Clustering Coefficient of Incomplete Watts-Strogatz Networks

Beverly V. Gemao[a*], David Paul M. Jinayon[b]

[a] *Department of Physics, Mindanao State University – Iligan Institute of Technology A. Bonifacio Avenue, 9200 Iligan City, Philippines*
[b] *Philippine Science High School - Caraga Region Campus, Ampayon, 8600 Butuan City, Philippines*

*E-mail: beverly.gemao@g.msuiit.edu.ph

INTRODUCTION

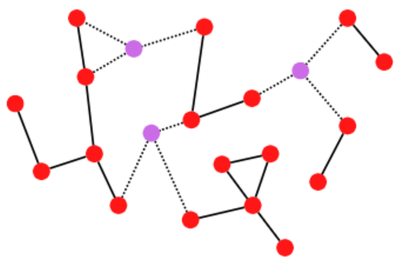

Figure 1. The fragmentation of a network when nodes are missing (pink). The dashed lines represent the missing connections due to incomplete node data.

- **In real systems, it is common occurrence that networks will have missing components.**
- **Analyzing such systems may affect our observations and inferences of its underlying structure.**

METHODOLOGY

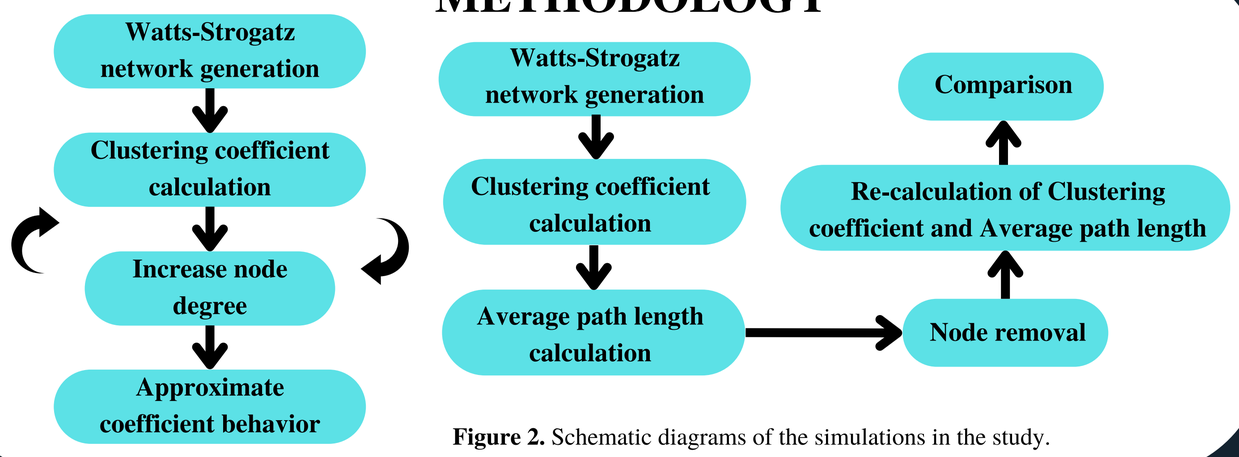

Figure 2. Schematic diagrams of the simulations in the study.

RESULTS AND DISCUSSION

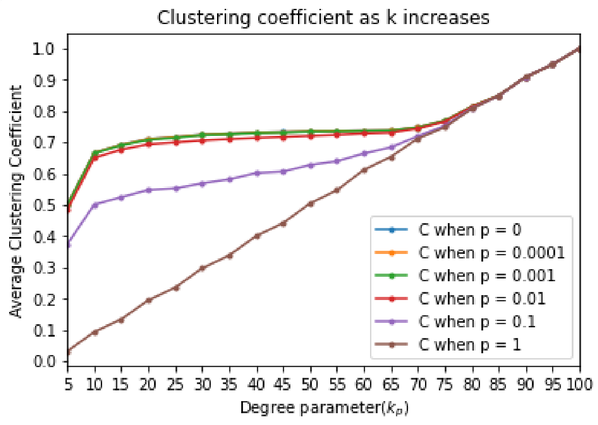

Figure 3. The behavior of the average clustering coefficient C as the initial degree k increases and with different rewiring probability p values.

