Quantitative Characterization of Complex Systems—An Information Theoretic Approach

Aditya Akundi
The University of Texas Rio Grande Valley

Eric Smith

Article

Quantitative Characterization of Complex Systems—An Information Theoretic Approach

Aditya Akundi [1,*] and **Eric Smith** [2]

1. Complex Engineering Systems Laboratory, Department of Informatics and Engineering Systems, The University of Texas Rio Grande Valley, Edinburg, TX 78539, USA
2. Department of Industrial, Manufacturing, and Systems, The University of Texas at El Paso, El Paso, TX 79968, USA; esmith2@utep.edu
* Correspondence: satya.akundi@utrgv.edu

Abstract: A significant increase in System-of-Systems (SoS) is currently observed in the social and technical domains. As a result of the increasing number of constituent system components, Systems of Systems are becoming larger and more complex. Recent research efforts have highlighted the importance of identifying innovative statistical and theoretical approaches for analyzing complex systems to better understand how they work. This paper portrays the use of an agnostic two-stage examination structure for complex systems aimed towards developing an information theory-based approach to analyze complex technical and socio-technical systems. Towards the goal of characterizing system complexity with information entropy, work was carried out in exploring the potential application of entropy to a simulated case study to illustrate its applicability and to establish the use of information theory within the broad horizon of complex systems. Although previous efforts have been made to use entropy for understanding complexity, this paper provides a basic foundation for identifying a framework to characterize complexity, in order to analyze and assess complex systems in different operational domains.

1. Introduction

Complex systems are referred to as systems that are composed of many independent system elements playing a key role in the whole system's behavior. Although the independent system elements follow their own logic and behavior, and interact among themselves, such elements define the dynamic behavior observed in a collectively complex system. Based on the principles of complexity science, the complexity of any system can be characterized by understanding the relations, interactions, and behavior of the constituent system elements. Systems can be formalized by two distinctions, disorganized complexity, and organized complexity [1]. Disorganized complexity is a result of many constituent independent system elements. In such a case, the system elements interact among several others, contributing to random system behavior. To understand such a system, statistical and probabilistic methods are usually used. By comparison, organized complexity comprises non-random behavior of the system elements where the number of parts need not be large for the system to be knowledge emergent. Properties of such a system can be understood with the help of simulations and various modeling techniques. It is to be noted that the whole idea of a complex system is still uncertain, based on the fact that it differs in terms of its definition, understanding, and concept between authors according to their perspectives [2]. To provide a few frequently contested definitions [3] of complex systems over the years:

> "*Complex systems are those with many strongly interdependent variables. This excludes systems with only a few effective variables, the kind we meet in elementary dynamics. It*

also excludes systems with many independent variables; we learn how to deal with them in elementary statistical mechanics. Complexity appears where coupling is important but doesn't freeze out most degrees of freedom" [4].

"Complex system can be defined as a system with large number of components, often called agents or constituent system elements, that interact, adapt and learn" [5].

"Complex systems exhibit several defining characteristics, including feedback, strongly interdependent variables, extreme sensitivity to initial conditions, fractal geometry, and self-organized criticality, multiple metastable states, and a non-gaussian distribution of outputs" [6].

"Systems that tend to be robust and at the same time capable of producing large events. They can attain equilibria, both fixed pints and simple patterns, as well as produce long random sequence" [7].

"Complex Systems are composed out of interdependent parts and violate the ergodic theorem. They have many degrees of freedom whose time dependence is very slow on a microscopic scale" [8].

The varying perspectives on complex systems are spearheading the development of several metrics in the literature to measure systems complexity. These include identifying system complexity based on requirements difficulty, stakeholder relationships, and several similar other measures. However, these measures are not exhaustive, considering the difficulty in accurately measuring systems complexity [3]. As an attempt to contribute to the scientific body of knowledge that addresses understanding systems complexity, this paper illustrates a framework to characterize complex systems based on the concept of Information Entropy, using a two-stage examination method.

The novelty of the proposed approach lies in establishing a translatable framework across complex system domains to use information entropy. This includes incorporating a well-established and validated statistical foundation that shall be able to assess the relationships between sub-systems/components of a complex system, provide information to understand a system, assess and identify the interaction patterns of components/sub-systems in a system, and be able to differentiate between the input and output information of a system. Further, the approach illustrated enables the identification of driving and reacting components/sub-systems of a complex system.

A parallel can be drawn between information entropy and system complexity, where, as a system evolves or changes its state, the entropy will also change. This helps to identify entropy in a complex system as a function of the systems components, and their interactions. Supporting this relationship is the theorization that entropy is a special case applicable to system of systems based on combination of change in information and system disorder, where the increase in information leads to an increase in disorder, also known as entropy [9]. Further, information in a system can be contemplated as an enabler for sub-systems to interact among themselves [9]. System components (in this case, sub-systems of a complex system) and their interactions take information as an anchor that drives a change in a system, thus supporting the statement "Entropy in a complex system is a function of the system components and their interactions".

To test the applicability of the parallel drawn to information entropy, in this paper a research question is formulated:

Q1. Can information entropy help in quantitatively characterizing a complex system?

To address this question, first a two-stage examination method is identified to introduce a structured thought regarding the application of information theory to different systems. Then, a simulation-based case study is introduced with tailored information entropy-based metrics developed to aid in portraying the application of information entropy in complex networks, systems with low structural rigidity (i.e., unpredictable sub-system interactions), and high information gain.

The reminder of this paper is structured as follows. Section 2 expands on the agnostic two-stage examination framework used. Section 3 provides a background on the concept

of entropy and its interpretations, along with the foundation of entropy from information theory justifying the approach adopted in answering the research question. Section 4 provides an overview and definition of complex networks, followed by the case study illustrating entropy-based metric development and its application to complex networks in Section 5. Finally, the discussion and conclusions on the applicability of information theory to understand and characterize complex systems are identified in Section 6.

