SNAM: A Heterogeneous Complex Networks Generation Model

Bassant E. Youssef
Bradley Department of Electrical and Computer Engineering, Virginia Tech
Email: bassant@vt.edu

Mohamed R. M. Rizk
Alexandria University
IEEE Senior Member
Email: mrmrizk@ieee.org

***Abstract*— Complex networks are found in various fields. Complex networks are characterized by having a scale-free power-law degree distribution, a small average path length (small world phenomenon), a high average clustering coefficient, and showing the emergence of community structure. Most proposed complex networks models did not incorporate all of these four statistical properties of complex networks. Additionally, models have also neglected incorporating the heterogeneous nature of network nodes. Moreover, even proposed heterogeneous complex network models were not general for different complex networks. Here, we define a new aspect of node-heterogeneity that was never previously considered which is the node connection standard heterogeneity. In this paper, we propose a generation model for heterogeneous complex networks. We introduce our novel model "settling node adaptive model" SNAM. SNAM reflects the heterogeneous nature of nodes' connection-standard requirements. Such novel nodes' connection standard criterion was not included in any previous network generation models. SNAM was successful in preserving the power law degree distribution, the small world phenomenon and the high clustering coefficient of complex networks.**

***Keywords*—Complex network modelling; BA model; preferential attachment; Heterogeneous nodes**

I. Introduction

Complex networks are ubiquitous in many areas. The Internet, the World Wide Web (WWW), social networks, food web (or food chain) networks and many other networks are complex networks [1]. Researchers studied and analyzed data extracted from complex networks which led to the discovery of their distinct features and behavioral patterns. Solid awareness of these features can lead to an improved understanding of the network's structure and dynamics. Devising a mathematical model for complex networks can aid in making decisions about complex networks management and help allocating their resources. It can be used to answer research questions such as discovering the mediator for disease transmission in sexual networks, predicting future connections between websites in the WWW, identifying critical nodes or links in power grid networks; etc. Therefore, finding a faithful mathematical model that is capable of mimicking the structure, dynamics and evolution of complex networks is paramount. Researchers used advanced computer capabilities to analyze real large databases to identify essential properties for modeling complex networks in the process of creating such mathematical model [2, 3].

Statistical properties of complex networks were identified as: small world effect, high average clustering coefficient, scale free power law degree distributions, and the emergence of community structure [3, 4]. Small world effect means that for a certain fixed value of the nodes' mean degree, the value of the average path length scales logarithmically, or slower, with network size. A node's clustering coefficient C can be defined as "the average fraction of pairs of neighbors of a node that are also neighbors of each other", where C lies between 0 and 1 [1]. The average clustering coefficients in real complex network tend to have high values. Community structure emerges when nodes in a community have denser connections within themselves than to vertices of different communities [5]. Degree distribution which is the fraction of vertices in the network with degree k follows a scale free power law distribution in real complex networks. Scale free power law distributions, $P(k) \sim k^{-\gamma}$, have a power law (PL) exponent γ independent of the size of the network and its values are in the range of $1 < \gamma < \infty$ [3, 4].

Various models tried to find a faithful model for complex networks. The most influential models in the complex-network modeling field are: Erdös and Rényi (ER), Watts and Strogatz (WS), and Barabási and Albert (BA). Networks generated according to the ER random graph model have small average path length but they have Poisson degree distributions and are characterized by having clustering coefficients lower than that found in real complex networks [3, 4]. Networks generated by WS small-world network model have a short average path length and a high clustering coefficient. However, it lacks modeling the scale free property for the networks' degree distribution [2, 3, and 4]. Thus, the scale-free power-law degree distribution of real complex networks was not represented in the ER or the WS models, rendering both models to be inaccurate in modeling the four characteristics of real complex-networks. This motivated Barbasi and Albert to induce the scale free property for node-degree distribution in their highly acclaimed model [2]. The BA model uses a Preferential Attachment (PA) connection algorithm that reflects the belief that nodes usually prefer to connect to higher-degree structurally-popular nodes [2]. BA model succeeded in preserving the PL degree distribution and small world phenomenon of real complex-networks. Networks generated by the BA model show a power-law heavy-tail degree distribution, if and only if, the model has the following two properties; growth (where new nodes are continuously added to the network) and preferential attachment (PA). The BA model starts with a small number of nodes (m_0), which is referred to as the seed network. A new node is added at

QSHINE 2014, August 18-20, RHODES, Greece

DOI 10.4108/icst.qshine.2014.256361

each time step. The new node preferentially attach to other m nodes, (where $m \leq m_0$) using a connection function based on the old nodes' normalized degrees. Thus, new node i connects preferentially to an old node j having degree Dj using a connection function (CF) based on the normalized degree of the node dj, where $dj = \frac{D_j}{\sum_j D_j}$.