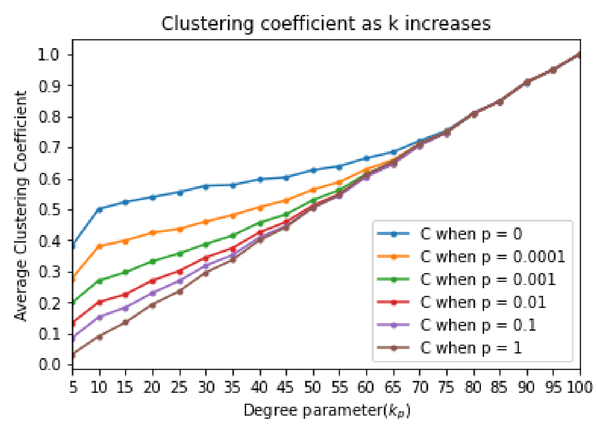

Figure 4. The behavior of the average clustering coefficient C as the initial degree k increases and with a different set of p-values.

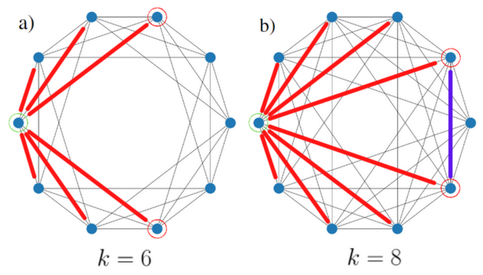

Figure 5. Two networks with $N=10$ with a degree limit of 6. (a) shows when the degree is equal to the limit and (b) shows what happens when the degree exceeds the limit. We observe that extra connections between neighbors appear in (b).

- **As the node degree k increases and the rewiring probability approaches 1, the clustering coefficient shows linear relationship with k.**
- **The clustering coefficient can be approximated by,**

$$C_i = \frac{3(k_i - 2)}{4\,(k_i - 1)}\,, \text{where } k << N$$

- **Beyond a certain limit, the clustering coefficient will then follow,**

$$C_i = \frac{2\left(L_i + L^{\dagger}\right)}{k_i\,(k_i - 1)}$$

- **where the correction factor has the following values:**
  - **Case 1:** $k_{lim} = \frac{2}{3}(N-1)$

$$L^{\dagger} = \frac{\frac{3}{2}k^{\dagger}\left(\frac{3}{2}k^{\dagger} + 1\right)}{2}$$

  - **Case 2:** $k_{lim} = \frac{2}{3}(N-2)$

$$L^{\dagger} = \frac{\frac{3}{2}k^{\dagger}\left(\frac{3}{2}k^{\dagger} - 1\right)}{2}$$

  - **Case 3:** $k_{lim} = \frac{2}{3}(N-3)$

$$L^{\dagger} = \frac{\left(\frac{3}{2}k^{\dagger} - 1\right)\left(\frac{3}{2}k^{\dagger} - 2\right)}{2}$$