2. Agnostic Two-Stage Examination Structure for Complex Systems

The two-stage system examination strategy used is illustrated in Figure 1. First, the concepts of entropy were explored in terms of thermodynamic sense, statistical sense, disorder sense, information sense, and homogeneity sense. Second, domain-specific methodologies of complexity in complex systems, for the considered case study, were studied. This involves understanding what complexity is, and later identifying intra- and inter-system dependencies.
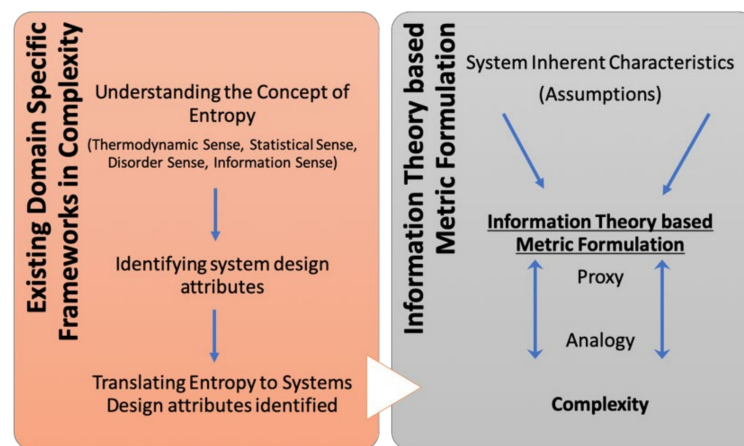

Figure 1. Agnostic two-stage complex systems examination structure.

With a thorough understanding of system characteristics, exploring the benefits and drawbacks of the entropy-based complexity measures helped confirm the need for exploiting various characteristics of a system to establish new techniques to measure and understand the system. This followed engineering entropy-based metrics specific to the case study aimed at providing a proxy that reflects an analogy of what complexity is. To expand, an overview of entropy is first provided, and a foundation for the case study on complex networks, i.e., systems with low structural rigidity (unpredictable sub-system interactions) and high information gain is then built.

3. Understanding Entropy—An Overview

The concept of entropy is often thought of as abstract and, at the same time, difficult to present because of its different applications by various authors in many research fields. Throughout the literature, it is observed that many authors mostly seem to consider entropy to be either a state of disorder or a loss of information. To better understand the distinct sense of the term entropy, a principled polysemy approach [10] adapted from the field of linguistics will help the readers to understand the basic co-existence of many meanings of the word entropy. Five distinct meanings of the word entropy were identified by Haglund et al. [10], viz., thermodynamic sense, statistical sense, disorder sense, information sense, and homogeneity sense. In the thermodynamic sense, entropy can be interpreted as a state of a system that tends to increase towards a maximum value. The statistical sense of entropy presumes a probabilistic approach of identifying, counting, and monitoring the transitions between the microstates of the system. In contrast to the statistical sense, the

disorder sense does not use a probabilistic approach but rather considers the snapshot of a given situation by representing spatial configurations and messiness.

Entropy in terms of information theory can be seen as the average rate of information added by the next element, calculated by considering the complete set of symbols and their probabilities. The information sense of entropy, according to Haglund et al., relates to the information needed to produce or interpret a message using its elements, such as digits, symbols, letters, and words. This entropy model shares the relationship between a message and its constituent elements. To better understand the information sense of entropy, refer to the following text extracted from Shannon's work on mathematical theory of communication [10]:

> "*If a source can produce only one particular message its entropy is zero, and no channel is required. For example, a computing machine set up to calculate the successive digits of π produces a definite sequence with no chance element*" [11,12].

Contrary to the statistical sense, where characteristics are shared by microstates, entropy in the information sense can be used to predict the next or an upcoming message or a symbol in a message based on conditional probabilities, thereby stressing the probabilistic prediction of the next symbol of a message. The main difference emphasized here is the characteristic of entropy according to the statistical sense based upon a system's description through its constituent elements' relationships. The information sense and the disorder sense share this property, whereas the thermodynamic sense does not. Entropy in thermodynamics represents the inherent disorder in a system over a period of time as the system heads towards thermodynamic equilibrium.

In information theory, entropy helps to quantify the information [11]. Quantifying information implies analyzing the information present and measuring its associated uncertainty. Higher values of entropy signify lesser order in a system and lower values of entropy signify a more ordered system. Shannon entropy H is given as:

$$H = -K \sum P_i \log_2 P_i \tag{1}$$

where $K = 1$, P_i is the probability of a symbol appearing in a given stream of symbols, and the use of the logarithm base two corresponds to expressing information entropy in terms of bits.

This paper takes its foundation for entropy from information theory, as given by Shannon. Information theory plays an important role in understanding the meaning of entropy. This is because it is preferable to consider information that can be defined both quantitatively and subjectively. Moreover, information can be used as an anchor to understand what is changing in a spontaneous process. It is to be noted that information theory helps to precisely establish a measure of information defined in terms of probabilities based on some given evidence [13].

Several attempts in the literature are observed on the use of entropy as a tool in understanding uncertainty in domain specific complex systems. To name a few, examples include: in the biological domain, the use of entropy for predicting diverse states of kidney failure disease considering the effects of hypertension and stress [14], and assessing causal relationships between pairs in gene regulatory networks [15]; in image segmentation, a generalized application of entropy in continuous and discrete variables [16]; analysis of financial markets by analyzing the relationship between nine different stock indices from U.S, Europe, and China using entropy measures [17]; detecting fatigue cracks in aluminum alloys [18]; fabric defect detection [19]; and, more interestingly, the quantification of information in cosmological probes of large scale structures [20]. A lack of a unifying approach for characterizing complex systems across domains justifies the framework used to characterize complex systems based on the concept of information entropy, underlying the assumption that entropy of a complex system is a function of its components and their interactions.

4. Complex Networks—An Overview and Definition

A network, as derived from graph theory, is a simplest form representing a collection of points that are joined together in pairs by lines. In every such representation, the set of points are referred to as nodes or vertices, and the lines used to connect the points together are referred as edges. Networks can be mapped from contexts across several domains, such as physical sciences, biological sciences, and social sciences, that consist of several components linked together. It is observed in the literature that physical, biological, and social systems, when mapped as networks, often provide new insights into the behavior and structure of the system in question [21]. The question of when a considered network is said to be complex can be addressed using the principles of complexity science. Several contrasting views [22–26] on complex and complicated systems have been defined throughout the literature from the lens of complexity science. Drawing upon the views of various researchers on characteristics of complicated and complex systems and, looking back into networks topologies and their characteristics [21], in this study a complex network is defined as a network mapped from a system that primarily constitutes interactions observed among systems components, interactions of human actors in the system, and the influence of human interaction on system components. Continual human interactions in such networks play a primary role in influencing the network structure and its evolution. The inherent ambiguity of human interactions and their individual perceptions introduce non-linear behavior, emergence, and self-organization in networks, thereby introducing complexity. Table 1 illustrates the distinguishing characteristics of complicated and complex networks.

Table 1. Characteristics of complicated and complex networks.

