Controllability Robustness of Henneberg-Growth Complex Networks

LOU, Yang; YANG, Dong; WANG, Lin; TANG, Chang-Bing; CHEN, Guanrong

Published in:
IEEE Access

Published: 01/01/2022

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1109/ACCESS.2022.3141011

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEo database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Download date: 08/10/2023
Controllability Robustness of Henneberg-Growth Complex Networks

YANG LOU\textsuperscript{1,2}, (Member, IEEE), DONG YANG\textsuperscript{3}, LIN WANG\textsuperscript{2,4}, (Senior Member, IEEE), CHANG-BING TANG\textsuperscript{5}, (Member, IEEE), AND GUANRONG CHEN\textsuperscript{6}, (Life Fellow, IEEE)

\textsuperscript{1}Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China
\textsuperscript{2}Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China
\textsuperscript{3}Department of Management of Complex Systems, Ernest and Julio Gallo Management Program, School of Engineering, University of California, Merced, CA 95343, USA
\textsuperscript{4}Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China
\textsuperscript{5}Department of Electronics and Information Engineering, Zhejiang Normal University, Jinhua 321004, China
\textsuperscript{6}Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China

Corresponding author: Yang Lou (felix.lou@ieee.org)

This work was supported in part by the National Natural Science Foundation of China under Grant 62002249 and Grant 61873167; in part by the Foundation of Key Laboratory of System Control and Information Processing, Ministry of Education, China, under Grant Scip202103; in part by the Lam Woo Research Fund at Lingnan University under Grant LWP20012; in part by the Research Committee of Lingnan University through Faculty Research under Grant DB22A3; and in part by the Hong Kong Research Grants Council through the GRF under Grant CityU11206320.

\section*{ABSTRACT}
It is important for a networked-system to have both good controllability and robustness, where the former measures the ability that the networked system can be properly steered via control to any target state from any initial state within a finite time duration, while the later reflects how well this system can regain its controllability after being destructively attacked. Empirical observations suggest that multi-loop structures are beneficial for enhancing the controllability robustness. The Henneberg-growth mechanism in social networks offers a natural manner to generate multi-clique structures that can be developed to multi-loops by assigning proper edge directions. In this paper, a series of random polygon networks are generated, forming random triangle, rectangle, pentagon and hexagon networks. Then, their controllability robustness is investigated. A realistic measure is designed for characterizing their controllability robustness, which can be used to filter out trivial performances after the network is severely destructed, so that the computation and analysis become much more efficient. Extensive simulation results suggest that, for random polygon networks, 1) non-loop polygon structures are inferior to the loop polygons, confirming that the multi-loop structures are indeed beneficial for controllability robustness; 2) polygons with more sides possess better controllability robustness; and 3) the correlation between controllability robustness and connectivity robustness is weak, implying that the two objectives cannot be enhanced in the same way.

\section*{INDEX TERMS}
Complex network, controllability, Henneberg-growth, polygon, robustness.

\section*{I. INTRODUCTION}
Network controllability measures the ability of a networked-system that can be steered by external input from any initial state to any target state within a finite duration of time [1]–[5]. In engineering applications, it is necessary for networked systems to be controllable and also be strong against malicious attacks or random failures, which take place in the forms of node- and/or edge-removals, causing significant damages to the systems. In the networking world today, more and more frequent and severe attacks and failures occur everywhere [6]–[10]. Regarding the controllability robustness of networked systems, which refers to the ability of a network to regain its controllability after being destructively attacked, strong robustness is desirable and often necessary.

If a given networked system is not controllable, then some external control inputs are needed to be placed at some network nodes (called driver nodes), so as to make it controllable. Thus, the network controllability can be measured by counting how many driver nodes are necessary. As a normalization, the ratio of the minimum number of driver nodes versus the total number of nodes in the network is commonly used. By nature, the number of needed driver nodes fills the rank deficit of the network controllability matrix, while...
the ratio quantifies the “level” of controllability by taking into account the network scale. With this measure, a network with good controllability requires a small ratio to maintain or regain its controllability, implying a low external control cost. Actually, controllability robustness can be measured by different means and in different forms. One effective method to evaluate it is through the process of retaining the network controllability against sequential attacks. This is quantified by using the entire processing results of the aforementioned “levels” of network controllability subject to sequential attacks, visualized by a curve (called controllability curve) of all values of the ratio between the number of driver nodes and the network size, or by the averaged controllability over the entire attack process.

It took quite a long time for people to understand the intrinsic relation between topology and controllability and its robustness of a general directed network [5], [9], [11], [12]. Recently, it is empirically found that the multi-chain and multi-loop structures are beneficial for enhancing the network controllability robustness [13]–[16]. For example, the $q$-snapback model with multi-loop structures [14], [16] shows stronger controllability robustness against various node- and edge-removal attacks than other typical network models [15].

In social networks, on the other hand, many acquaintances are built in a triangular form. It is commonly observed that, when being introduced into a community, a newcomer is likely to acquaint an individual and simultaneously one of his/her direct neighbors. Thus, a strengthened connectivity is established in the network. Meanwhile, some newcomers may influence the existing acquaintances within the community, causing some acquaintances to be disconnected. These two sides of the new connectivity naturally form a planar structure, known as the Henneberg-growth framework [17]–[19], which provides a precise mechanism for modeling undirected social networks such as Facebook [19]. Inspired by the Henneberg-growth mechanism, directed network models were constructed, namely the random triangle (RT) and random rectangle (RR) networks [15], which contain many loops, namely 3-node triangles in RT and 4-node rectangles in RR. Simulation results reveal that RT and RR possess strong resistance against node- and edge-removal attacks [15].