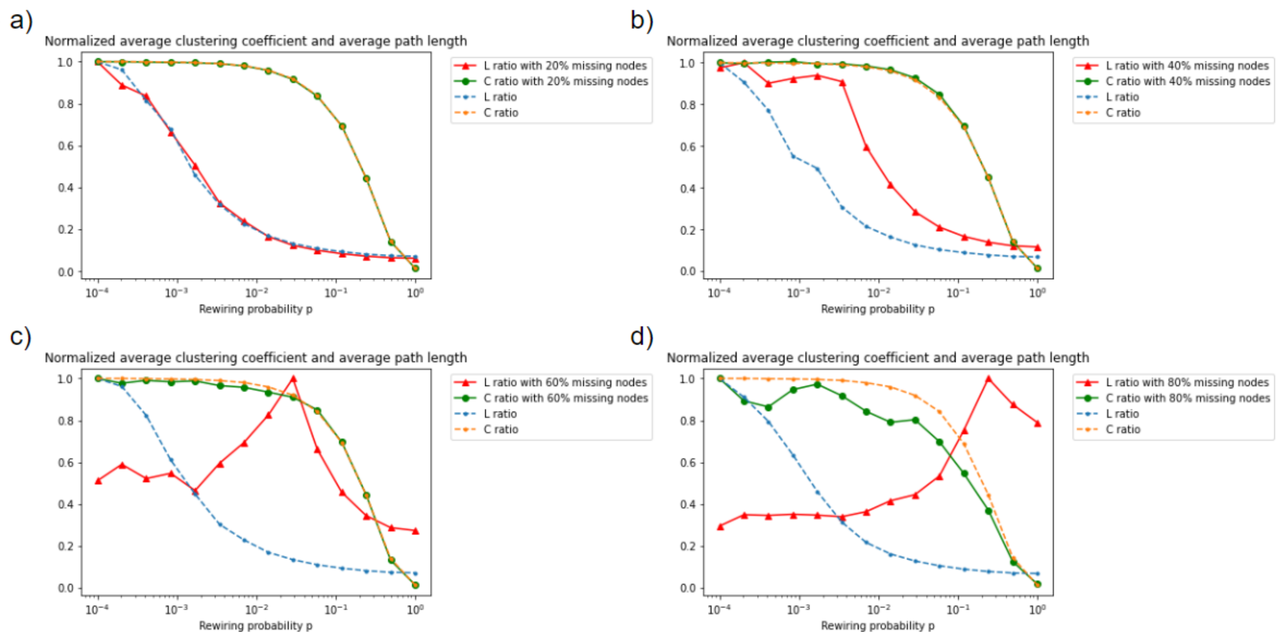

Figure 5. Comparison between the recreated Watts-Strogatz graph with complete (broken lines) and missing nodes (solid lines). As the number of nodes missing increases, a change in the behavior of the average path length can be observed.

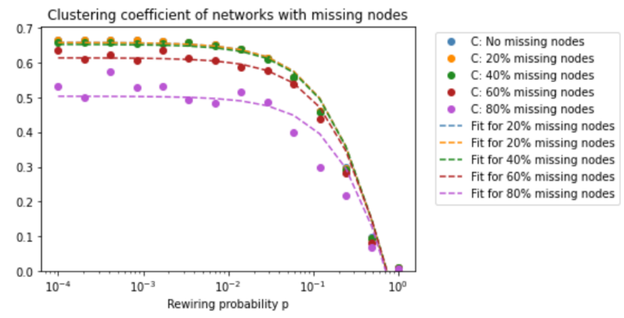

Figure 6. Comparison of the clustering coefficients of the complete network and the networks with nodes missing. The dashed lines are the curve fits of the networks.

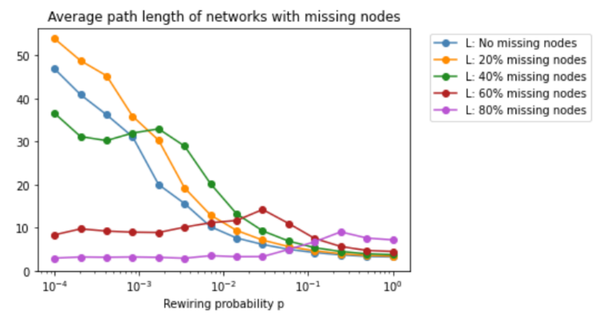

Figure 7. Comparison between the average path lengths of the complete network and the networks with nodes missing.

- **Clustering coefficients of the networks with 20% and 40% nodes missing do not vary significantly from when there were no missing nodes with average deviation values of 0.082 and 0.150 respectively.**
- **Networks with 60% and 80% nodes missing still follow the same behavior but with lesser clustering coefficient values having average deviations 0.259 and 0.371 respectively.**

SUMMARY

- Clustering coefficient of Watts-Strogatz network increases as node degree k increases and when probability of rewiring p approaches 1. For random networks, the behavior is linear. For regular networks, the behavior follows our approximation.
- When removing nodes, the local clustering coefficient varies and variability increases as more nodes are removed.
- When nodes are removed, we observe a change in behavior of the average path length but the clustering coefficient generally follows the same trend..
- The path lengths are dependent on the size of the largest connected component considered.

ACKNOWLEDGEMENT

The authors would like to acknowledge **Mindanao State University Iligan Institute of Technology (MSU IIT)**, **Premier Research Institute of Science and Mathematics (PRISM),** and **Department of Science and Technology - Science Education Institute (DOST-SEI)** for providing the necessary facilities for this study.