Category	Complicated Networks	Complex Networks
Nodes	Structured and predefined set of connections	Entity of random node connections
Characteristic	Highly tuned	Minimally tuned
Topology	Highly organized structure	Continually emergent structure based on dynamics of interactions
Robustness	Robust to targeted perturbations	Robust to random perturbations
Relationships	Easier to understand network relationships	Difficult to understand network relationships
Human influence	Human interactions based on a predefined set of rules	Random human interactions
Network Examples	Software system, automotive system, electronic circuit networks	Transportation, social, communication and biological networks

Figure 2 summarizes an abstract fundamental representation of how we define a complex network. Complexity in a network is illustrated as result of the continual relationship between an actor (human) and the components (sub-systems) of a system; that is, complexity in a network depends upon the extent of interactions a human actor has in a system and the cascading influence human interactions initiate throughout that lead to network emergence.

Complex Network = f (System component interactions, Human interactions, Human & system component interactions)
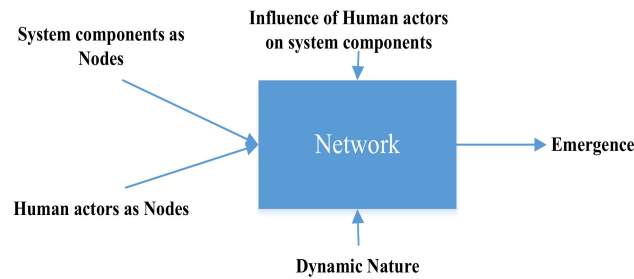

Figure 2. Abstract fundamental representation of a complex network.

To some extent, complex networks can be characterized by the interactions among its constituent nodes. These interactions evolving together result in the dynamic nature of the network. Many complex networks have a number of properties that are common in nature, leaving aside their domain specific characteristics. Displaying small world phenomena is one such characteristic, and the other is the highly heterogeneous nature of interaction patterns in many cases [27]. Heterogeneity in complex networks can be characterized by observing the degree distribution based on the node interaction dynamics. Many observed networks are reported to be scale-free networks, in which the degree distributions are observed to follow a power law. Examples of such networks are electronic circuits, cellular metabolism networks, research collaboration networks, and connectivity in the World Wide Web [28].

5. Case Study—Entropy Metric Formulation and Analysis of Simulated Complex Networks

5.1. Entropy to Understand Interaction Dynamics

To understand the dynamic structure of a given complex network as a whole, it is important to analyze and explore the main conceptual underpinnings of a network, which are the nodes and their interactions. Because complex networks are continually emergent and non-linear in nature, exploring the dynamics of constituent nodes and their interaction patterns helps in providing insight about how a network structure evolves. Considering the degree distribution of every individual node in a network, let P_{k_i} be the dynamic degree of a node where it gives the probability of a particular node having k links at a given time-step i. An entropy measure for a static network is first defined, which then is expanded to consider the dynamic nature of complex networks.

Suppose there are n nodes in a network space N, $N = (N_1, N_2, \ldots, N_n)$. Now let K be a class of N where it defines the number of links of each observed node n of a network. $K = (K_1, K_2, \ldots, K_n)$. Let I be the proportion of interactions that are represented by the number of links associated with each node n:

$$I_n = \frac{K_n}{(K_1 + K_2 + \ldots + K_n)} = \frac{K_n}{\sum_{i=1}^{n} K_i} \tag{2}$$

Entropy is mathematically represented as:

$$H(X) = -\sum_{i=1}^{m} p(x) \log_2 p(x) \tag{3}$$

where $0 < H(X) < \log_2 m$ and $0 < H(X) < \log_2 H(X)$. It is to be noted that the value of $P \log P = 0$ when $P = 0$.

Applying this measure to the number of nodes constituent in the network space we obtain:

$$H(X) = -\sum_{i=1}^{n} I_n \log_2 I_n \tag{4}$$

Now, considering the dynamic nature of a complex network, let j be the number of time-steps over which a network is observed. At every time-step j, each and every node n of the network will have a degree k. Now, considering the degree of nodes at individual time-steps of a network, entropy is given by:

$$H(X) = \sum_{t=1}^{j} \left(-\sum_{i=1}^{n} I_{n_t} \log_2 I_{n_t}\right) \tag{5}$$

Here, a complex network is quantified to be a function of the number of nodes present (N_n) and the number of links (I_n) associated with the nodes. The uncertainty based on the distribution of the node degrees over a time period is used as a base indicator to calculate network entropy. This also takes into consideration the diversity in network interactions. When calculated, this index represents absolute order in the network when it takes a value of 0, and represents disorder and absolute diversity when it takes a maximum value (the maximum value depends upon the type of complex network being considered).

5.2. Transfer Entropy to Understand the Transfer of Information in Network over Time

The use of entropy in the field of networks is a well-studied area. Entropy has been identified as a possible metric for playing a crucial role in analyzing the complexity of a network [29]. To further extend entropy towards measuring uncertainty between two random variables X and Y, a measure of mutual information is used. Mutual information is based only on the present state/symbol of each variable; considering all such present states, mutual information, $I(X, Y)$, is defined by:

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log \frac{P(x, y)}{P(x) \times P(y)} \tag{6}$$

However, the fact that the measure of mutual information is symmetrical, i.e., $I(X, Y) = I(Y, X)$, implies that the future state of a random variable has a casual effect on the past state. In the case of complex networks, where a network consisting of several nodes evolves over a period of time, the structure of a network at a time t depends upon the evolution of the network until time $t-1$, not the other way around. Thus, in order to address the symmetrical limitation of the mutual information measure, the concept of transfer entropy was proposed [30].

Transfer entropy (TE) was coined [31] to address the time symmetric limitation of a mutual information measure. Considering two sample spaces of information represented in time by $X = \{x_1, x_2, x_3, \ldots, x_t\}$ and $Y = \{y_1, y_2, y_3, \ldots, y_t\}$, transfer entropy is defined as the additional amount of information gained for the next observation of one of the two processes being considered, given the past observation of the other process.

Following Murcio et al. [30], considering two systems X and Y, we first define the entropy rate (i.e., entropy based on time t) assuming that y_{t+1} depends upon both x_t and y_t as:

$$H_A = -\sum_t p(y_{t+1}, y_t, x_t) \log p(y_{t+1}/y_t, x_t) \tag{7}$$

We now define the entropy rate in which y_{t+1} depends only on y_t:

$$H_B = -\sum_t p(y_{t+1}, y_t, x_t) \log p(y_{t+1}/y_t) \tag{8}$$

Hence, the transfer of information from X to Y is defined as *TE* (X, Y) i.e., transfer of information from X to Y: *TE* $(X, Y) = H_B - H_A$.