In this paper, the RT and RR models are further extended to a series of random polygon networks, and their controllability robustness is systematically investigated.

The main contributions of this paper are summarized as follows:

1) A series of directed network models generated based on the Henneberg-growth mechanism are proposed and investigated, namely, RT, RR, random pentagon (RP) networks, and random hexagon (RH) networks. In these models, newcomers are connected to the existing community in the form of a group rather than one by one individually. Different patterns of internal connections of newcomers in the group are compared. Specifically, newcomers are organized as either an $\alpha$-chain ($\alpha = 1, 2, 3$ or 4) or a $\beta$-clique ($\beta = 1, 2$, or 3).

Then, never before, the controllability robustness of these networks are carefully investigated and compared in detail.

2) A realistic measure based on the network connectivity is proposed to evaluate the network controllability robustness. Under this new measure, if a network is severely destructed by node- or edge-removal, then the attack process will stop or its network controllability will not be measured any further.

3) The conventional connectivity robustness of the proposed network models are also compared. The simulation results reveal that the multi-loop structures, which are beneficial to controllability robustness, are not beneficial to connectivity robustness.

The rest of the paper is organized as follows: Section II introduces some preliminaries, including some basic concepts and definitions. Section III elaborates in detail the various mechanisms of Henneberg-growth in complex networks. Section IV demonstrates simulation results with analysis and discussions. Finally, Section V concludes the investigation.

II. PRELIMINARIES

A general linear time-invariant (LTI) system is described by

$$\dot{x} = Ax + Bu,$$

where $A$ and $B$ are constant matrices of compatible dimensions, $x$ is the state vector, $u$ is the control input, and $N$ is the dimension of $A$. Conceptually, this LTI system is state controllable if and only if there exists a control input $u$ that can drive the state $x$ from any initial state to any target state in the state space in finite time. As a criterion, the LTI system is state controllable if and only if the controllability matrix $[BAA^2B \cdots A^{N-1}B]$ has a full row-rank [20].

The concept of structural controllability is a slight generalization of the state controllability, to deal with two parameterized matrices $A$ and $B$, in which the parameters characterize the structure of the underlying system: if there are specific parameter values that can ensure the system to be state controllable, then the system is structurally controllable.

A. CONTROLLABILITY ROBUSTNESS

Clearly, without control input $u$, or $B \equiv 0$, the LTI system is by no means controllable. When considering a network of many LTI systems, the node system with control input is called a deriver node. The network controllability can be measured by the density of driver nodes, $n_D$, defined by

$$n_D \equiv \frac{N_D}{N},$$

where $N_D$ is the minimum number of driver nodes needed to retain a full control of the network, which can be calculated using either the minimum inputs theorem (MIT) [3] (for directed networks), or the exact controllability theorem (ECT) [4] (for both directed and undirected networks), defined as follows:

$$n_D = \begin{cases} \frac{1}{N} \cdot \max\{1, N - |E^*|\}, & \text{using MIT [3]}, \\ \frac{1}{N} \cdot \max\{1, N - \text{rank}(A)\}, & \text{using ECT [4]}, \end{cases}$$

where $E^*$ is the set of the minimum number of edges that need to be removed to make the network state controllable.
where \(|E^*|\) is the number of edges in the maximum matching \(E^*\). Under node-removal attacks, the controllability robustness is measured by

\[ n_D(i) \equiv \frac{N_D(i)}{N - i}, \quad i = 0, 1, \ldots, N - 1, \quad (3) \]

where \(N_D(i)\) is the number of driver nodes needed to retain the network controllability after a total of \(i\) nodes have been removed. This measure plots a controllability curve for a network under attack, where the horizontal axis is the number (or proportion) of the removed nodes and the vertical axis is the corresponding density of driver nodes.

A commonly used overall controllability robustness measure is calculated by \(R = \frac{1}{T} \sum_{i=0}^{N-1} n_D(i)\), which counts the entire process from attacking the first node to the last node. This measure is not practical, however, since in the later stage of an attack, the networked system has been ‘severely’ or ‘sufficiently’ destructed, such that it already lost its functionality as a ‘network’ or ‘system’. Thus, measuring the controllability for such a destructed network is no longer meaningful. An example is given in Figs. 1 (d), (e), and (f), where the network has been severely destructed and lost its original functionalities such as controllability and even connectivity.

As a remedy, a more realistic measure of controllability robustness is proposed here, as follows:

\[ R_T = \frac{1}{T + 1} \sum_{i=0}^{T} n_D(i), \quad (4) \]

where \(T (T < N)\) represents the number of node-removals when the threshold of ‘sufficient destruction’ is reached. Here, \(T\) separates the attack process into two parts: before \(T\) is reached, the network is considered as normal, and after \(T\) is reached, the network is deemed breakdown. The controllability robustness performance of a network will be measured only before this threshold is reached.

Figure 1 shows an illustrative example of node-removal attacks. The notations ‘\#node’, ‘\#driver’, ‘\#lcc’ and ‘\#dc’ denote the number of nodes, the minimum number of driver nodes needed, the size of the largest connected component (LCC) [8] and the number of disconnected components, respectively. It can be observed that \#node and \#lcc are monotonically non-increasing, while \#dc is increasing in the early-stage but decreasing in the later-stage. For a connected network, successive node-removals will result in disconnection and the number of disconnected components will increase. However, if a network has been severely disconnected into many components, further node-removals will only delete small components (e.g., isolated nodes) and thus the number of disconnected components will decrease gradually.

