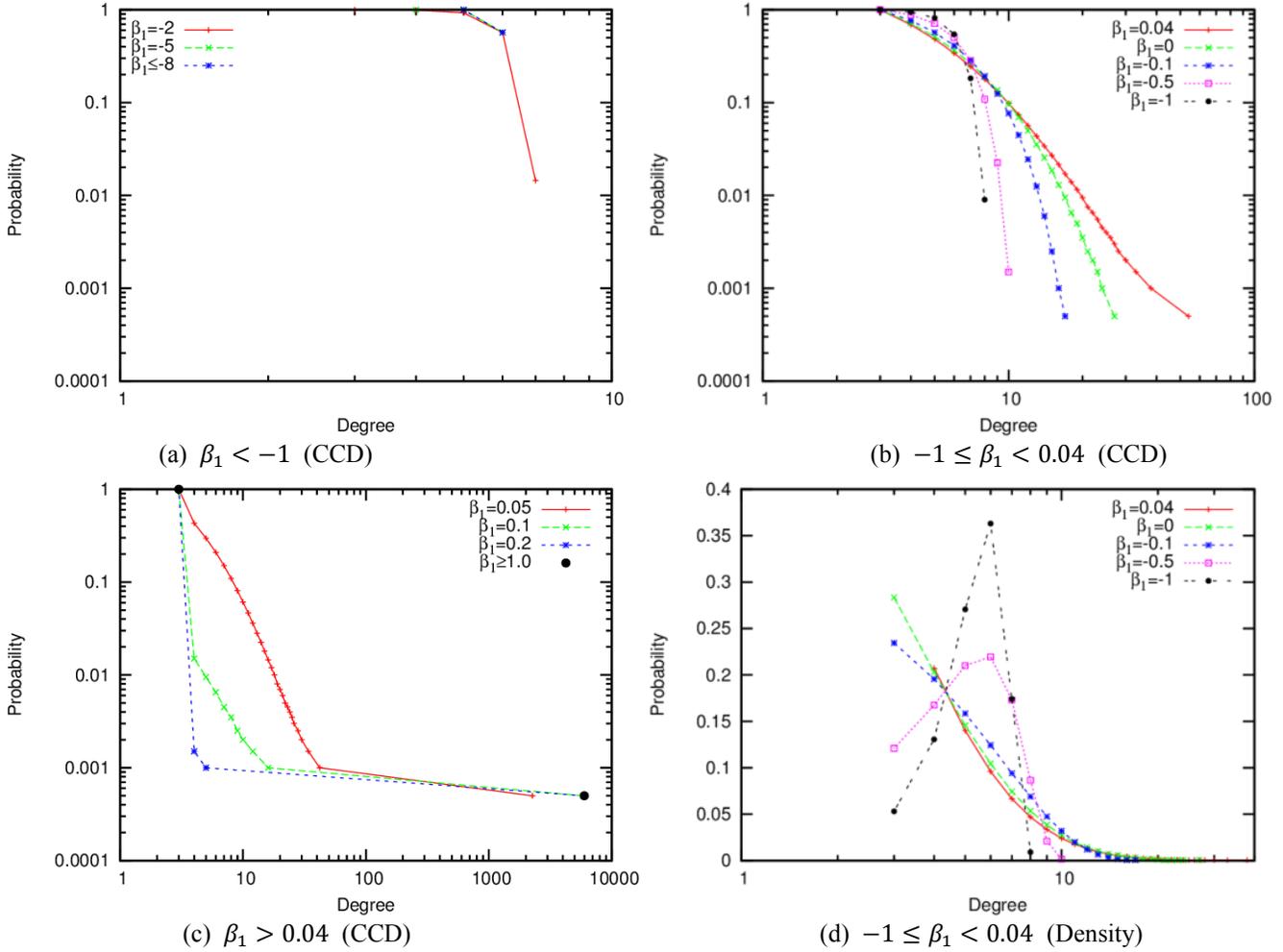


Fig. 2 Theoretical results of degree distributions at different β_1 values

close to 0. Thus, only the first node's degree increases at the rate of m , as indicated by Equation (6).