TE, the transfer of information between two random variables X and Y, is given by [30,31]:

$$TE\ (X,Y)\ =\ \sum_{t=1} p(y_{t+1}, y_t, x_t) \log \frac{p(y_{t+1}, y_t, x_t)\ \times\ p(y_t)}{p(y_t, x_t)\ \times\ p(y_{t+1}, y_t)} \tag{9}$$

Moreover, the transfer entropy from Y to X can be inferred similarly based on the above as:

$$TE\ (Y,X)\ =\ \sum_{t=1} p(x_{t+1}, x_t, y_t) \log \frac{p(x_{t+1}, x_t, y_t)\ \times\ p(x_t)}{p(x_t, y_t)\ \times\ p(x_{t+1}, x_t)} \tag{10}$$

To analyze the flow of information generated from one node to the other of a considered network based on the observed individual node degree evolution over several time-steps of network formation, the concept of transfer entropy is used. When applied to a network evolution scenario, TE helps in understanding how much of the information generated at one node is responsible for the information obtained by the other node.

5.3. Simulation Models Used to Apply the Metrics

To calculate entropy and TE values based on individual node degree evolution, data was generated using the NetLogo agent-based simulation platform. The advances observed in the use of agent-based modeling (ABM) to represent social systems compared to traditional equation-based models, and their ability to mimic heterogenous subsystems, autonomous entities, and nonlinear relationships, affirmed the use of ABM [32]. Tracing back to the principles of complexity science, a system is a sum of its parts as a whole but not the other way around. Two random networks i.e., preferential attachment and small world networks, were generated. Data was captured individually for each network based on the number of nodes present along with the individual node degrees (number of connections) over several time-steps. ABM helps to model and represent individual components of a system considering all the possible attributes and their behavior instead of representing the state of a whole system. ABM models were used to generate network models where individual agents (in this case network nodes) are describes as individual entities that interact with each other locally in their environment [33]. The characteristics of the networks generated, and the calculations of transfer entropy values, are discussed below. The focus is to model network dynamics of the system that arises when individual nodes interact with each other, and on how such individual nodes affect the overall network. This simulates a few characteristics of complex systems and their independent sub-systems, such as degrees of freedom, emergence, and adaptability.

5.3.1. Preferential Attachment Networks

Generated networks using the NetLogo software (version 5.3.1) illustrate the behavior of real-world networks (such as connection from and to a website, social networks, and collaboration networks), where few nodes have a large number of connections, whereas all the other nodes have only a few. This phenomenon in which network nodes prefer to connect to the popular existing nodes is called preferential attachment. This model starts with two nodes connected initially, and thereby every new originating node randomly picks an existing node with which to connect, with some inherent bias, i.e., the chance of a node being chosen to be connected to by another node is directly proportional to the number of connections (degree) it already has. The networks that arise from this phenomenon often follow a power law distribution, i.e., the distribution of the number of connections of each node is not normal [34]. Barabasi and Albert originally coined this mechanism for creating scale-free networks, so the networks created by this mechanism are called Barabasi scale-free networks. Figure 3 illustrates the evolution of interactions in the network over fifteen different snapshots.
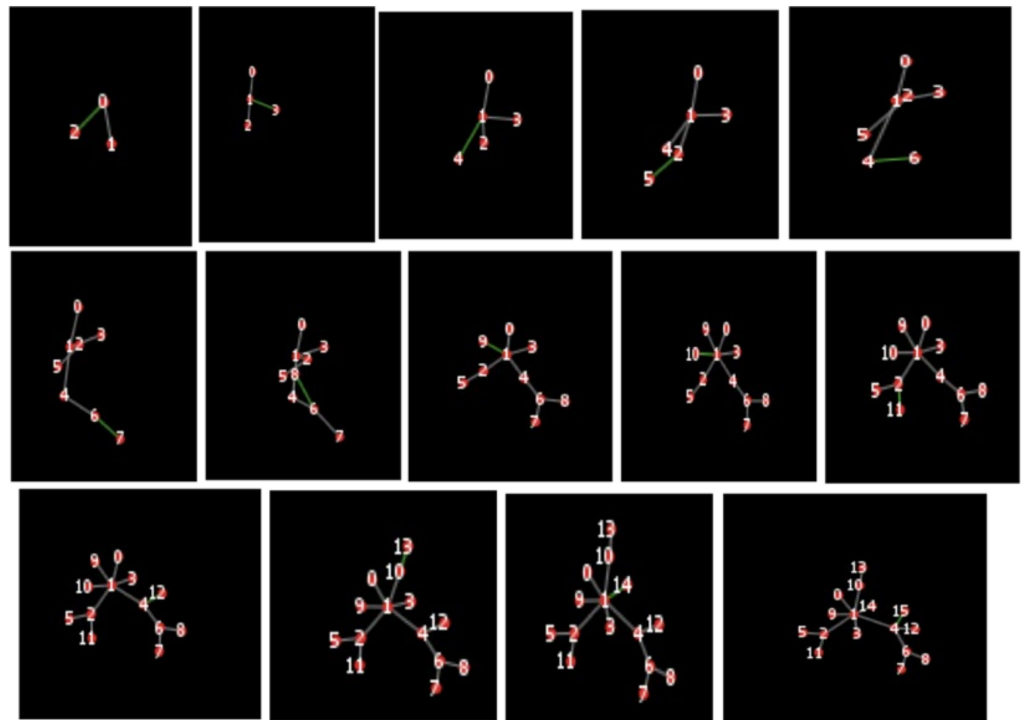
Figure 3. Emergence of network based on preferential attachment phenomenon.

To analyze the flow of information generated from one node to another of the network based on the individual node degree evolution over several time-steps of network formation, the concept of transfer entropy (TE) is used. To generate TE values based on individual node degree evolution, data on the number of nodes present along with the individual node degrees over 15 different time-steps were captured using NetLogo [35]. Table 2 illustrates the degree evolution of nodes portrayed in Figure 3.

Table 2. Time series showing the node degrees generated at each time-step.

