

PySGN: A Python package for constructing synthetic geospatial networks

Boyu Wang¹, Andrew Crooks¹, Taylor Anderson², and Andreas Züfle³

¹ Department of Geography, University at Buffalo, Buffalo, New York, USA ² Geography & Geoinformation Science Department, George Mason University, Fairfax, Virginia, USA ³ Department of Computer Science, Emory University, Atlanta, Georgia, USA ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [↗](#)

Submitted: 05 June 2025

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

PySGN (Python for Synthetic Geospatial Networks) is an open-source Python package designed to construct synthetic undirected geospatial networks that consider the coordinates of nodes. By incorporating the spatial information, the package extends classic network generation models, such as Erdős-Rényi (Erdős & Rényi, 1960), Watts-Strogatz (Watts & Strogatz, 1998) and Barabási-Albert (Barabási & Albert, 1999). It creates synthetic geospatial networks that mimic the spatial relationships observed in real-world networks (e.g., decaying connectivity with respect to node distance) and can be used for various simulation and analysis tasks. PySGN integrates with the PyData ecosystem, utilizing libraries, such as GeoPandas (Bossche et al., 2024) and NetworkX (Hagberg et al., 2007), to facilitate the creation, utilization, and analysis of synthetic geospatial networks.

PySGN offers flexibility in parametrization, allowing users to define networks by an average node degree or expected degree for each node. Users can also define custom constraints and control the rewiring probability, distance decay, and other key parameters, making it adaptable to various research contexts. PySGN is intended for researchers and practitioners in fields, such as urban planning, epidemiology, and social science, who require robust tools for simulating and analyzing complex geospatial networks.

Statement of Need

The need for synthetic geospatial networks arises from their utility in social simulations, including modeling of transportation systems, pedestrian movements, and the spread of infectious diseases (Züfle et al., 2024). Traditional synthetic populations often lack the integration of geographic social networks, which are crucial for accurately capturing social connections and spatial dynamics, to explore the effects of spatial proximity on social interactions, mobility patterns, and network robustness (Jiang et al., 2024).

PySGN addresses this gap by providing a tool that not only generates geographically explicit networks but also incorporates key network properties, such as clustering, preferential attachment and spatial decay. These features allow users to explore different network properties and configurations (e.g., average node degree). This is essential for a variety of simulation scenarios, where understanding spatial relationships and social dynamics is critical for analyzing and modeling complex systems. This makes PySGN suitable for diverse applications, including infrastructure resilience studies, agent-based modeling, and geospatial data analysis.

Acknowledgements

The algorithms implemented in PySGN are based on the work of Alizadeh et al. (2017) and Gallagher et al. (2023), with several improvements and modifications, including bug fixes, performance enhancements, and additional features. We would like to thank the authors for their contributions to the field of synthetic geospatial network generation.

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/Interior Business Center (DOI/IBC) contract number 140D0423C0025. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI/IBC, or the U.S. Government.

References

- Alizadeh, M., Cioffi-Revilla, C., & Crooks, A. (2017). Generating and analyzing spatial social networks. *Computational and Mathematical Organization Theory*, 23, 362–390. <https://doi.org/10.1007/s10588-016-9232-2>
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512. <https://doi.org/10.1126/science.286.5439.509>
- Bossche, J. V. den, Jordahl, K., Fleischmann, M., Richards, M., McBride, J., Wasserman, J., Badaracco, A. G., Snow, A. D., Ward, B., Tratner, J., Gerard, J., Perry, M., cjof, Hjelle, G. A., Taves, M., Hoeven, E. ter, Cochran, M., Bell, R., rraymondgh, ... Gardiner, J. (2024). *Geopandas/geopandas: v1.0.1* (Version v1.0.1). Zenodo. <https://doi.org/10.5281/zenodo.12625316>
- Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. In *Publ. math. inst. hung. acad. sci* (Vol. 5, pp. 17–60).
- Gallagher, K., Anderson, T., Crooks, A., & Züfle, A. (2023). Synthetic geosocial network generation. *Proceedings of the 7th ACM SIGSPATIAL Workshop on Location-Based Recommendations, Geosocial Networks and Geoadvertising*, 15–24. <https://doi.org/10.1145/3615896.3628345>
- Hagberg, A., Swart, P. J., & Schult, D. A. (2007). *Exploring network structure, dynamics, and function using NetworkX*. Los Alamos National Laboratory (LANL). <https://doi.org/10.25080/tcww9851>
- Jiang, N., Yin, F., Wang, B., & Crooks, A. T. (2024). A large-scale geographically explicit synthetic population with social networks for the united states. *Scientific Data*, 11(1), 1204. <https://doi.org/10.1038/s41597-024-03970-1>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, 393(6684), 440–442. <https://doi.org/10.1038/30918>
- Züfle, A., Pfoser, D., Wenk, C., Crooks, A., Kavak, H., Anderson, T., Kim, J.-S., Holt, N., & Diantonio, A. (2024). In silico human mobility data science: Leveraging massive simulated mobility data (vision paper). *ACM Transactions on Spatial Algorithms and Systems*, 10(2), 1–27. <https://doi.org/10.1145/3672557>