Third, for $-1 \leq \beta_1 < 0$, nodes with lower degree have greater probability to be linked. At each step, there is always a new node with lower degree entering the network. Hence, most nodes have a degree close to the average degree, $\langle d \rangle$, leading to a small-world network [30] with Gaussian distribution (see Fig. 2(d)), which has a pronounced peak at $\langle d \rangle = 6$, and decays exponentially for large d . Finally, after a certain value (the solution indicates that for $\beta_1 \leq -8$), the degree distribution indicates that all the nodes have the same degree around average degree $\langle d \rangle$. For example, the degree distributions at $\beta_1 = -8$ show that the network only has nodes with degrees 5 and 6 (see Fig. 2(a)). This is because for $\beta_1 = -8$, the resulting choice probabilities are almost the same for each node. So the rate at which each node's degree changes is almost the same as indicated by Equation (6). This corresponds to following network formation process: the node with lowest degree has the largest probability to be connected. The newly added node is more likely to choose the node added in the previous step

because that node has the lowest degree. So, a network with "chain" structure emerges.

To verify the solutions from continuum theory, synthetic networks are generated using different values of β_1 and the corresponding degree distributions are compared with predictions from the theoretical results in Section 4.1 In the following section, we present theoretical analysis of the network robustness.

4.2. Network robustness against random failure and targeted attack

The system-level performance considered in this paper is the network's robustness. Robustness means the persistence of a system's characteristic behavior under perturbations [31]. For example, in the ATS, the robustness implies whether the system is still functional (e.g., transporting passengers) when some of the airports cannot perform their intentional functionalities (e.g., because of congestion of flights). In network science, the perturbations generally considered are either from random failure of nodes or targeted attack on nodes. In random failure,

nodes are randomly selected and removed from the network. For example, this corresponds to situation in the ATS where airports get affected by severe weather conditions. On the other hand, targeted attack involves elimination of nodes with certain properties. We consider targeted attacks on nodes with the highest degree, which can be used to represent a hub airport attacked by terrorist or its service is brought down by hackers. In both scenarios, after a certain fraction of nodes, say f_c , is attacked, the network becomes fragmented. The robustness of the network is proportional to the fraction of nodes that need to be removed before fragmentation occurs. So higher the f_c , higher the robustness. To measure when the network is fragmented, the size of largest connected component (LCC) is adopted. The presence of a LCC is an indicator of a network that is at least partly performing its intended function, while the size of the LCC indicates how much of the network is working. Theoretical analysis of network robustness is based on percolation theory [11]. Details about how percolation transition can be solved using degree distribution can be referred in [32].

for $\beta_1 = -1$, $f_c = 0.82$ for $\beta_1 = 0$ and $f_c = 0.74$ for $\beta_1 = 0.04$. Compared with the robustness of other networks reported in [32], the obtained results indicate that the overall network robustness stays high as β_1 changes. Fig. 3(b) shows the percolation transition of targeted attack on three representative networks. The critical fraction points for the three networks are calculated as $f_c = 0.969$ for $\beta_1 = -1$, $f_c = 0.398$ for $\beta_1 = 0$, and $f_c = 0.0028$ for $\beta_1 = 0.04$. The robustness against targeted attack of Gaussian network ($\beta_1 = -1$) is higher than the exponential network ($\beta_1 = 0$), and the robustness of the exponential network is higher than the network with power-law degree distribution ($\beta_1 = 0.04$). This indicates that as β_1 increases from negative to positive, the network robustness decreases gradually.

In summary, we employ the continuum theory and percolation theory in this section to analyze the structure and robustness of networks generated with the DBDC model. These theories enable us to establish the relationship between node-level preferences, the network structure, and the robustness. The results show significant influence of the node-level preferences on network structure and robustness. In the following section, the conclusions are verified by using experimental analysis on synthetic networks generated using the DBDC model.

5. SIMULATION STUDY

5.1. Simulation results of network structures

To illustrate the effects of β_1 on the network structure, sample networks generated with 200 nodes and $m = 1$ are shown in Fig. 4. The results visually verify our theoretical conclusions that as β_1 increases from negative to positive, the network structure transitions from “chain” to exponential random, then to scale-free, and finally to a giant-hub.