Networks generated using the BA model have a scale-free power-law degree-distribution and their average path lengths exhibit the small world phenomenon. However, BA model generates networks with a constant PL exponent value of $\gamma = 3$, unlike real networks where the exponent values differ according to the network type and ranges between $1 < \gamma < \infty$. Additionally, BA modeled network average clustering coefficient is lower than that observed in real complex networks of the same size [3, 4, 5].

The BA model was still inaccurate in representing all four properties observed in real complex-networks. This motivated many researchers to introduce modifications to the BA model in an attempt to remedy the model's shortcomings. Accordingly, devising a model that can represent all four properties of complex-network is still an ongoing research [5]. Additionally, most models have assumed that nodes have the same properties and neglected incorporating the heterogeneous nature of network nodes. Moreover, even proposed heterogeneous complex network models did not integrate it with other structural properties of the network in the analysis and growth algorithms of such networks. Also, models were not general for different types of complex networks. Therefore, finding a faithful general heterogeneous complex network model that preserves real complex network statistical properties is still a challenge. In this paper we aim to devise a mathematical model that preserves the statistical properties of complex-networks. Additionally, we include a factor that, we claim, was undermined in most contemporary complex-network models which is the node heterogeneity, [4, 5]. We identify two types of node-heterogeneity; node characteristics heterogeneity and node connection standard heterogeneity. Node characteristics heterogeneity reflects the different properties or attributes that network-nodes have. Node connection standard heterogeneity reflects the difference in each node's requirements to make a connection. The contribution in this paper can be summarized as:

1) Accounting for node heterogeneity in the graph-theory by incorporating node-attributes as one of the elements defining a network graph. Accordingly, our model defines the network graph, G as a set of three elements; G= {V, E, A}, where V is the number of nodes in the network, E is the number of edges and A is the set of attributes assigned to each network node.
2) Based on (1) we propose the Settling Node Adaptive Model "SNAM" for generating complex-networks. SNAM acknowledges the heterogeneous nature of nodes by integrating the attribute-similarity with the structural popularity measure within the CFs.
3) "SNAM" departs from the BA algorithm while acknowledging the node heterogeneity. SNAM" introduces the idea of heterogeneous node connection-requirements as a criterion for connecting nodes

Our proposed models will be validated using Matlab simulation [6]. The success of each proposed model to mimic real complex networks will be verified by examining the generated network statistical properties, namely the average path length, clustering coefficient, and degree distribution.
The rest of the paper is organized as follows: section two presents the related work, section three presents our proposed models and their simulation results, and section four is the conclusion and future work.

II. Related Work

Several researchers have proposed mathematical models that address the heterogeneous nature of the nodes composing a network. The success of these models in generating networks that mimic real complex-networks was examined by observing the statistical properties of these networks. This section will review a subset of these attempts.

Bianconi and Barabási in [7] introduced the term node fitness to represent nodes' different abilities to attain connections. Their work was motivated by the observation that the nodes' abilities to attract connections do not depend only on their degrees (based on the nodes' ages). WWW nodes that provide good content are likely to acquire more connections irrespective of their ages. In citation networks, a new paper with a breakthrough is likely to have more connections than older papers. Thus each node should be assigned a parameter that describes its competitive nature to attain connections. In their model, node j upon birth is assigned a fitness factor η_j, following some distribution $\rho(\eta)$, which represents its intrinsic ability to attain connections. Bianconi and Barabási model followed the BA PA connection algorithm with a modified PA function. The model has the PA function value for connecting an old node j to a new added node i depending on the old-node degree Dj, and its fitness value η_j. When $\rho(\eta)$ follows a uniform distribution, the degree distribution is a generalized power law, with an inverse logarithmic correction. The average clustering coefficient and average path length values of networks generated by this model were not calculated in the presented work.

Shaohua et al. in [8] observed that nodes with common traits or interests tend to interact. They introduced an evolving model based on attribute-similarity between the nodes. Each of the network nodes has an attribute set. Node-attributes can be described by a true or false function as in fuzzy logic. Shaohua et al. used fuzzy similarity rules to define a similarity function between attribute sets of two nodes. A connection is established between two nodes if their attributes similarities fall within a certain sector. Despite that this model satisfies the small world property; its degree distribution does not follow a power law.