The above observation suggests that there is a peak in the curve of \#dc, which can be used to indicate whether or not the network is severely destructed. Specifically, in this paper, the threshold \(T\) is set to be the number of node-removals when the number of disconnected components begins to decrease, namely the turning point of the curve of the disconnected component numbers.

Let \(D(i)\), where \(1 \leq D(i) \leq N\) for \(i = 0, 1, \ldots, N - 1\), denote the number of disconnected components, when a total number of \(i\) nodes have been removed from the network. The turning point of \(D(i)\) is also the maximum value of \(D(i)\), namely, \(T = \arg\max_{i} D(i)\). To eliminate the influence of randomness, the threshold will be averaged over multiple independent attack simulations.

Given the measure \(R_T\) as presented in Eq. (4), which is used as the measure for controllability robustness in this paper, the smaller the \(R_T\) value in a network, the better the controllability robustness the network.

**B. CONNECTIVITY ROBUSTNESS**

Connectivity robustness refers to the ability of a complex network maintaining its connectivity against destructive node- or edge-removal attacks, which is typically measured by the normalized largest connected component (LCC) [8]. The LCC of a directed network is the largest weakly connected sub-network, where a directed graph is weakly connected if its underlying graph is connected, namely, it remains to be connected after all the directed edges are changed to be undirected.

Taking the threshold \(T\) into account, the network connectivity robustness measure is calculated as follows:

\[ Q_T = \frac{1}{T + 1} \sum_{i=0}^{T} q(i), \quad (5) \]

where \(q(i) = \frac{N_{LCC(i)}}{N - i}\) represents the proportion of nodes in the LCC, namely the number of nodes \(N_{LCC(i)}\) in the LCC divided by the current network size \(N - i\), when a total of \(i\) nodes have been removed from the original network.

**III. THE HENNEBERG-GROWTH OF COMPLEX NETWORKS**

The Henneberg-growth process in social network modeling consists of two steps: 1) the construction of triangular acquaintance by a newcomer and a random pair of two friends that are already in the community; 2) the random elimination of an existing acquaintance in the community. As a graph, the first step forms undirected triangles by adding edges, while the second step randomly removes edges. In this paper, the Henneberg-growth process is slightly extended: the undirected edges are changed to be directed; and the newcomers join the network in the form of a small group, with 1, 2, 3, or 4 individuals, respectively. Clearly, the directed edges make the network connections more complicated.

**A. GENERATION PROCESS**

In this paper, four random polygon network models are proposed. Fig. 2 (a) shows an illustrative example of RT, where a newcomer is an individual node, which will be connected to two existing adjacent nodes in the network; thus, each
FIGURE 1. An example of node-removal attacks, where ‘#node’ represents the number of nodes; ‘#driver’ represents the minimum number of driver nodes needed; ‘#lcc’ represents the size of the largest connected component; and ‘#dc’ represents the number of disconnected components.

(a) #node=8
   #driver=3
   #lcc=8
   #dc=1
(b) #node=7
   #driver=3
   #lcc=3
   #dc=3
(c) #node=6
   #driver=3
   #lcc=3
   #dc=3
(d) #node=5
   #driver=4
   #lcc=2
   #dc=4
(e) #node=4
   #driver=3
   #lcc=2
   #dc=3
(f) #node=1
   #driver=1
   #lcc=1
   #dc=1

FIGURE 2. [color online] Example of Henneberg-growth: (a) RT: newcomers as individuals (1-cliques); (b) RR: newcomers as 2-cliques (or 2-chains); (c) RP: newcomers as 3-chains; (d) two variants of RP: newcomers as looping 3-cliques (left) and detouring 3-cliques (right); (e) RH: newcomers as 4-chains; (f) different types of formed polygons. Blue nodes and edges represent newcomers and new connections, respectively; black nodes and edges represent the existing nodes and edges in the network.

(a) individual newcomer: random triangle network
(b) pair-wise newcomers: random rectangle network
(c) group-wise newcomers: random pentagon network
(d) random pentagon network variants: loop and detour
(e) group-wise newcomers: random hexagon network
(f) types of directed polygons

In a nutshell, a type-I topology forms a polygon loop, while a type-II topology breaks the loop by a reversed edge. In the time, a triangle will be formed. In Fig. 2 (b), a pair of newcomers are connected to two existing adjacent nodes in the network; thus, a rectangle will be formed. Similarly, random pentagons and hexagons are generated, where newcomers join the network in the form of as a 3-node chain and a 4-node chain, as shown in Figs. 2 (c) and (e), respectively. It is worth mentioning that for RP, the newcomers can be also organized as a 3-clique (other than a 3-node chain), as shown in Fig. 2 (d); while for RH, the newcomers are organized only as a 4-chain, since a directed 4-clique has as many as 729 possible configurations or 24 acyclic configurations [21].

Different types of directed polygons formed as above are illustrated in Fig. 2 (f). If an individual newcomer and two existing nodes together form a type-I triangle (loop), then it is denoted as RT [15]; otherwise, if they form a type-II triangle (detour), the resultant network is denoted as RT-II. Similarly, if a group of newcomers and two existing nodes form a type-I rectangle, pentagon, or hexagon, then the resultant network is named RR, RP, and RH, respectively.