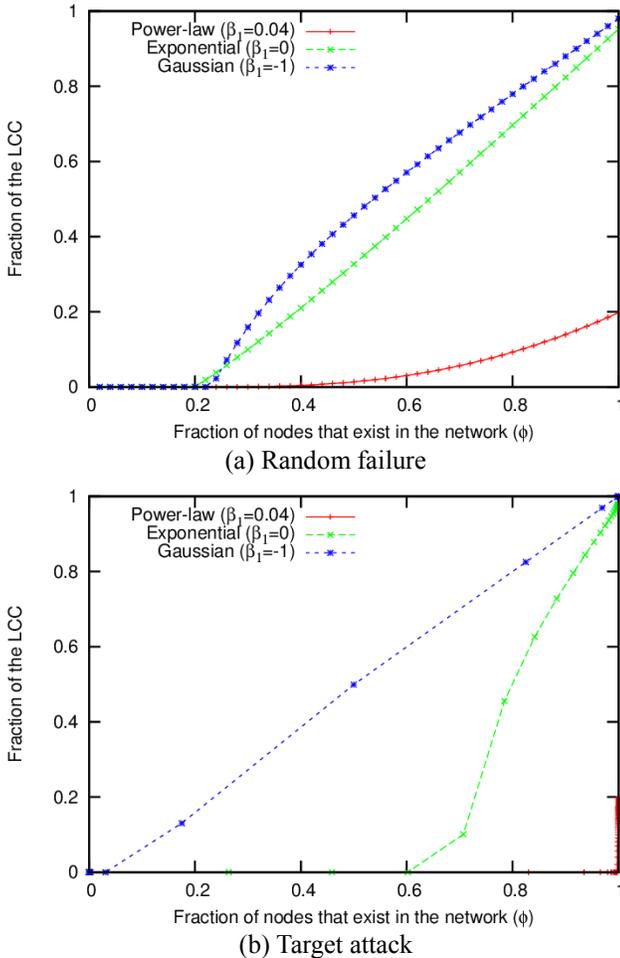


Fig. 3 Percolation transition of the three critical networks

Fig. 3(a) shows numerical results of percolation transition of the three critical networks for random failure of nodes, $f_c = 0.78$

Table 1 Best fitting of degree distribution of synthetic networks generated with the DBDC model

Critical β_1	Fitting Function for CCD ($F(d) = p(d_i > d)$)	R^2	Network structure
$\beta_1 = -20$	$4.5679e^{-0.397d}$	0.996	Exponential random
$\beta_1 = -2$	$\Phi(5.99, 0.871)$ truncated between [3,8]	0.982	Small-world network
$\beta_1 = -1$	$\Phi(5.988, 1.296)$ truncated between [3,10]	0.959	Transition from exponential random to small-world
$\beta_1 = 0$	$2.6233e^{-0.298d}$	0.996	Exponential random
$\beta_1 = 0.04$	$1.08e^{-0.202d}$	0.988	Transition from scale-free to exponential random
$\beta_1 = 0.05$	$46.444d^{-2.673}$	0.976	Scale-free
$\beta_1 = 15$	$0.8544e^{-0.212d}$	0.85	Exponential random

We further investigate the effects of β_1 on degree distributions empirically for the same model settings (network size $N =$

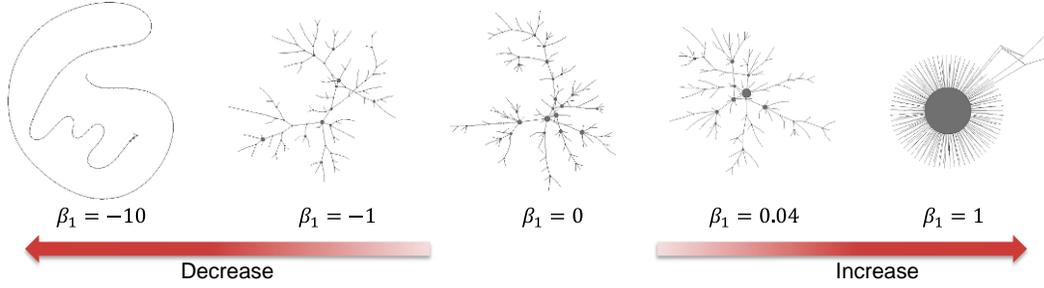


Fig. 4 Network visualization for representative β_1 values (Model parameters: network size = 200 and $m = 1$)

2000 and $m = 3$) as in Section 3.1. The degree distributions of the critical values of β_1 identified through theoretical analysis (i.e., $\beta_1 = -1, 0$ and 0.04) is analysed and the best fitting curve is presented in Table 1. Based on the results, it is observed that the overall trends of the network structural evolution comply with the ones presented in Fig. 2. But the critical values of β_1 at which the networks transition are slightly different. For example, the theoretical solution indicates that $\beta_1 = 0.04$ corresponds to the network with power-law degree distribution.