Yixiao Li et al. in [9] argued that every vertex is identified with a social identity represented by a vector whose elements represent distinctive social features. The new node added at each time step connects with probability p to the group closest to its social identity and to the other groups with probability (1-p). The higher degree node is attached to the new node within a group using PA. Random linking to neighbors of the previously attached old node is repeated until the new node establishes its m links. Their generated network follows power-law degree distribution and used triad formation to produce high average clustering coefficients. The authors claimed that using triad formation produced high average clustering but they did not present values for it and they did not measure their generated networks' average path length. Additionally, the model did not increase the length of the attribute vector to more than one.

While [6, 7, 8] based their connection algorithm on the PA attachment algorithm, some authors experimented with models that were not based on the BA PA algorithm such as those presented in [10] and [11]. Kleinberg et al. in [10] used a copying mechanism which entails randomly choosing a node then connecting its m links to neighbors of other randomly chosen nodes. The model was found to preserve power-law distributions using heuristics only. They argued that analytical tools were unable to prove this conclusion, because the copying mechanism generated dependencies between random variables. Krapivsky et al. [11] argued that an author, in a citation network, citing a paper is most likely going to cite one of its references as well. In their model, when a new node 'i' is added to the network, its edge attaches to a randomly chosen node 'j' with probability (1-r). Then with probability r this edge from the new node 'i' is redirected to the ancestor node 'o' of the previous randomly chosen node 'j'. The rate equations of the model show that it has a power-law degree distribution with degree exponent decreasing with the increase of the probability r value. Other statistical properties were not studied. These models were able to generate networks having a power-law degree distribution without using the PA algorithm of BA. However, they are not applicable to all complex-networks. Whether the node is copying its connections from a random node or connecting to the ancestor of a node previously connected to it, is not applicable for some types of complex networks. Additionally, the choice of the nodes from which the links are copied or the choice of the ancestors of the node is made randomly without regards to nodes-heterogeneous characteristics or their heterogeneous connection-standards.

Our previous paper [12] introduced the integrated attribute similarity models "IASM". IASM is a growing network model. It uses a preferential attachment algorithm to connect the nodes. The CF in IASM depends on the attribute similarity between newly arriving nodes and old network nodes as well as the structural popularity of old nodes. Two different structural popularity measures are used in IASM simulation. In IASM_A, a node's structural popularity is based on the number of connections that the node has, i.e. the node-degree, while in IASM_B, the structural popularity is based on the node's Eigen vector centrality. . IASM preserved the power law degree distribution and the small world phenomenon but it did not reflect the high average clustering coefficient and the emergence of community structure. We enhance the IASM by adding a triad formation step which results in increasing the clustering coefficient values.

III. Settling Node Adaptive Model (SNAM)

1) Introduction:

Nodes, users or entities, in real complex-networks have different profiles and characteristics. Connections between nodes affect the network dynamics, and their future evolution. We argue that nodes having different characteristics influence the density and the pattern of connections within a network. The notion of node-attributes is used to highlight the node-distinct characteristics. Attribute set is extracted from the characteristics or profiles of the network node. In our models, nodes are assigned their attributes upon their arrival to the network. Accordingly, the network graph G is now defined by a three-element set G = {V, E, A}, where V is the number of nodes in the network, E is the number of edges and A is the set of attributes defining the profiles/characteristics of all the network nodes. SNAM is a growing generation model with nodes constantly being added to the network during its evolution.

SNAM's connection algorithm uses attribute-similarity between the newly added node and the old node attribute(s) in the connection function (CF). Including the attribute similarity in CF makes it dependent on the attributes of both of the newly added node and the old node rather than having the CF dependent only on the old node's fitness/degree.). SNAM integrates the attribute-similarity between new node and old nodes with the structural popularity of old nodes in the CF. The node structural popularity is a measure of the node's popularity based on its network position and connections. SNAM uses the normalized node-degree as the structural popularity measure. SNAM departs from the classic PA connection algorithm presented in BA. SNAM reflects the idea that nodes are not only differentiated by their attributes but also according to their connection-standard requirements. Connection-standard requirements for the nodes represent the minimum CF values that a node find satisfactory to connect with other nodes.