Given a re-directed edge that breaks the original directed loop, the corresponding polygon is denoted as type-II rectangle, pentagon, and hexagon, respectively. The corresponding resultant networks are named RR-II, RP-II, and RH-II, respectively.

In a nutshell, a type-I topology forms a polygon loop, while a type-II topology breaks the loop by a reversed edge.
case where newcomers are organized as 3-cliques, as shown in Fig. 2 (d), and if the 3-clique is a loop, then it is named RP-L; otherwise, if it is a detour then it is named RP-D.

The random polygon network generation process will first construct a number of polygons such that all nodes are connected. To exactly control the number of edges, if the summation number of edges has not reached the predetermined number, a similar generation mechanism will be applied, namely, for a random \( q \)-side polygon network, \( \alpha \) non-adjacent nodes are randomly picked from the network as if they were newcomers. These \( \alpha \) nodes will be internally connected as a group and then connected to another pair of adjacent nodes, forming a \( q \)-side polygon, where \( q = \alpha + 2 \). This process repeats until the number of edges is greater than or equal to the predetermined number. On the other hand, if the summation number of edges has exceeded the predetermined number, then a random edge-removal is performed repeatedly, until the number of edges reach the stop criterion.

Algorithm 1 Random \( q \)-Size Polygon Network Generation

\[\begin{align*}
\text{Input} & : \text{number of nodes} N; \text{number of edges} M; q; e \\
\text{Output} & : \text{resultant network} G \\
1 & G \leftarrow 2\text{-clique} ; \\
2 & n \leftarrow 2; // \text{temporal number of nodes} \\
3 & m \leftarrow 1; // \text{temporal number of edges} \\
4 & \alpha \leftarrow q - 2; \\
5 & \text{while } n < N \text{ do} \\
6 & \quad \text{randomly choose a } 2\text{-clique in the network} ; \\
7 & \quad \text{connect } \alpha \text{ newcomers as a group using } e \text{ edges} ; \\
8 & \quad \text{connect } \alpha \text{-newcomer group to the } 2\text{-clique, forming } \\
9 & \quad \text{a } q\text{-size polygon} ; \\
10 & \quad n \leftarrow n + \alpha ; \\
11 & \quad m \leftarrow m + e + 2 ; \\
12 & \text{end} \\
13 & \text{while } m < M \text{ do} \\
14 & \quad \text{randomly choose a } 2\text{-clique in the network} ; \\
15 & \quad \text{randomly choose } \alpha \text{ non-adjacent nodes in the } \\
16 & \quad \text{network}; // \text{as if they were } \\
17 & \quad \text{newcomers} \\
18 & \quad \text{connect the } \alpha \text{ nodes as a group using } e \text{ edges} ; \\
19 & \quad \text{connect the } \alpha\text{-group to the } 2\text{-clique} ; \\
20 & \quad m \leftarrow m + e + 2 ; \\
21 & \text{end} \\
22 & \text{while } m > M \text{ do} \\
23 & \quad \text{randomly pick an edge and remove it} ; \\
24 & \quad m \leftarrow m - 1 ; \\
25 & \text{end} \\
\end{align*}\]

Algorithm 1 shows the pseudocode of the Henneberg-growth network generation, where a \( q \)-side polygon is generated in each time step, of which the complexity is \( O(N) \). Source codes of this work are available for the public.\(^1\)

\(^1\)https://fylou.github.io/sourcecode.html

![FIGURE 3.](image)

**FIGURE 3.** [color online] Out-degree distributions of ER, RT, and RH networks with different average degrees \((N = 1000)\).

![FIGURE 4.](image)

**FIGURE 4.** [color online] Max out-degree of ER, RT, RR, RP, and RH networks with different average degrees \((\langle k \rangle = 4, 6, 8, \text{ and } 10)\) and network sizes \((N = 1000, 600, \text{ and } 1600)\).

The generation starts from a pair of connected nodes (a 2-clique). In the main generation part from Lines 5–11, \( \alpha (\alpha = q - 2) \) newcomers are introduced and connected to the network, where \( \alpha = 1 \) for RT, \( \alpha = 2 \) for RR, \( \alpha = 3 \) for RP, and \( \alpha = 4 \) for RH, respectively. The isolated newcomers are connected by \( e \) edges, and then connected to the 2-clique that exists in the network.

Algorithm 1 will generate a network with exact \( N \) nodes and \( M \) edges. If the temporal summation number of edges \( m < M \), then \( q \) non-adjacent nodes are randomly selected as if they were the newcomers, which are connected using \( e \) internal edges, and then they are connected to another 2-clique, as shown in Lines 12–18.

Since each step will generate \( e + 2 \) edges, edge removals are usually necessary, where a random edge will be removed at each time step, as shown in Lines 19–22.

### B. BASIC FEATURES

The family of random polygon networks are homogeneous, with out- and in-degrees following Poisson distributions. As shown in Fig. 3, when the average degree is low \((\langle k \rangle = 3)\), RT and RH possess slightly more concentrated degree distributions around the average degree as compared to the Erdös-Rényi (ER) random-graph [22]. However, when the average degree increases, the shapes of distributions of these three networks tend to be similar to each other.