However, in the generated network, the corresponding β_1 is found to be 0.05. This can be attributed to the idealizations in the theoretical model, including the assumption of degree as a continuous variable and the numerical solution of the rate equation. So far, the empirical results verify the conclusion drawn in Section 3.1.

5.2. Simulation results of network robustness

To simulate the network fragmentation process, 1% of the existing nodes are removed during each step. In random failure, the nodes to be removed are randomly selected, whereas in targeted attack, the nodes with the highest degrees are selected. Changes in LCC are monitored during the node-removal process. The network is said to be fragmented when the LCC reduces below 2% of the LCC of the original network. The network's robustness is quantified in terms of the critical fraction point, f_c , corresponding to network fragmentation. Following the approach of Beygelzimer et al. [33], ten networks are generated at each β_1 value, and the random failure process is carried out ten times for each network. 95% confidence intervals are calculated using the *t-distribution*.

Empirical results of network robustness against random failure and targeted attack are shown in Fig. 5. It is observed that the robustness against random failure is higher than the robustness against targeted attack. This verifies the conclusion from theoretical analysis presented in Section 3.2. However, the network robustness is highly sensitive to β_1 in the range of $0.05 \leq \beta_1 \leq 2$, as highlighted in the small figure in Fig. 5. The average robustness ranges from 0.658 to 0.859. This is because β_1 in this range corresponds to the network with many hubs. Since these hubs occupy most of the edges in the network, the robustness is highly dependent on when these hubs are

removed. Since nodes are randomly removed, the variation in network robustness is large.

For the targeted attack, the network robustness increases logarithmically (The fitting function is $y = 0.0311 \ln(x) + 0.4908, R^2 = 0.9803$) as β_1 decreases from 0 to -5 , which corresponds to the transition from exponential random to small-world network. The robustness reaches a maximum value of 0.533 at $\beta_1 = -5$. At $\beta_1 = 0$, the robustness is 0.382, which corresponds to the robustness of an exponential random network. The network robustness decreases for $0 \leq \beta_1 \leq 0.32$ and reaches its minimum value of 0.02 at $\beta_1 = 0.32$, where the largest giant-hub network structure is formed. There is a sharp increase in robustness from 0.02 to 0.373 for $0.32 \leq \beta_1 \leq 2$.

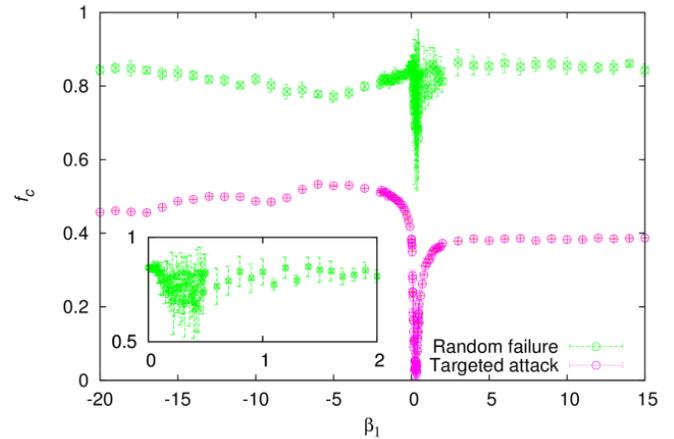


Fig. 5 Empirical results of network robustness against random failure (the sub-figure inside is the zoom-in of the network robustness against random failure in the range of $\beta_1 \in [0, 2]$)

In summary, the network robustness decreases monotonically as network transitions from small-world network to exponential random network, and further to a giant-hub structure. This result verifies the outcomes from theoretical analysis in Section 3.2. Additionally, for $\beta_1 \geq 2$, the robustness is almost constant, with an average value of 0.383. For $\beta_1 \leq -5$, the network robustness gradually decreases to 0.457 at $\beta_1 = -20$. Both values are close to the robustness of random network, indicating that the network

structure returns to the exponential random network as $|\beta_1| \rightarrow \infty$. This provides a complete picture of how the node-level preferences β_1 affect the network robustness.