To evaluate our models, we generate networks based on each model using MATLAB simulation. For each of the generated networks, values for the power law exponent, the average path length and the average clustering coefficients were measured and assessed against values reported for a variety of real complex-networks [3,4].

In SNAM, each new network-node upon birth possesses its own distinct attribute-set (attribute vector) that represents the interests or engagements of the node in the network's L interests or activities. The CF does not depend solely on a specific characteristic of the old node but on the characteristics of both the new and the old nodes. SNAM is a growing network model. SNAM start with a seed network of size m_o shown in figure 1. Then at each time step a new node is added with m edges to be connected to it, where $m \le m_o$. Each node is assigned an attribute vector having L elements. Each element takes binary values of 1or 0 representing the presence or absence of an attribute in the attribute-vector respectively. Our proposed attribute similarity is equal to the normalized summation of the inner product of the new-node and old-node attribute vectors. The algorithm of SNAM is shown in the flow chart in figure2.

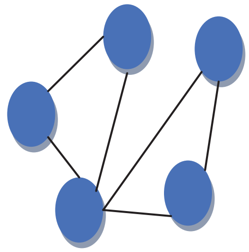


Fig1. Seed network, m_o =5

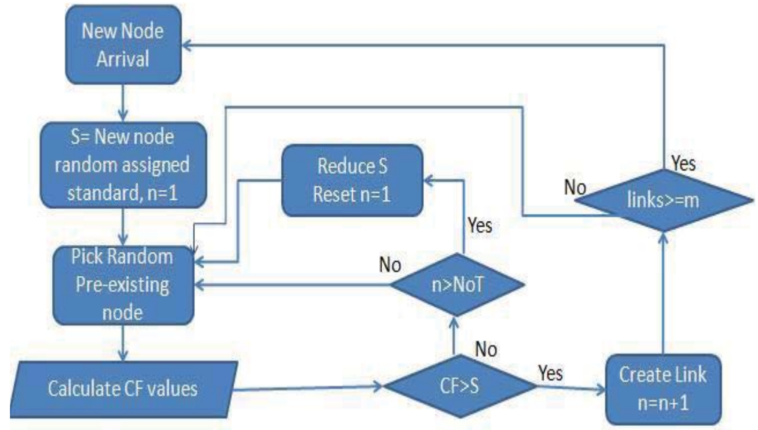


Fig2. Flow chart of SNAM algorithm

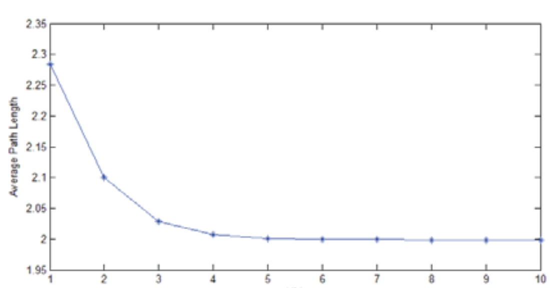
To the extent our knowledge, all previously proposed models assumed that all newly arriving nodes have the same requirements when connecting to old nodes. In reality, nodes may have different views of the same value of a connection-function (CF) calculated based on attribute similarity and/or structural popularity. A network node may have high connection standard and does not settle for the CF value offered by the tested old node, thus rejecting the connection. Another new node may have lower standards and considers the same CF value acceptable. To reflect this, we assign a characteristic that reflects the node's standards. This characteristic represents the minimum acceptable value of the CF for each node. All old pre-existing nodes whose CF values with the new node are below the newly arriving node standard will not be attached to that new node. Arriving node must then test other old pre-existing nodes to find the ones that satisfy its connection standard.

Thus, in SNAM, each arriving node, upon birth will be assigned a value representing its own connection standard value which is derived from uniform distribution. Arriving node will calculate its CF values with old nodes. Hence, the CF obtained values will not be used to deploy the preferential attachment algorithm but will be used to examine if the randomly chosen old nodes will meet the arriving node's standards. A newly arriving node will calculate the CF corresponding to random chosen nodes. The new node will establish connections with the old nodes whose CF values are equal to or higher than its connection-standard. The used CF depends on the normalized degree values and/or attribute similarity.