Although following Poisson distributions, RT has a greater maximum out-degree than other random polygon networks.
FIGURE 5. [color online] Network assortativity, algebraic connectivity, and heterogeneity of out-degrees of RT, RR, RP, and RH, with different average degrees ($\langle k \rangle = 4, 6, 8, \text{ and } 10$) and network sizes ($N = 600, 800, 1000, 1200, 1400, \text{ and } 1600$).

FIGURE 6. [color online] Initial controllability ($n_D(0)$) of RT, RR, RP, RH, ER, and SF, with different average degrees ($\langle k \rangle = 4, 6, 8, \text{ and } 10$) and network sizes ($N = 600, 800, 1000, 1200, 1400, \text{ and } 1600$).

FIGURE 7. [color online] Controllability curves of ER, SF, RT, and RH under different attack strategies. A dashed line represents the curve turning point of the number of disconnected components for the corresponding network in the same color.

and ER, as shown in Fig. 4, while all the random polygon networks except RT possess similar maximum out-degrees as ER. Note that degree distribution is not the key issue that influences the controllability robustness [13].

The algebraic connectivity (AC), assortativity (AS), and heterogeneity of out-degrees (HO) values of RT, RR, RP, and RH with different average degrees ($\langle k \rangle = 4, 6, 8, \text{ and } 10$) and network sizes ($N = 600, 800, 1000, 1200, 1400, \text{ and } 1600$) are shown in Fig. 5. As can be seen from the figures, these values are generally not influenced by the network size changes. Figures 5 (a–d), (e–h), and (i–l) show that AC increases, but AS and HO decrease, as the average degree increases.

Figure 6 shows the initial controllability $n_D(0)$ (see Eq. (3)) of RT, RR, RP, RH, ER, and generic scale-free (SF) networks [23], with different average degrees and network sizes. A lower value of $n_D(0)$ represents better initial controllability. It can be seen from Figs. 6 (a) and (e) that RT and ER (Erdős-Rényi random graph [22]) possess similar initial controllability performances. Given a fixed value of average degree, RR, RP, RH possess a common feature that their $n_D(0)$ values decrease as the network size increases.
In contrast, ER and RT trend to keep the same $n_D(0)$ values, while for SF, its $n_D(0)$ value decreases as the network size increases. This implies that for ER and RT, the (initial) controllability will remain the same as the network size scales, as long as the average degree remains the same; while for RR, RP, and RH, the (initial) controllability will be enhanced as the network size increases. This is because the required number of driver nodes is very low (e.g., only one driver node is needed to ensure the controllability of the entire network), and the increase of the denominator in Eq. (3) will decrease the density of the driver nodes. As for SF, given the same average degree, as the network size increases, the (initial) controllability degenerates due to its heterogeneity.

IV. EXPERIMENTAL STUDIES

This paper investigates the controllability robustness of different random polygon networks generated by the Henneberg-growth mechanism, namely RT (with variant RT-II), RR (with variant RR-II), RP (with variants RP-II, RP-L, and RP-D), RH (with variant RH-II). To test the scalability as the network size changes, the number of nodes $N$ is set to 500, 1000, and 2000, respectively. For each $N$, the average degree $\langle k \rangle$ is 3, 5, and 10, respectively. Note that there are two different settings of network sizes in simulations: 1) the above settings for comparisons of controllability robustness (as well as connectivity robustness); and 2) for the basic feature studies, the network size is set to $N = 600, 800, 1000, 1200, 1400$, and $1600$, with different average degrees $\langle k \rangle = 4, 6, 8$, and 10, respectively. The different settings further confirm the scalability. In simulations, since the attack simulations are very time-consuming for large-scaled networks, it is only applied to networks with $N = 500$, while the other three attack strategies are applied to all topologies with different sizes.

The experimental studies are organized as follows: Subsection IV-A experimentally verifies the effectiveness of using the threshold $T = \arg \max_i D(i)$ as the indicator of network destruction. In Subsection IV-B, the controllability robustness of type-I polygon networks (namely, RT, RR, RP, and RH) is investigated, together with ER and scale-free (SF) networks [23], [25], [26]. Small-world (SW) networks [27] are not compared since SW networks are generated from global loop structures, while ER, SF, and the random polygon networks are generated based on local structures. They have very different generation mechanisms and, therefore, are not comparable. Subsection IV-C compares the corresponding type-I and type-II networks, in a neck-to-neck manner. In Subsection IV-D, the scenarios of organizing newcomers as different $\beta$-cliques ($\beta = 1, 2$, or $3$) are investigated. Subsection IV-E compares the initial critical edges in different topologies. Finally, in Subsection IV-F, the connectivity robustness performances of these networks are presented and briefly discussed, using the measure (5).

To eliminate the effect of randomness, all data presented in this section are averaged from 30 repeatedly independent simulations.

A. THRESHOLD

Figure 7 shows the controllability curves of ER, SF, RT, and RH under different node-removal attack strategies. In each subplot, there are vertical dashed lines, each representing the turning point of the number of disconnected components for the same colored network. Figures 7 (a–e) show the controllability curves under TBA; Fig. 7 (f) shows the curves under CRI; Figs. 7 (g–k) show the curves under TDA; and Fig. 7 (l) shows curves under RND. It can be observed that the vertical dashed lines are very close to the turning points of the controllability curves under TDA and TBA. As for
TABLE 1. Neck-to-neck controllability robustness comparison of type-I and type-II networks under TBA, TDA, RND, and CRI. A ‘+’ represents that the type-I network outperforms its corresponding type-II network; while a ‘–’ represents that the type-II network outperforms its corresponding type-I network.