The results on the network robustness shed light on making the system more robust, assuming the model represent the growth mechanisms underlying real systems. As a guiding principle, β_1 values resulting in giant-hub networks should be avoided due to low robustness in targeted attack and fluctuation of robustness in random failure. There are three critical values of β_1 for acquiring a network with high robustness: $\beta_1 = \{0, 2, -2\}$. For $\beta_1 = 0$, an exponential random network is obtained with high robustness against targeted attack. The second critical value is at $\beta_1 = 2$ where the resulting network has a robustness value of exponential random network and meanwhile, the network robustness does not increase but is maintained after this value. The third critical value is in the transition of network structure from exponential random to small-world as β_1 decreases. For $\beta_1 \leq -2$, the robustness does not increase significantly. This implies that the incentives should be well designed to influence the individuals' preferences so that a network with high robustness to targeted attack can be achieved without costing additional effort.

6. COMPARATIVE STUDY AND FURTHER INSIGHTS

The approach that integrate continuum theory and percolations theory to analyze the network models is general enough to be applied to many other network generation models. We applied this approach to the GPA model that has been widely used in network science community. Detailed analysis on the GAP model can be referred to [6]. Through the comparison, we found the GPA model is limited in modeling the diverse structures of systems in the real world. However, as the results indicate, the proposed DBDC model is simple, yet shows complex transitions of network structures and performance. The richer model provides the capability of capturing diverse types of complex networked systems, as compared to the GPA model. Besides, there are several other differences between the DBDC model and GPA model. The first main difference between the DBDC and GPA model is in the linking choice probabilities. The choice probability of DBDC model in Equation (5) can be written as:

$$P_i^{DBDC} = \frac{e^{\beta_1 d_i}}{\sum_{j=1}^N e^{\beta_1 d_j}} = \left(1 + \sum_{j=1, j \neq i}^N e^{\beta_1 (d_i - d_j)} \right)^{-1}$$

Similarly, the linking probability of GPA model is:

$$P_i^{GPA} = \frac{d_i + A}{\sum_{j=1}^N d_j + A} = \left(1 + \sum_{j=1, j \neq i}^N \frac{1 + \frac{A}{d_i}}{\frac{d_i}{d_j} + \frac{A}{d_i}} \right)^{-1}$$

where A is called additional attractiveness which describes the additional influence on nodes' linking probability besides degree. Therefore, in the DBDC model, the probability of linking is dependent on the *difference* between their degrees. On the

other hand, in the GPA model, the linking probability is dependent on the *ratio* of the degrees of nodes. Hence, the two models provide different explanations for how the systems are driven to evolve.

In addition, the DBDC model and GPA models provide different interpretations of how parameters can be influenced. The models suggest two different ways to affect node-level behaviors. The first approach is to provide incentives to influence individuals' preferences, e.g., the preference to degree, β_1 . The second approach is to give incentives to have individuals change their characteristics, e.g., the node's additional attractiveness, A . Further, the DBDC model is more general than GPA. Actually, the GPA model can be modeled as a special case of the DBDC model. The choice probability of the DBDC model transforms to the GPA model if the utility function defined in Equation (4) is defined as $V_j = \ln(\beta_1 d_j + \beta_0)$. Finally, the DBDC model can be extended by including other network-metric and non-network-metric variables, linear or nonlinear utility functions, homogeneous or heterogeneous preferences. Thus it provides a general and extensible framework for network generation.

7. CLOSING THOUGHTS – TOWARDS SCIENTIFIC FOUNDATIONS FOR COMPLEX SYSTEMS ENGINEERING

As described in Section 1, in complex networked systems, individual entities interact with each other so that the systems evolve in a bottom-up manner. The interconnectivity and interdependency among the individual entities (either physical components/subsystems or human such as users, designers, and other stakeholders) lead to systems with complexity and vulnerabilities. Further, the increasing involvement of human activities in the loop makes systems more difficult to model and analyze. Therefore, engineering complex systems requires comprehensive systems approach to analyze not only the traditional technical issues, but also the policy issues and the human behaviors.