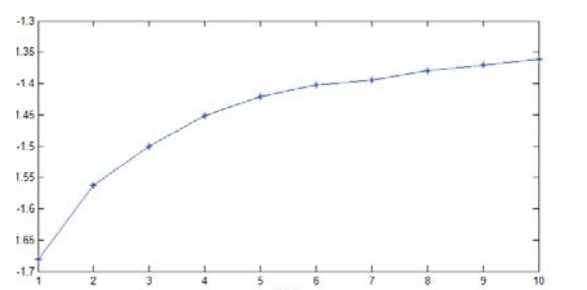
2) Simulation:

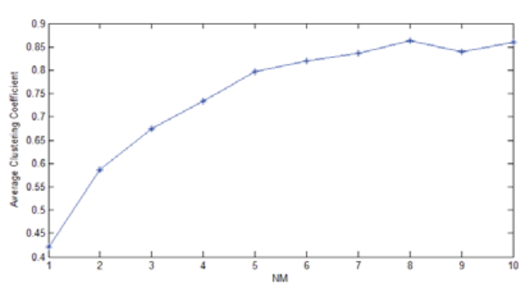
The network starts with a seed network m_o. A new node arrives at each time step and each new node 'i' is assigned a random connection-standard value "S_i", where $0 < S_i \leq 1$. If for a chosen pre-existing old node 'j' the CF value exceeds S_i then i will establish a connection to j, otherwise i rejects the connection to j and another old node is tested. This testing of other test nodes continues up to a maximum number of tests 'NoT' or until the new node achieves a predefined number of connections referred to as m. The arriving node has to lower its standard after reaching its maximum number of tests 'NoT' if it has not made m connections.

a)



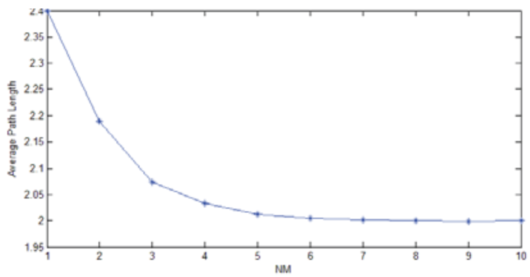
b)



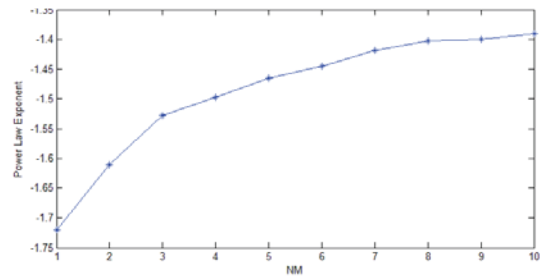


c)

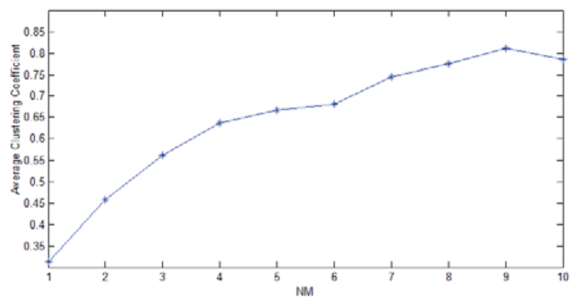
Fig. 3 SNAM algorithm with a normalized degree CF (β=1) a) Average Path length, b) Power law Exponent of Degree distribution c) Average clustering coefficients



a)

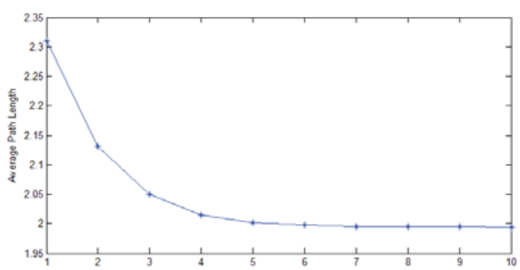


b)

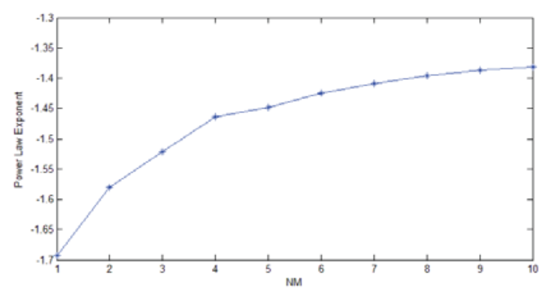


c)