<table>
<thead>
<tr>
<th>N=1000</th>
<th>RT</th>
<th>RT-II</th>
<th>RR</th>
<th>RR-II</th>
<th>RP</th>
<th>RP-II</th>
<th>RH</th>
<th>RH-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBA</td>
<td>(k) = 2</td>
<td>0.396</td>
<td>0.477</td>
<td>+</td>
<td>0.325</td>
<td>0.367</td>
<td>+</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.301</td>
<td>0.361</td>
<td>+</td>
<td>0.324</td>
<td>0.266</td>
<td>+</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.192</td>
<td>0.298</td>
<td>+</td>
<td>0.148</td>
<td>0.169</td>
<td>+</td>
<td>0.145</td>
</tr>
<tr>
<td>TDA</td>
<td>(k) = 2</td>
<td>0.468</td>
<td>0.510</td>
<td>+</td>
<td>0.388</td>
<td>0.429</td>
<td>+</td>
<td>0.383</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.333</td>
<td>0.364</td>
<td>+</td>
<td>0.283</td>
<td>0.301</td>
<td>+</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.185</td>
<td>0.190</td>
<td>+</td>
<td>0.171</td>
<td>0.173</td>
<td>+</td>
<td>0.162</td>
</tr>
<tr>
<td>RND</td>
<td>(k) = 2</td>
<td>0.300</td>
<td>0.399</td>
<td>+</td>
<td>0.256</td>
<td>0.299</td>
<td>+</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.187</td>
<td>0.251</td>
<td>+</td>
<td>0.163</td>
<td>0.188</td>
<td>+</td>
<td>0.160</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.092</td>
<td>0.105</td>
<td>+</td>
<td>0.082</td>
<td>0.091</td>
<td>+</td>
<td>0.085</td>
</tr>
<tr>
<td>TBA</td>
<td>(k) = 2</td>
<td>0.400</td>
<td>0.480</td>
<td>+</td>
<td>0.333</td>
<td>0.372</td>
<td>+</td>
<td>0.301</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.317</td>
<td>0.347</td>
<td>+</td>
<td>0.236</td>
<td>0.267</td>
<td>+</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.197</td>
<td>0.217</td>
<td>+</td>
<td>0.145</td>
<td>0.174</td>
<td>+</td>
<td>0.146</td>
</tr>
<tr>
<td>TDA</td>
<td>(k) = 2</td>
<td>0.465</td>
<td>0.505</td>
<td>+</td>
<td>0.396</td>
<td>0.428</td>
<td>+</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.320</td>
<td>0.358</td>
<td>+</td>
<td>0.289</td>
<td>0.306</td>
<td>+</td>
<td>0.281</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.179</td>
<td>0.193</td>
<td>+</td>
<td>0.167</td>
<td>0.170</td>
<td>+</td>
<td>0.166</td>
</tr>
<tr>
<td>RND</td>
<td>(k) = 2</td>
<td>0.300</td>
<td>0.335</td>
<td>+</td>
<td>0.256</td>
<td>0.302</td>
<td>+</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.178</td>
<td>0.250</td>
<td>+</td>
<td>0.162</td>
<td>0.184</td>
<td>+</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.088</td>
<td>0.102</td>
<td>+</td>
<td>0.079</td>
<td>0.087</td>
<td>+</td>
<td>0.082</td>
</tr>
<tr>
<td>CRI</td>
<td>(k) = 2</td>
<td>0.491</td>
<td>0.504</td>
<td>+</td>
<td>0.411</td>
<td>0.453</td>
<td>+</td>
<td>0.397</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.317</td>
<td>0.377</td>
<td>+</td>
<td>0.271</td>
<td>0.305</td>
<td>+</td>
<td>0.269</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.146</td>
<td>0.178</td>
<td>+</td>
<td>0.115</td>
<td>0.149</td>
<td>+</td>
<td>0.131</td>
</tr>
<tr>
<td>TBA</td>
<td>(k) = 2</td>
<td>0.491</td>
<td>0.547</td>
<td>+</td>
<td>0.325</td>
<td>0.366</td>
<td>+</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.291</td>
<td>0.354</td>
<td>+</td>
<td>0.233</td>
<td>0.288</td>
<td>+</td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.195</td>
<td>0.208</td>
<td>+</td>
<td>0.147</td>
<td>0.164</td>
<td>+</td>
<td>0.142</td>
</tr>
<tr>
<td>TDA</td>
<td>(k) = 2</td>
<td>0.474</td>
<td>0.517</td>
<td>+</td>
<td>0.395</td>
<td>0.433</td>
<td>+</td>
<td>0.380</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.335</td>
<td>0.366</td>
<td>+</td>
<td>0.295</td>
<td>0.310</td>
<td>+</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.188</td>
<td>0.194</td>
<td>+</td>
<td>0.173</td>
<td>0.176</td>
<td>+</td>
<td>0.172</td>
</tr>
<tr>
<td>RND</td>
<td>(k) = 2</td>
<td>0.305</td>
<td>0.398</td>
<td>+</td>
<td>0.258</td>
<td>0.304</td>
<td>+</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>(k) = 5</td>
<td>0.181</td>
<td>0.249</td>
<td>+</td>
<td>0.162</td>
<td>0.183</td>
<td>+</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>(k) = 10</td>
<td>0.088</td>
<td>0.107</td>
<td>+</td>
<td>0.085</td>
<td>0.090</td>
<td>+</td>
<td>0.083</td>
</tr>
</tbody>
</table>

* + in total: 30, 30, 30, 28

* – in total: 0, 0, 0, 2

CRI, a similar phenomenon happens only to SF networks, while for other networks there are no clear turning points in their controllability curves. Due to the nature of RND, the controllability curves of all networks under RA have no turning points.