In this regard, it is important to have a scientific foundation that facilitates the bottom-up systems engineering. The lack of theoretic foundations in current systems engineering impedes the effective adaptation of systems engineering practices to ever increasing system complexity, to a broad range of application domains, and to new enabling technologies [34]. The systems engineering community has been aware of this trend. For example, a recent National Science Foundation (NSF) workshop brought experts in systems engineering to address the task of laying a theoretical foundation for the practice of systems engineering. The workshop pointed out that the systems engineering discipline is so broad that there should not be just one universal theory. Participants in this workshop listed over 30 theories and knowledge areas, ranging from mathematics and physical sciences to human and social sciences, which the foundation can be built on.

Towards developing such theoretical foundation, many initiatives have been taken. Davendralingam et al. [5] categorize these initiatives into four classes: a) the *application* of theories to solve specific problems in one domain; b) *abstraction* of theories from application domains to domain-independent knowledge; c) the *integration* of theories from different domain; and d) the *translation* of a theory from one domain to another domain. In these four classes of initiatives, current systems engineering has been primarily focused on the first two initiatives, i.e., the *application* and the *abstraction*. Though they are important, Davendralingam and coauthors [5] argue that the key to laying theoretical foundations for systems engineering depends on the other two classes of initiatives, i.e., the *integration* of theories and the *translation* of theories. These two classes are important because effective and cohesive integration helps determine the appropriate level of abstraction such that foundational research can be performed while ensuring the applicability of theories across domains. Consequently, translation can be achieved which enables useful systems engineering approaches to expand into new domains for lasting impact.

Achieving meaningful integration is whereas challenging primarily due to the breadth and diversity of many knowledge areas. As Collopy and Mesmer mentioned in [34] that “*a foundation is more than recognition of related knowledge areas. A foundation should identify specific elements within each knowledge area that are necessary, and then tie the knowledge areas together in a theoretical construct or model*”.

In this paper, we take initial steps towards achieving a meaningful integration. We show an implementation of the decision-centric framework to demonstrate how it facilitates the integration of theories from different domains, such as the random utility theory, continuum theory and the percolation theory, to support the modeling and analysis of complex networked systems. Specifically, a DBDC model is proposed for modeling the evolution of complex networked systems. The network structures, and the robustness against random failure and targeted attack are analyzed. The results establish a connection from the node-level preferences to the system-level performance. Therefore, a holistic framework that quantitatively captures the mappings across all the five levels (shown in Fig. 1) of complex networked systems is provided. The model is domain-independent and at an appropriate level of abstraction of system architecture.

The contributions of this paper are three-fold. First, the proposed DBDC model is domain-independent. It helps obtain diverse network topologies which represent various types of systems’ structures. This helps in obtaining surrogate models for different types of real-world systems. Second, the model enables node-level preferences estimation in terms of the β values. Thereby, the node-behaviors in complex systems can be better understood. Third, once the node-level preferences are obtained, current

system’s performance can be evaluated. Appropriate incentives could be designed to modify individuals’ preferences so that a desired system performance can be achieved.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support from the NSF through the grant 1360361.

REFERENCES

1. R. Albert and A. L. Barabasi, “Statistical mechanics of complex networks,” *Review of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.
2. B. D. Hawthorne and J. H. Panchal, “Policy design for sustainable energy systems considering multiple objectives and incomplete preferences,” in *2012 ASME International Design Engineering and Technical Conferences*, Chicago, IL.
3. Wang, M., Chen, W., Huang, Y., Contractor, N., and Fu, Y., “A Multidimensional Network Approach for Modeling Customer-product Relations in Engineering Design”, in *ASME 2015 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Boston, MA, 2015.
4. Varaiya, P., “Smart cars on smart roads: problems of control,” in *Automatic Control, IEEE Transactions on*, vol.38, no.2, pp.195-207, 1993.
5. N. Davendralingam, Z. Sha, K. Moolchandani, A. Maheshwari, J. H. Panchal, and D. DeLaurentis, “Scientific Foundations for Systems Engineering – Challenges and Strategies”, in *ASME 2015 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Boston, MA, 2015.
6. Z. Sha and J. Panchal, “Towards the design of complex evolving networks with high robustness and resilience,” in *Procedia Computer Science*, vol. 16, 2013, pp. 522–531.
7. Z. Sha, K. Moolchandani, A. Maheshwari, J. Thekinen, J. H. Panchal, and D. DeLaurentis, “Modeling Airline Decisions on Route Planning Using Discrete Choice Models”, in *15th AIAA Aviation Technology, Integration, and Operations Conference*, Dallas, TX, 2015.
8. Z. Sha, and J. H. Panchal, “Estimating Linking Preferences and Behaviors of Autonomous Systems in the Internet using Discrete Choice Models”, in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC) Conference*. San Diego, CA.
9. Z. Sha, J. Panchal, “Estimating the node-level behaviors in complex evolutionary systems,” *ASME Journal of Mechanical Design*, vol. 136 (6), 061003(11), 2014.
10. E. Cascetta, “Random utility theory,” in *Transportation Systems Analysis, ser. Springer Optimization and Its Applications*. Springer, 2009, vol. 29, pp. 89–167.
11. D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. CRC Press, July 1994.
12. M. Colby and K. Tumer, “Learning-based Coordination of Large Heterogeneous Distributed Sensor Networks”, in *Proceedings of the 2015 International Society of Automation Power Industry Division Symposium*. Kansas City, MO.
13. A. L. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