Time-Steps Node	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	3	4	4	4	4	4	5	6	6	6	6	7	7
2	0	1	1	1	2	2	2	2	2	1	1	1	2	2	2
3	0	0	1	1	1	1	1	1	1	2	1	1	1	1	1
4	0	0	0	1	1	2	2	2	2	1	3	1	1	1	1
5	0	0	0	0	1	1	1	1	1	2	1	3	1	1	1
6	0	0	0	0	0	1	2	3	3	1	2	1	3	1	1
7	0	0	0	0	0	0	1	1	1	3	1	3	1	3	1
8	0	0	0	0	0	0	0	1	1	1	3	1	3	1	3
9	0	0	0	0	0	0	0	0	1	1	1	3	1	3	1
10	0	0	0	0	0	0	0	0	0	1	1	1	3	1	4
11	0	0	0	0	0	0	0	0	0	0	1	1	1	3	1
12	0	0	0	0	0	0	0	0	0	0	0	1	1	1	3
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

To calculate the TE values (see Table 3), joint probabilities are calculated for the emerging node degrees of Table 2. Thus, to calculate the TE from node 0 to node 1 at time-step t1, we need to find the probability p (y2, y1, x1) as required for calculating TE, where y2 corresponds to the value located in the second column, second row; y1 corresponds to the

value in the first column, second row; and x1 corresponds to the value in the first column, first row; i.e., *p* (1,1,1). We now count the number of matching combinations of these values that exist in the two rows corresponding to node 0 and node 1. In the graph (Figure 4), the X-axis represents the node numbers and the Y-axis represents the corresponding TE values. It is read as: for the red line, Mark 0 is the TE value from node 0 to node 1, Mark 1 is the TE value from node 1 to node 2, and so on. For the blue line, mark 0 is the TE value from node 1 to node 0, mark 1 is the TE value from node 2 to node 1, and so on.

Table 3. TE values calculated based on node degrees from Table 2.

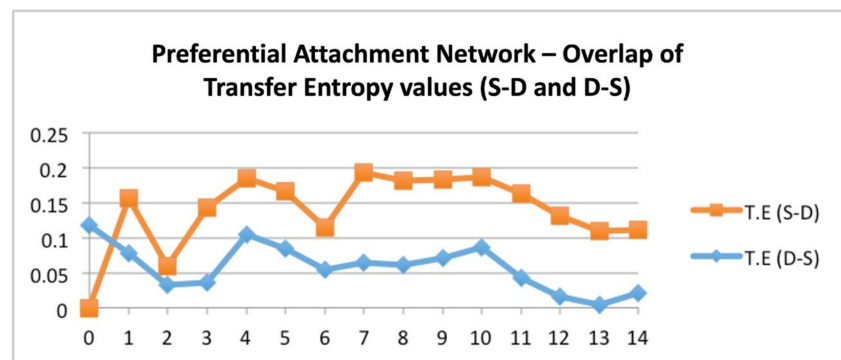| Source Node	Destination Node	Transfer Entropy (S-D)	Transfer Entropy (D-S)
0	1	0.000	0.117
1	2	0.155	0.077
2	3	0.058	0.032
3	4	0.143	0.035
4	5	0.185	0.104
5	6	0.166	0.084
6	7	0.114	0.054
7	8	0.193	0.064
8	9	0.182	0.061
9	10	0.183	0.070
10	11	0.186	0.086
11	12	0.163	0.043
12	13	0.131	0.016
13	14	0.109	0.003
14	15	0.111	0.021

Figure 4. Transfer entropy values between the nodes of the preferential attachment network.

It is seen that, in general (see Figure 4), information flow from the source nodes to the destination nodes initially rises and then dominates information flow from destination nodes to source nodes. At the early stages of the graph, the trend line of S-D, when observed, shows that TE between node 0 and 1 is zero with node 0 as the source and node 1 as the destination.

This simply suggests to the initial network structural formation where at the initial time-step nodes 0 and 1 are connected, on which the simulated network builds over the considered time series. Before calculating the network entropy measure, we introduce a measure of connectedness, also called the reproductive number in complex networks. The reproductive number, R_0, helps to understand the context of information spread in a network [36]. R_0, also known as the connectedness of a network, can be calculated as follows.

Consider a node *n* with degree *k* to communicate an idea or spread information to its neighbors with a probability *r*. The expected number of nodes it will pass the information

to will be $r \times (k-1)$, by excluding the nodes it previously communicated the information to. Taking a weighted average over all the nodes we obtain:

$$R_0 = r \times \frac{\sum_{i=1}^n K_i(K_i - 1)}{\sum_{i=1}^n K_i} \tag{11}$$

If the value of R_0 is greater than 1, it implies that the number of nodes receiving the information grows exponentially; if R_0 is less than 1 it implies that the information will dissipate rapidly over a network.

Table 4 and Figure 5 portray that, as the network evolves, entropy increases, implying that in a preferential attachment network with more nodes evolving over a period of time, the network as a whole tries to reach a disordered/more diverse state from an ordered state initially. As entropy (diversity) increases, the reproduction number initially increases and then decreases. This implies that when the network starts to evolve, there is a greater chance of information spread, and when the network starts evolving to more than four nodes at time-step 3, the information spread slowly decreases.

Table 4. Temporal evolution of entropy and R_0 values for the generated preferential attachment network model.

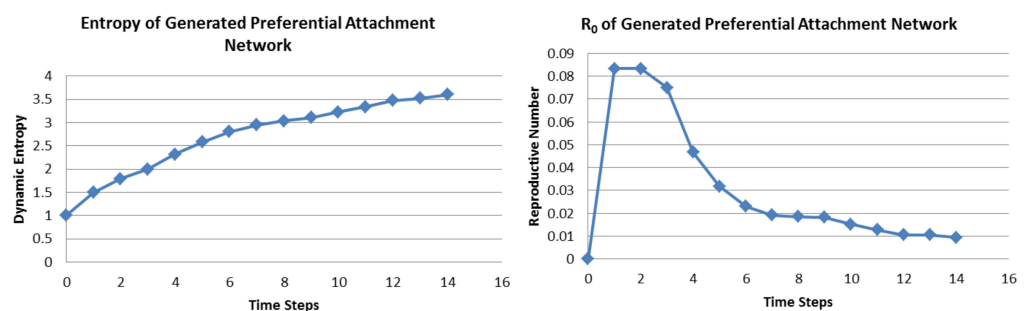| Time-Step	Entropy	R_0
0	1.00	0.00
1	1.50	0.083
2	1.79	0.083
3	2.00	0.075
4	2.32	0.046
5	2.58	0.031
6	2.80	0.022
7	2.95	0.019
8	3.038	0.018
9	3.10	0.018
10	3.23	0.015
11	3.34	0.012
12	3.47	0.010
13	3.52	0.010
14	3.60	0.009

Figure 5. Temporal evolution of entropy (**left**) and R_0 measures (**right**) of the generated preferential attachment network.