As shown in Fig. 7, the curve turning point of the number of disconnected components shows a strong correlation to the network controllability under the targeted node-removal attacks. This kind of turning point has a clear physical meaning that, when the number of disconnected components of a network begins to decrease, it implies that the network cannot be further disconnected, consequently all further attacks will only clear some disconnected components such as isolated nodes. Therefore, the number of disconnected components naturally forms a good indicator for determining the threshold of the level of network destruction.

FIGURE 9. [color online] Controllability robustness comparison of RP, RP-II, RP-L and RP-D, with different average degrees (⟨k⟩ = 3, 5, and 10) and network sizes (N = 1000, 500, and 2000), under different attack strategies (TBA, TDA, RND, and CRI).
It is observed from Fig. 8 that polygons with greater numbers of sides are more beneficial than small-sized loops. Next, the directions of edges within polygons are investigated. As shown in Fig. 2, type-I topologies consist of many loops. The overall controllability robustness performances of the type-I polygon networks (including RT, RR, RP, and RH), ER, and SF are compared, as shown in Fig. 8, using the measure (4). The results can be summarized as follows: 1) it is clear that the homogeneous topologies (including ER, RT, RR, RP, and RH) possess significantly better controllability robustness than SF with a heterogeneous topology. The performance differences among ER, RT, RR, RP, and RH are not significant. 2) For sparse networks (⟨k⟩ = 3), as the number of sides of a polygon increases, the controllability robustness is enhanced. This suggests that bigger-sized multi-loops (such as hexagons) are more beneficial to controllability robustness than smaller-sized multi-loops (such as triangles). Here, a polygon with a greater number of sides is referred to as a ‘bigger-sized loop’, and vice versa. 3) For each topology, its controllability robustness is enhanced as its average degree increases; meanwhile, the performance differences among the homogeneous networks become smaller. For dense networks (⟨k⟩ = 10), there is no clear performance difference among the homogeneous networks under random attacks (see Figs. 8 (c), (f), and (j)).

Redirecting one edge in a type-I network results in a corresponding type-II network, where the polygon structure is kept the same as the type-I network, except that the original directed loop is now destroyed.

Table 1 shows the neck-to-neck controllability robustness comparison of type-I and type-II networks. A ‘+’ sign represents that the type-I network outperforms its corresponding type-II network; while a ‘−’ represents that the type-II outperforms its corresponding type-I network. According to the total numbers of ‘+’ and ‘−’ signs, it is clear that type-I networks are better than type-II networks, in terms of controllability robustness.

In overall 30 × 4 neck-to-neck comparisons, there are only two cases where type-II networks outperform their corresponding type-I networks, namely RH-II (with N = 1000, ⟨k⟩ = 10) and RH-II (with N = 2000, ⟨k⟩ = 10), win their corresponding RH networks under TDA, by 0.001 and 0.002, respectively. This observation is consistent with the previous investigations where multi-loop structures are beneficial to enhance the network controllability robustness.

D. NEWCOMERS AS CLIQUES

One distinguished extension of the Henneberg-growth model developed in this paper is that the newcomers join the community (network) in the form of a small group, rather than individually. In the following, the scenarios of newcomers as β-cliques, β = 1, 2, or 3, are compared.

For β = 1 and 2, the resultant networks are RT and RR, respectively. The case of β = 3 leads to two variants of RP networks, the first one is that the three newcomers are internally connected as a loop, which results in the RP-L (‘L’ for loop) structure; and the second one is the case that the newcomers are internally connected as a detours, denoted by RP-D (‘D’ for detour). Illustrative examples of RP-L and RP-D are shown in Figs. 2 (d) and (f).

Figure 9 shows the controllability robustness comparison of RP, RP-II, RP-L, and RP-D. For targeted attacks, including TBA, TDA, and CRI, as shown in Figs. 9 (a, d, h), (b, e, i), and (g), respectively, RP-L outperforms the other topologies when the average degree is low (⟨k⟩ = 3). While for random attacks, it is observed that RP-II performs worse than the others, as shown in Figs. 9 (b) and (j). Similarly, as the networks become denser, the differences of controllability robustness become smaller. An obvious conclusion form Fig. 9 is that newcomers as loop 3-cliques are better than that as detour 3-cliques, regarding the controllability robustness.

E. PROPORTION OF CRITICAL EDGES

An edge is said to be critical if its removal will increase the required number of driver nodes by one [3]. Given the same network size and average degree, a lower proportion of critical edges will ensure a better controllability robustness in the early stage of an attack [24].

Figure 10 shows the comparison of proportions of critical edges in the networks. It is clear that if the average degree is as

\[
\langle k \rangle = 2, \quad \langle k \rangle = 3, \quad \langle k \rangle = 4, \quad \langle k \rangle = 5, \quad \langle k \rangle = 6.
\]
sparse as \(k\) = 2, type-I networks (including RT, RR, RP, and RH) expose a higher proportion of critical edges than type-II networks in the early attack stage. However, as the average degree increases, the proportion of critical edges in type-I networks decreases faster than type-II networks, implying a significant enhancement of controllability robustness in the early attack stage. Note that the proportion of critical edges may significantly change due to different attack strategies.