14. C. Herrera and P. J. Zufiria, "Generating scale-free networks with adjustable clustering coefficient via random walks," in *2011 IEEE Network Science Workshop (NSW)*, 2011, pp. 167–172.
15. P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Physical Review E*, vol. 65, no. 2, p. 026107, 2002.
16. K. Klemm, V. M. Eguiluz, "Growing scale-free networks with small-world behavior," *Physical Review E*, vol. 65, no. 5, p. 057102, 2002.
17. M. O. Jackson, "A survey of models of network formation: Stability and efficiency," California Institute of Technology, Division of the Humanities and Social Sciences, Working Papers 1161, 2003.
18. E. Tardos and T. Wexler, "Network formation games and the potential function method," in *Algorithmic Game Theory*, pp. 487–516, 2007.
19. A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker, "On a network creation game," in *Proceedings of the twenty second annual symposium on Principles of distributed computing, ser. PODC '03*. New York, NY, USA: ACM, 2003, pp. 347–351.
20. E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden, "The price of stability for network design with fair cost allocation," in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, ser. FOCS '04*. Washington, DC, 2004, pp. 295–304.
21. A. Vetta, "Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions," in *Proceedings of the 43rd Symposium on Foundations of Computer Science, ser. FOCS '02*. Washington, DC, 2002.
22. J. Derks, J. Kuipers, M. Tennekes, and F. Thuijsman, "Local dynamics in network formation," Maastricht University, Department of Knowledge Engineering, Tech. Rep., 2008.
23. T. A. B. Snijders, "Stochastic actor oriented models for network change," *The journal of Mathematical Sociology*, vol. 21(1-2), pp. 149–172, 1996.
24. M. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, 1985.
25. H. C. W. L. Williams, "On the formation of travel demand models and economic evaluation measures of user benefit," *Environment and Planning*, vol. 9, no. 3, pp. 285–344, 1977.
26. W. Chen, C. Hoyle, and H. J. Wassenaar, *Decision-Based Design: Integrating Consumer Preferences into Engineering Design*. Springer, 2013.
27. K. Train, *Discrete Choice Methods with Simulation*, 2nd ed. New York, NY: Cambridge University Press, 2009.
28. R. Albert and A. L. Barabasi, "Topology of evolving networks: Local events and universality," *Physical Review Letters*, vol. 85, no. 24, pp. 5234–5237, 2000.
29. T. Lindeberg, "Scale-space for discrete signals," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 234–254, 1990.
30. D. J. Watts, S. H. Strogatz, "Collective dynamics of 'small-world' networks", *Nature* 393 (6684): 440–442. 1998.
31. J. E. Alex K. S. Ng, "Structural robustness of complex networks," in *NetSci 2006*, 2006.
32. M. Newman, *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc., 2010.
33. A. Beygelzimer, G. Grinstein, R. Linsker, and I. Rish, "Improving network robustness by edge modification," *Physica A: Statistical Mechanics and its Applications*, vol. 357, no. 3–4, pp. 593–612, 2005.
34. P. D., Collopy, B. L. Mesmer, "Report on the Science of Systems Engineering Workshop", in *53rd AIAA Aeronautics Science Meeting*, Kissimmee, FL, 2015.