5.3.2. Small World Networks

The network generated here using NetLogo simulation software is based on the phenomenon of small word networks. This phenomenon implies that a given person is only a small number of (a few) connections away from any other person in the world. A popular example of this phenomenon is the famous Kevin Bacon network (also known as a six-degree separation network), which relates to a network generated based on actors appearing in the same movie. However, small world networks are not only limited to

networks of people, but also apply to several other real-time networks such as power grids. The network model generated here is based on a few assumptions and conditions under which a small world network is formed. It is developed based on the model suggested by Duncan Watts and Steve Strogatz [37]. The model starts by initially generating a network in which each node is connected to its two neighbors on either of its sides [35].

After this initial random network is generated, at each and every time-step, a random connection is picked and then rewired i.e., a random end of a connected pair of nodes is changed. The probability that a random edge is chosen and rewired is based on the rewiring probability value assigned to the network generation model. Figure 6 illustrates the evolution of interactions in the network over fifteen different snapshots. Tables 5 and 6 illustrate the node degree evolution observed for the generated small world network and their respective transfer entropy values.
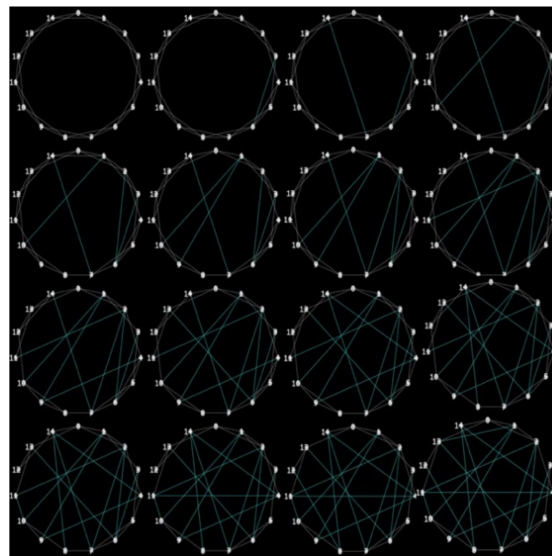

Figure 6. Small world network emergence generated.

Table 5. Time series showing the node degrees generated at each time-step of a small world network.

Time-Steps Nodes	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
1	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5
2	4	4	4	4	5	5	5	6	6	6	6	6	6	6	6
3	4	4	4	4	4	4	3	3	3	3	3	3	3	3	2
4	4	3	3	3	3	3	3	4	4	4	4	4	4	4	4
5	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3
6	4	5	5	5	5	5	5	5	6	5	5	5	5	5	5
7	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5
8	4	4	4	4	3	3	3	3	3	3	3	3	3	4	4
9	4	4	3	3	3	4	4	4	4	4	3	3	3	3	3
10	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
11	4	5	4	3	3	3	3	2	2	2	2	3	4	4	4
12	4	3	4	4	4	4	4	3	3	3	3	2	2	2	2
13	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5
14	4	4	5	5	5	4	4	4	3	4	5	5	5	4	4

Table 6. TE values calculated based on node degrees from Table 5.

Source Node	Destination Node	Transfer Entropy (S-D)	Transfer Entropy (D-S)
0	1	0.000	0.0002
1	2	0.069	0.0006
2	3	0.045	0.0882
3	4	0.206	0.0106
4	5	0.022	0.0137
5	6	0.010	0.0059
6	7	0.006	0.0185
7	8	0.012	0.0077
8	9	0.021	0.0609
9	10	0.000	0.0000
10	11	0.000	0.0022
11	12	0.084	0.1661
12	13	0.052	0.0267
13	14	0.028	0.0226

As observed in Figure 7, Node 0 has no information originating from it. This simply reflects the rigidity of the node, which does not change the degree over the network emergence (the node is connected to the same neighbors throughout), and thereby does not generate any new information.
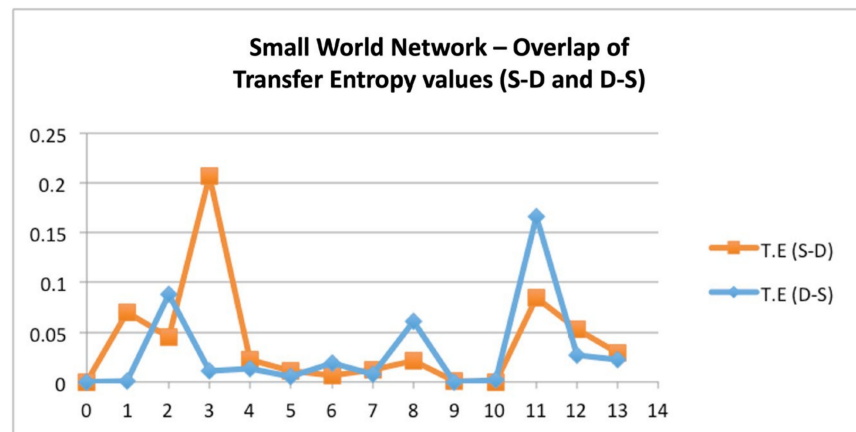

Figure 7. Transfer entropy values between the nodes of small world network.

Figure 8 and Table 7 illustrate the entropy measure and R_0 calculated at every time-step of the dynamic network. It is observed that, as the network evolves, entropy decreases, implying that the small world network tries to reach an ordered state from a highly disordered initial state. As entropy (diversity) reduces the reproduction rate R_0 increases. This implies that when this network tends to reach absolute order (i.e., less entropy) the spread of information increases over the network.
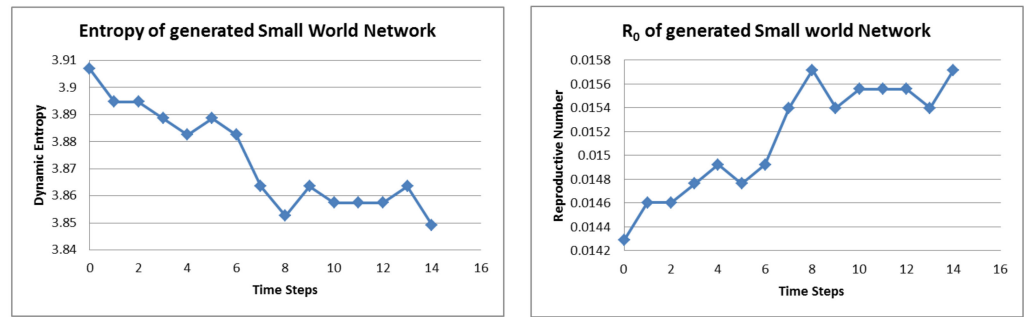
Figure 8. Dynamic entropy (**left**) and R_0 measures (**right**) of the generated small world network.