F. CONNECTIVITY ROBUSTNESS OF RANDOM POLYGON NETWORKS

Figure 11 shows the connectivity robustness comparisons of ER, SF, RT, RR, RP, and RH, with different average degrees \((k) = 3, 5, \text{ and } 10\) and network sizes \(N = 1000, 500, \text{ and } 2000\), under different attack strategies (TBA, TDA, RND, and CRI).

TABLE 2. Neck-to-neck connectivity robustness comparison of type-I and type-II networks under TBA, TDA, RND, and CRI. A ‘+’ represents that the type-I network outperforms its corresponding type-II network; while a ‘–’ represents that the type-II network outperforms its corresponding type-I network.
Table 2 shows the neck-to-neck connectivity robustness comparison of type-I and type-II networks under TBA, TDA, RND, and CRI. For RT networks, type-I networks perform slightly better than type-II networks, regarding the connectivity robustness under attacks. This shows that the correlation between connectivity robustness and controllability robustness is low. As the number of sides of polygons increases, type-I networks become better than type-II networks, namely RR, RP, and RH are better than RR, RP-II and RH-II, respectively. This reveals a weak positive correlation between connectivity robustness and controllability robustness.

V. CONCLUSION
In this paper it is empirically observed that the multi-loop structures are beneficial to enhancing network controllability robustness, and the Henneberg-growth mechanism in social networks offers a good approach to generating multiple loop structures in a complex network.

This paper investigates the controllability robustness of a series of random polygon networks, which are generated using an extended Henneberg-growth mechanism, where newcomers join in the form of a group, rather than join individually. A realistic measure of network controllability robustness is proposed, which filters the controllability values when a network is severely destructed. The threshold of determining the level of destruction is based the number of components of the network under a sequential of attacks.

Extensive attack simulations reveal that not only the construction of different polygons, but also the edge directions, lead to different controllability robustness performances: 1) loop polygons are more beneficial than non-loop polygons to enhance the controllability robustness. 2) Polygons with more sides are more beneficial than those with fewer sides to enhance the controllability robustness. 3) Connectivity robustness and controllability robustness have a very weak correlation with each other under various attacks, suggesting that these two features should be optimized independently.

Looking forward, it is foreseeable that the random polygon networks have potential applications not only in industrial automation, but also in biological systems and brain science, regarding its feedback control, multi-loop structures, controllability and information transmission.

REFERENCES
YANG LOU (Member, IEEE) received the Ph.D. degree from the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, in 2017. From 2017 to 2021, he was a Postdoctoral Research Fellow with the Department of Electrical Engineering and the Centre for Chaos and Complex Networks, City University of Hong Kong. He is currently a Research Assistant Professor with the Department of Computing and Decision Sciences, Lingnan University, Hong Kong. His current research interests include complex networks, evolutionary computation, and machine learning.

DONG YANG received the Ph.D. degree in electronic engineering from the City University of Hong Kong, 2019. He is currently a MCS Postdoctoral Scholar at the University of California at Merced, Merced, CA, USA. His current research has been investigating social science from different perspectives, from the mechanisms of social networks to information spreading on social networks, to predicting the relationships between users, with graph-embedding methods. He is developing new analytical techniques with complex networks and machine learning to explore some universal patterns and fundamental mechanisms in many real systems.

LIN WANG (Senior Member, IEEE) received the B.S. and M.S. degrees from the School of Mathematical Sciences, Shandong Normal University, Jinan, China, in 2003 and 2006, respectively, and the Ph.D. degree in operations research and control theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2009. She is currently a Professor with the Department of Automation, Shanghai Jiao Tong University, Shanghai, China. Her current research interests include multi-agent systems, adaptive complex networks, and coordination of multiple manipulators.

CHANG-BING TANG (Member, IEEE) received the B.S. and M.S. degrees in mathematics and applied mathematics from Zhejiang Normal University, Jinhua, China, in 2004 and 2007, respectively, and the Ph.D. degree from the Department of Electronic Engineering, Fudan University, Shanghai, China, in 2014. He is currently an Associate Professor with the Department of Electronics and Information Engineering, Zhejiang Normal University. His research interests include game theory, blockchain and its applications, networks, and distributed optimization. He was a recipient of the Academic New Artist Doctoral Post Graduate from the Ministry of Education, China, in 2012, and the Academician Pairing Training Program for Young Talents of Zhejiang Province, in 2019.

GUANRONG CHEN (Life Fellow, IEEE) received the M.Sc. degree in computer science from Sun Yat-sen University, Guangzhou, China, in 1981, and the Ph.D. degree in applied mathematics from Texas A&M University, College Station, TX, USA, in 1987. He has been a Chair Professor and the Founding Director of the Centre for Chaos and Complex Networks, City University of Hong Kong, since 2000, prior to that he was a tenured Full Professor at the University of Houston, Houston, TX, USA. He is a member of the Academy of Europe and a fellow of The World Academy of Sciences. He has been a Highly Cited Researcher in engineering according to Clarivate Web of Science, since 2009. He received the State Natural Science Award of China, in 2008, 2012, and 2016, respectively. He was awarded the 2011 Euler Gold Medal, Russia; conferred Honorary Doctorates by the Saint Petersburg State University, Russia, in 2011; and the University of Le Havre, Normandy, France, in 2014.