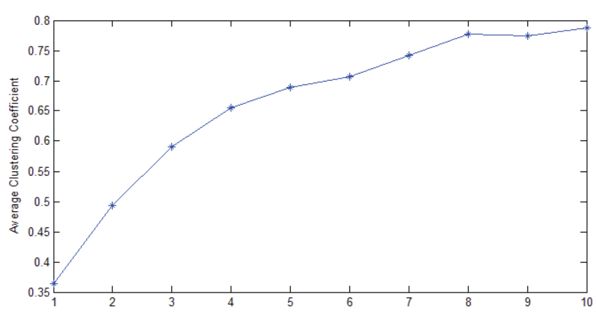
Fig. 4 SNAM algorithm with a normalized degree with added attribute similarity CF (w = β= 0.5) a) Average Path length, b) Power law Exponent of Degree distribution c) Average clustering coefficients



a)



b)



c)

Fig. 5 SNAM algorithm with a normalized degree with multiplied attribute similarity CF (α =1) a) Average Path length, b) Power law Exponent of Degree distribution c) Average clustering coefficients

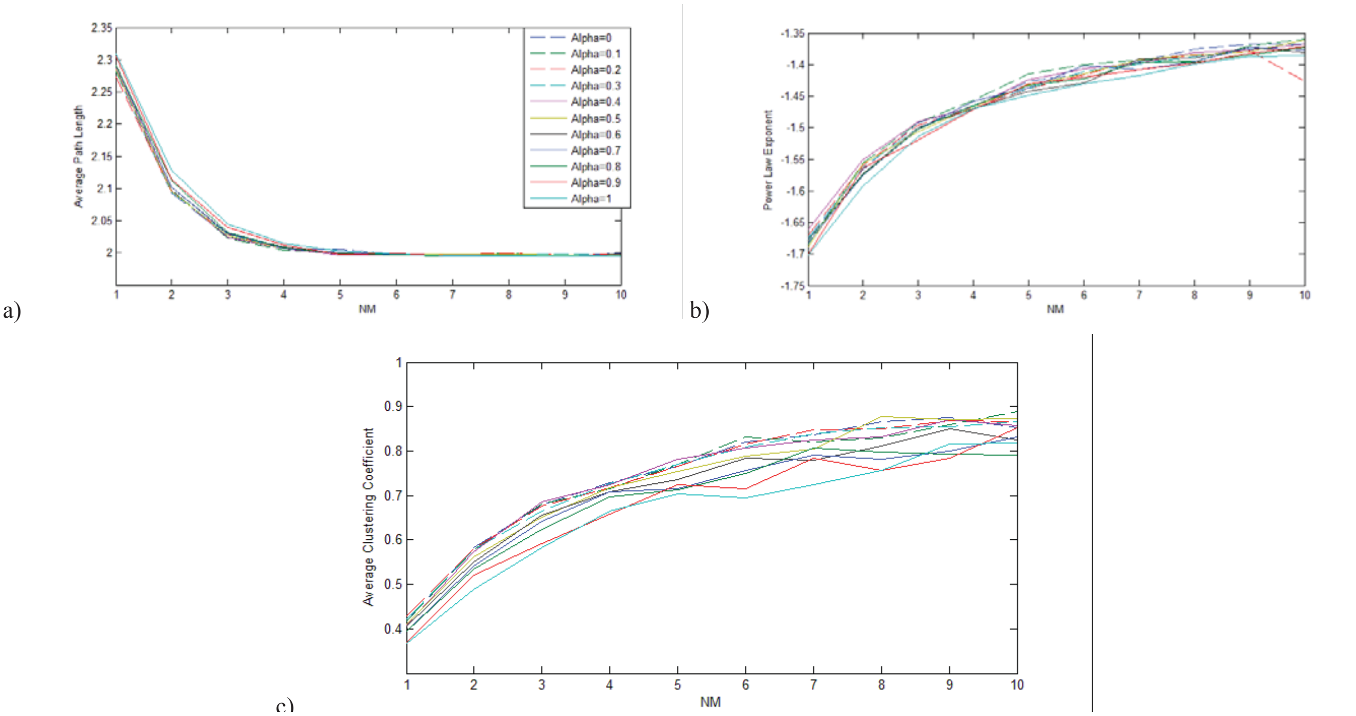


Fig.6 SNAM algorithm with varying coefficient values for the normalized degree with multiplied attribute similarity CF a) Average Path length, b) Power law Exponent of Degree distribution c) Average clustering coefficients

Notably, we can observe that unlike BA, the arriving node does not have to calculate the corresponding CF values for all pre exiting nodes when making connections. In case of SNAM the CF values corresponding to only the random chosen test nodes are calculated, Thus, the simulation time needed to generate an network by SNAM was less than that needed to generate a network of the same size by both BA and IASM[12]. In the present model, we experiment with the maximum number of tests 'NoT' required for the arriving node 'i' to lower its standard if 'i' has not already established its m connections during the "NoT" tests.