Table 7. Temporal evolution of entropy and R_0 values for the generated small world network model.

Time-Step	Dynamic Entropy	R_0
0	3.906	0.014286
1	3.894	0.014603
2	3.894	0.014603
3	3.888	0.014762
4	3.882	0.014921
5	3.888	0.014762
6	3.882	0.014921
7	3.863	0.015397
8	3.852	0.015714
9	3.863	0.015397
10	3.857	0.015556
11	3.857	0.015556
12	3.857	0.015556
13	3.863	0.015397
14	3.849	0.015714

The framework of information theory used has no restrictions on the type of complex system it is applied to because the interactions among the components/subsystems of a complex system can be viewed as information transfer. Perceived applications of the proposed framework across various operational domains can include addressing time critical system of systems (SoS), such as testing autonomous system adaptation platforms for understanding and assessing the impact of adding additional system functionalities [38], mission critical weapon SoS [39], assessing complexity in manufacturing systems [40], identifying software code complexities [41], and assessing commercial off the shelf products [42]. However, the use of TE may lead to the curse of dimensionality [43] with an exponential increase in the number of nodes and connections, thus limiting the scope of the proposed method. This may be plausibly addressed using decomposed transfer entropy [44], thus providing the opportunity for further investigation of the proposed approach and its limitations.

6. Discussion and Conclusions

The simulated case study considered in this paper considers complex networks. To examine the use the information theoretic approach, we mapped the similarity of a complex network to a general communication channel. A typical communication channel consists of a source that sends information, a channel that acts as a medium to transmit the information, and a destination that receives the transmitted information. In terms of complex networks, network nodes and actors behave as both source and destination, constantly transmitting and receiving information, whereas the edges connecting the nodes of a complex network act as the medium of information transfer. The application of the concepts from information theory helped in understanding, at a system and sub-system level, the interdependencies

among the considered system constituents. This provided system analysis insights which led to the development of specific metrics for the tailored entropy-based case study.

Table 8 illustrates a summarized representation of how information theoretic concepts and their applications were relevant in understanding the case study used. This paper presenting the case study is framed according to the application of information theory concepts with a goal to better understand and assess complex systems. We present the insight drawn from the case study in answering the research question identified i.e., Can information entropy help in quantitatively characterizing a complex system?

Table 8. Information theoretic concepts and their application towards complex systems.

Concepts and Measures from Information Theory	Complex System Characterization
Entropy	• Helps to understand the diversity of network node interactions • Helps to understand the context of information spread in a network over its evolution
Transfer Entropy	• Helps to identify the structural implication of nodes in a complex network on their information processing ability • Helps to identify the information processing ability of a network • Helps to quantify how much information is needed to describe a network of a given scale

Because not all complex systems have the same characteristics and behave in a similar fashion, it is indeed not possible to provide a generalized metric to quantitatively characterize a complex system. However, using information theory concepts tailored to the specific system being considered helps to answer and quantify the complexity, based on a set of assumptions. In such a case, a thorough understanding of the considered systems, sub-systems, and their interactions is required. Using the concept of entropy, complex systems can be mapped as a network of interactions, and the characteristics of the complex system can be quantified, thereby setting a foundation for understanding a complex system. The complexity can then be quantified based on a set of assumptions. For example, when a complex system is mapped as a network, information entropy can be used to quantify the network at a given snapshot based on the observed sub-systems' interaction patterns. Considering the network dynamics, at every instant the network emerges with either new nodes introduced into the network or rewiring of the connections among the nodes of the networks. Provided that the data on system interaction patterns at every instance of the system evolution is available, entropy-based measures can detect the precise time instance at which the system is exposed to external stimuli. This is because entropy can be seen as a measure of surprise in which, following sudden external stimuli to the network, a rapid change in the tailored entropy-based measure can be detected when observed over the system dynamics

The insight drawn is strictly based on the observations from the case study presented in this paper. Although the insights gained directly may not address complex system characteristics of every possible scenario, they set a foundation for a fact-based theoretical framework for explaining, understanding, and analyzing complex systems using information theory concepts. A more realistic application of the proposed framework to time-series evolution data of real-world complex systems would validate the use of entropy as a generalized measure to characterize complex systems. However, this may also lead to a need for identifying the influence of dimensionality, i.e., the role played by the increase in nodes and connections in a complex system's evolution, which needs to be further studied.

References

1.   Weaver, W. *Science and Complexity*; American Scientist: Boston, MA, USA, 1948; Volume 36, p. 536.
2.   Baranger, M. *Chaos, Complexity, and Entropy*; New England Complex Systems Institute: Cambridge, MA, USA, 2000.
3.   Potts, M.; Harvey, D.; Johnson, A.; Bullock, S. The Complexity Register: A Collaborative Tool for System Complexity Evaluation. *Eng. Manag. J.* **2021**, 1–17. [CrossRef]
4.   Boccara, N. *Modeling Complex Systems*; Springer: Berlin/Heidelberg, Germany, 2004.
5.   Holland, J.H. Studying Complex Adaptive Systems. *J. Syst. Sci. Complex.* **2006**, *19*, 1–8. Available online: http://hdl.handle.net/2027.42/41486 (accessed on 5 November 2021). [CrossRef]
6.   Kastens, K.A.; Manduca, C.A.; Cervato, C.; Frodeman, R.; Goodwin, C.; Liben, L.S.; Mogk, D.W.; Spangler, T.C.; Stillings, N.A.; Titus, S. How geoscientists think and learn. *Eos Trans. Am. Geophys. Union* **2009**, *90*, 265–266. [CrossRef]
7.   Page, S.E. *Diversity and Complexity*; Princeton University Press: Princeton, NJ, USA, 2010; Volume 2.
8.   Bar-Yam, Y. *Dynamics of Complex Systems*; CRC Press: Boca Raton, FL, USA, 2019.
9.   Johnson, J.J.; Tolk, A.; Sousa-Poza, A. A Theory of Emergence and Entropy in Systems of Systems. *Procedia Comput. Sci.* **2013**, *20*, 283–289. [CrossRef]
10.   Haglund, J.; Jeppsson, F.; Strömdahl, H. Different senses of entropy—Implications for education. *Entropy* **2010**, *12*, 490–515. [CrossRef]
11.   Shannon, C.E.; Warren, W. The mathematical theory of information. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
12.   Shannon, C.E. A mathematical theory of communication. *ACM SIGMOBILE Mob. Comput. Commun. Rev.* **2001**, *5*, 3–55. [CrossRef]
13.   Ben-Naim, A. *Entropy Demystified: The Second Law Reduced to Plain Common Sense with Seven Simulated Games*; World Scientific: Singapore, 2008.
14.   Waezizadeh, T.; Mehrpooya, A.; Rezaeizadeh, M.; Yarahmadian, S. Mathematical models for the effects of hypertension and stress on kidney and their uncertainty. *Math. Biosci.* **2018**, *305*, 77–95. [CrossRef]
15.   Tung, T.Q.; Ryu, T.; Lee, K.H.; Lee, D. Inferring gene regulatory networks from microarray time series data using transfer entropy. In Proceedings of the Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07), Maribor, Slovenia, 20–22 June 2007; IEEE: Piscataway, NJ, USA, 2007.
16.   Kittaneh, O.A.; Khan, M.; Akbar, M.; Bayoud, H.A. Average entropy: A new uncertainty measure with application to image segmentation. *Am. Stat.* **2016**, *70*, 18–24. [CrossRef]
17.   He, J.; Shang, P. Comparison of transfer entropy methods for financial time series. *Phys. A Stat. Mech. Appl.* **2017**, *482*, 772–785. [CrossRef]
18.   Karimian, F.; Modarres, M.; Bruck, H.A. A new method for detecting fatigue crack initiation in aluminum alloy using acoustic emission waveform information entropy. *Eng. Fract. Mech.* **2020**, *223*, 106771. [CrossRef]
19.   Liu, G.; Zheng, X. Fabric defect detection based on information entropy and frequency domain saliency. *Vis. Comput.* **2020**, *37*, 1–14. [CrossRef]
20.   Pinho, A.M.; Reischke, R.; Teich, M.; Schäfer, B.M. Information entropy in cosmological inference problems. *Mon. Not. R. Astron. Soc.* **2021**, *503*, 1187–1198. [CrossRef]
21.   Newman, M. *Networks: An Introduction*; OUP Oxford: Oxford, UK, 2010.
22.   Snowden, D.J.; Boone, M.E. A leader's framework for decision making. *Harv. Bus. Rev.* **2007**, *85*, 68. [PubMed]
23.   Alderson, D.L.; Doyle, J.C. Contrasting views of complexity and their implications for network-centric infrastructures. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2010**, *40*, 839–852. [CrossRef]
24.   Poli, R. A note on the difference between complicated and complex social systems. *Cadum* **2013**, *2*, 142.
25.   Smith, E.D. Gödel's Incompleteness and Consistency Theorems Elucidated with Principles of Abstraction Levels, Complementarity, and Self-Reference. *Philos. Math. Educ. J.* **2013**, *27*.
26.   Rouse, W.B. *Modeling and Visualization of Complex Systems and Enterprises: Explorations of Physical, Human, Economic, and Social Phenomena*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
27.   Solé, R.V.; Valverde, S. Information theory of complex networks: On evolution and architectural constraints. In *Complex Networks*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 189–207.
28.   Newman, M.; Albert-Laszlo, B.; Watts, D.J. *The Structure and Dynamics of Networks*; Princeton University Press: Princeton, NJ, USA, 2006.

29. Anand, K.; Bianconi, G. Entropy measures for networks: Toward an information theory of complex topologies. *Phys. Rev.* **2009**, *80*, 045102. [CrossRef]
30. Murcio, R.; Morphet, R.; Gershenson, C.; Batty, M. Urban transfer entropy across scales. *PLoS ONE* **2015**, *10*, e0133780. [CrossRef]
31. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **2000**, *85*, 461. [CrossRef]
32. An, L.; Grimm, V.; Sullivan, A.; Ii, B.T.; Malleson, N.; Heppenstall, A.; Vincenot, C.; Robinson, D.; Ye, X.; Liu, J.; et al. Challenges, tasks, and opportunities in modeling agent-based complex systems. *Ecol. Model.* **2021**, *457*, 109685. [CrossRef]
33. Railsback, S.F.; Volker, G. *Agent-Based and Individual-Based Modeling: A Practical Introduction*; Princeton University Press: Princeton, NJ, USA, 2011.
34. Barabási, A.-L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [CrossRef]
35. Wilensky, U. "NetLogo preferential attachment model." Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, USA. 2005. Available online: http://ccl.northwestern.edu/netlogo/models/PreferentialAttachment (accessed on 5 November 2021).
36. Easley, D.; Kleinberg, J. Networks, crowds, and markets: Reasoning about a highly connected world. *Significance* **2012**, *9*, 43–44.
37. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world'networks. *Nature* **1998**, *393*, 440–442. [CrossRef] [PubMed]
38. Štefanič, P.; Cigale, M.; Jones, A.C.; Knight, L.; Taylor, I.; Istrate, C.; Suciu, G.; Ulisses, A.; Stankovski, V.; Taherizadeh, S.; et al. SWITCH workbench: A novel approach for the development and deployment of time-critical microservice-based cloud-native applications. *Future Gener. Comput. Syst.* **2019**, *99*, 197–212. [CrossRef]
39. Zhao, Q.; Hu, W.; Xia, B.; Ding, J.; Li, M. The capability spaces complexity measure method of weapon system of systems. In Proceedings of the 2020 IEEE International Systems Conference (SysCon), Montreal, QC, Canada, 24 August–20 September 2020; IEEE: Piscataway, NJ, USA; pp. 1–5. [CrossRef]
40. Vidal, G.H.; Hernández, J.R.C.; Minnaard, C. Modeling and statistical analysis of complexity in manufacturing systems under flow shop and hybrid environments. *Int. J. Adv. Manuf. Technol.* **2021**, 1–10. [CrossRef]
41. Akundi, A.; Smith, E.; Tseng, T.-L. Information entropy applied to software based control flow graphs. *Int. J. Syst. Assur. Eng. Manag.* **2018**, *9*, 1080–1091. [CrossRef]
42. Akundi, A.; Smith, E.; Tseng, T.-L.; Rubió, I. Quantifying system structural complexity using design structure matrices. In Proceedings of the 2018 Annual IEEE International Systems Conference (SysCon), Vancouver, BC, Canada, 23–26 April 2018; IEEE: Piscataway, NJ, USA; pp. 1–8. [CrossRef]
43. Chen, L. Curse of Dimensionality. In *Encyclopedia of Database Systems*; Liu, L., Özsu, M.T., Eds.; Springer: Boston, MA, USA, 2009. [CrossRef]
44. Runge, J.; Heitzig, J.; Petoukhov, V.; Kurths, J. Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.* **2012**, *108*, 258701. [CrossRef]