NoT is initially taken as the integer value of half the seed size m_o. NoT value is dependent on the current size (CS) of the network. NoT is increased by one whenever the CS of network reaches certain predefined milestones. The value of these milestones is dependent on the final size of the network (N) and the number of milestones (NM) occurring during network evolution, where NM ranges between 1 and N. Thus, NoT is increased by one whenever the CS of the network is increased by (N/NM) nodes. The higher the value of NM, the more rapid is the increase in 'NoT'.

Our experimentation with 'NoT' parameter indicated that rapid increase of 'NoT' with network growth resulted in the presence of irregularities in the statistical characteristics of the generated network. Here, the maximum value of NM is 10 which correspond to increasing NoT upon the arrival of 10% of the final size nodes (100 nodes). This choice was made to avoid irregular statistical properties and has proved to give satisfactory results as shown in figures 3, 4,5,and 6.

The connection function CF is dependent on the old node i degree D_i and attribute similarities (A_{ij}) for both node i and new node j , namely: $CF = \alpha * \frac{D_j A_{ij}}{\sum_j D_j A_{ij}} + \beta * \frac{D_j}{\sum_j D_j} + w * \frac{A_{ij}}{\sum_j A_{ij}}$, where $\alpha + w + \beta = 1.0$, $0 \le \alpha \le 1$, $0 \le w \le 1$, and $0 \le \beta \le 1$. . α, w, and β are the coefficients used to give different weights to the different terms of the CF to test their influence. Simulation of SNAM starts with a seed network of size $m_o = 5$. The network size grows as new nodes arrive to the network, until reaching a predetermined final size N. In our simulation N=1000. Each newly arriving node has to establish m links with the preexisting network nodes, where $m=m_o=5$. Each new node in the network is randomly assigned an attribute vector of length L =10, whose elements are derived from a uniform distribution.

Matlab[6] simulations were performed for different combinations of the CFs' coefficients for both models. The simulation results show the average of 10 experiments with different random-seed generator values. CFs used can be based on normalized degree only ($\beta=1$, $\alpha = w = 0$), on degree with added attribute similarity ($\alpha = 0$ and $w=1-\beta$ where $0 \le \beta \le 1$), and on degree multiplied by the attribute similarity. Simulation results for the Average Clustering Coefficient (Av_CC), the Average Path length (Av_Pl), and the Exponent of PL (Exp_PL) for three combinations of the coefficients α, β and w in figures 3, 4, 5. These results will be compared to values found for a network of the same size (number of edges and number of nodes) generated by Erdös and Rényi (ER), and Barabási and Albert (BA) models shown in the next table

TABLE 1: Simulation Results for ER and BA models

	Average Path length	PL exponent	Average clustering coefficient
Erdös and Rényi (ER),	3	-	0.00998
Barabási and Albert (BA)	3	2.49	0.032

Figures (3a), (4a), (5a) for the average path length indicate that small world effect is preserved for three combinations of α, β and w. Average path length decreases with the increase of NM. It is obvious that SNAM produces shorter average path length values for all NM values for the three combination of α, β and w.

Figures (3b, 4b, 5b) show that the magnitude of PL exponents for the three variations remains in the range of $1.35 \leq \gamma \leq 1.75$ which is consistent with values found in real networks [1, 3, 4]. Additionally, the magnitudes of PL exponent saturates at values close to $\gamma \cong 1.35$ with the increase of NM. The PL exponent has no value in ER generated network as the degree distribution follows a Poisson distribution. Moreover, the magnitude of the PL exponent generated by SNAM is less than that of the BA model but within values reported in [2,3]. The decrease in the PL exponent is related to the formation of hubs which will have an effect on the obtained average clustering coefficients.

The average clustering coefficient values increase with the increase of "NM" for the three variations as seen in figures (3c), (4c), and (5c). Average clustering coefficient reach high values compared to those of BA model. The clustering coefficients in figure (3c) corresponding to degree only achieves higher values than those of figure (4c) and (5c) using additive attribute similarities and multiplicative attribute similarity CF respectively. The smallest average clustering values for SNAM corresponding to $NM=1$ was 0.36 which is much higher than that of BA (0.032).

We now examine the effect that changing the values of the coefficients α and β of the CF has on the statistical properties of the generated network. Figures (6.a), (6.b), and (6.c) show the effect of varying the coefficient of the multiplicative attribute similarity term 'α' and that of the normalized degree 'β' in the CF on resulting SNAM statistical properties. Figure (6.a) shows that the higher the value of 'α' (Alpha) the higher the average path length is. However, the small world phenomenon is preserved for all 'α, β' values. The PL exponent values remain between $1.35 \leq \gamma \leq 1.75$ for all 'α, β' values as seen in figure (6.b). Figure (6.c) also shows that the average clustering coefficient increases slightly with the decrease of α.

Thus, the SNAM generation model has preserved the PL degree distribution, has a small average path length, and has high clustering coefficient values. Parameter 'NoT" value can be used to generate a variety of complex networks with specific values of the clustering coefficient, the average path length and the PL exponent.

SNAM model will be useful in studying online social networks and mimicking their structure and dynamics. SNAM excels by including connection standard which can be used to represent individual differences between users. SNAM can be used for link prediction between users in online social networks.

IV. Conclusion

This paper took into consideration that complex networks mathematical models should incorporate their statistical properties and should also reflect the heterogeneous nature of network nodes. In this paper, we propose several mathematical models that pave our path to find a final mathematical model that can successfully mimic real complex networks. The proposed models have heterogeneous network nodes with assigned distinct attributes. Our work is the first to assign more than one attribute to each node. SNAM integrates the attribute similarity measure within the CF. SNAM uses the CF values to connect the nodes. The CF depends on the old node degree simultaneously with the attribute similarity between new node and old node. SNAM is the first model that has new arriving nodes having different connection-standard requirements. SNAM proved to be very promising as it generated a network that had a PL degree distribution, small average path length and high clustering coefficient values. The effect of using Eigen vector centrality instead of degree centrality on the emergence of community structure in SNAM is still to be examined in the future work. We are also working on implementing an algorithm to SNAM that would result in the emergence of community structure. Implementing an analytical model for SNAM is also part of our future work.

References

[1] Xiao Fan Wang; Guanrong Chen; , "Complex networks: small-world, scale-free and beyond," Circuits and Systems Magazine, IEEE , vol.3, no.1, pp. 6- 20, 2003, doi: 10.1109/MCAS.2003.1228503

[2]A.-L.Barabási,R. Albert , "Emergence of scaling in random networks " Science 286, 509–512 (1999).

[3]M. E. J. Newman," The structure and function of complex networks". SIAM Review 45, 167–256 (2003)

[4]R. Albert and A. Barabasi, "Statistical mechanics of complex networks", Rev. Modern Phys., 74 (2002), pp. 47–97.

[5]Emilio Ferrara, "Mining and Analysis of Online Social Networks" Ph.D. dissertation 2012.

[6] http://www.mathworks.com/products/matlab/

[7] Bianconi, G.and Barab´asi, A.-L., "Competition and multiscaling in evolving networks", Europhys. Lett. 54, 436–442 (2001).

[8]Shaohua Tao; XiaopengYue; , "The attributes similar-degree of complex networks" ,Future Computer and Communication (ICFCC), 2010 2nd International Conference on , vol.3, no., pp.V3-531-V3-535, 21-24 May 2010, doi: 10.1109/ICFCC.2010.5497519

[9] Yixiao Li; Xiaogang Jin; Fansheng Kong; Jiming Li, "Linking via social similarity: The emergence of community structure in scale-free network", Web Society, 2009. SWS '09. 1st IEEE Symposium on , vol., no., pp.124-128, 23-24 Aug. 2009 , doi: 10.1109/SWS.2009.5271769

[10] Kleinberg, J. M., Kumar, S. R., Raghavan, P., Rajagopalan, S., and Tomkins, A., "The Web as a graph: Measurements, models and methods", in Proceedings of the International Conference on Combinatorics and Computing, no. 1627 in Lecture Notes in Computer Science, pp. 1–18, Springer, Berlin (1999).

[11] Krapivsky, P. L., and S. Redner," Organization of Growing Random Networks", 2001, Phys. Rev. E 63, 0661

[12] Youssef , B, and Hassan, H, "IASM: An Integrated Attribute Similarity for Complex Networks Generation" in Proceedings of the International Conference on Information Networking (ICOIN), 2014.