An Online Visual Analytics Approach for Evaluating LLM-Generated Protein Interactions

Allison Austin [1, 2] Shilpika [1] Aditya Tanikanti [1] Venkatram Vishwanath [1] Michael E. Papka [1]

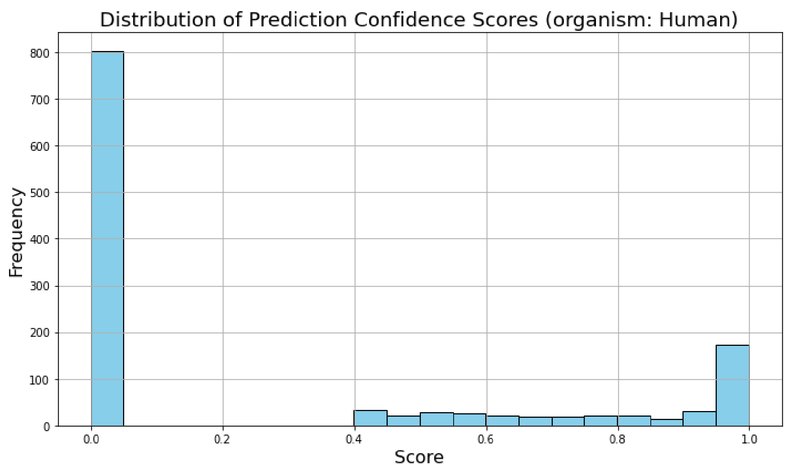
[1]Argonne Leadership Computing Facility, Argonne National Laboratory [2]University of California, Davis

Abstract

Protein-protein interaction (PPI) data has been widely used to enrich visual representations of PPI networks, and has been expanded to supporting large-language models (LLMs) for various tasks (e.g. protein generation, interaction prediction). While LLM-generated protein interactions have been extensively studied in recent years, existing tools for visualizing PPI networks do not support evaluation of LLM-generated data. In this work, we present a visual analytics tool for exploring and evaluating PPIs generated by large-language models. This tool incorporates standard PPI visualization and analysis techniques while also depicting prediction confidence for improved user understanding and discovery of AI-generated protein interaction data.

Analysis

We evaluated the prediction confidence of a small LLM-generated PPI network of 502 proteins and 1,224 interactions.



- Majority showed scores of 0.0 (**65.55%**), indicating most of the PPIs generated by the LLM were not found in STRING.
- Of the interactions in STRING (contained non-zero scores), approximately half (**47.98%**) had **high confidence scores** (between 0.9-1.0).

Visualization

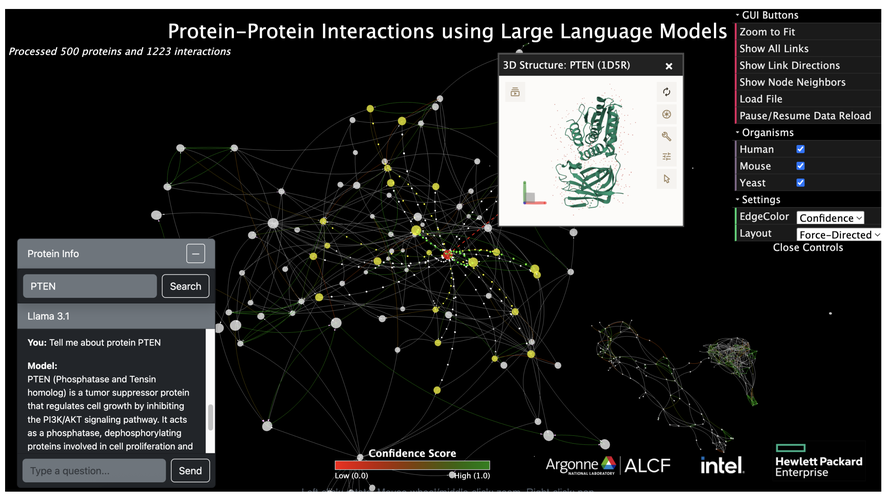


Figure 1. An overview of the UI of our system

Our web-based visualization tool features two different graph layout options: (1) spherical and (2) force-directed. The 3D network shows proteins (nodes) and interactions (edges). The colors of the edges show overall confidence score of LLM prediction (red: low, green: high). Gray edges indicate the interaction was not found in the STRING database [1]. The source code for this tool is available at `https://github.com/sshilpika/protein-graph-visualization`.

Chatbot

Users can prompt an AI chatbot to gain additional information about a set of proteins or interactions.

Dynamic search

To improve rendering and exploration, we enable filtering of interaction type and protein search.

Protein structure comparison

The UI also allows the user to view the 3D structure of a selected protein(s). Multiple 3D structures can be visualized at once for comparison. The 3D structure is rendered using MolStar Viewer [2].

Motivation

Existing tools for visualizing PPI networks have **limited support for LLM integration or evaluation** of LLM-generated network data in web-based tools for PPI analysis.

- Popular tools like Gephi [3], gnuplot [4], and Graphia [5] are not optimized for analyzing scale-free networks or biological data.
- Tools designed for molecular interaction networks (Cytoscape [6], OmicsNet [7]) have scalability concerns, do not evaluate confidence of LLM-generated data.
- Existing methods for validating functional relevance of PPI data typically rely only on enrichment analysis from publicly-available datasets, no streaming support.

Conclusions

We have successfully incorporated an LLM evaluation component into a PPI network visualization.

We have further enhanced user understanding of LLM-generated PPIs by adding a chatbot and an optimized 3D graph layout with the implementation of the concentric sphere–a more optimal graph layout for scale-free networks (PPIs).

Next Steps

- Perform more in-depth analysis on the PPI network (KEGG/GO enrichment).
- Add retrieval step to LLM workflow (Retrieval-Augmented Generation) to enable relevant querying of data from additional sources.
- Performance testing of large PPI networks (100,000+ nodes/edges).
- Conduct user studies to further test how the visualization enhances user understanding of the predictions.

References

[1] D. Szklarczyk *et al.*, "The string database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Research*, vol. 49, no. D1, p. D605–D612, 2021.

[2] D. Sehnal *et al.*, "Mol* viewer: modern web app for 3d visualization and analysis of large biomolecular structures," vol. 49, pp. W431–W437.

[3] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2009, pp. 361–362. [Online]. Available: https://gephi.org

[4] Thomas Williams and Colin Kelley and others, "Gnuplot: An interactive plotting program," http://www.gnuplot.info/, 2020.

[5] T. C. Freeman *et al.*, "Graphia: A platform for the graph-based visualisation and analysis of high dimensional data," *PLOS Computational Biology*, vol. 18, no. 7, 2022.

[6] K. Ono *et al.*, "Cytoscape web: bringing network biology to the browser," vol. 53, pp. W203–W212.

[7] G. Zhou and J. Xia, "OmicsNet: a web-based tool for creation and visual analysis of biological networks in 3d space," vol. 46, pp. W514–W522. [Online]. Available: https://doi.org/10.1093/nar/gky510

[8] A. Ahmed *et al.*, "Visualisation and analysis of large and complex scale-free networks," pp. 1–8, 2005.

[9] W. Kwon *et al.*, "Efficient memory management for large language model serving with pagedattention," in *Proceedings of the ACM SIGOPS 29th Symposium on Operating Systems Principles (SOSP)*, 2023.

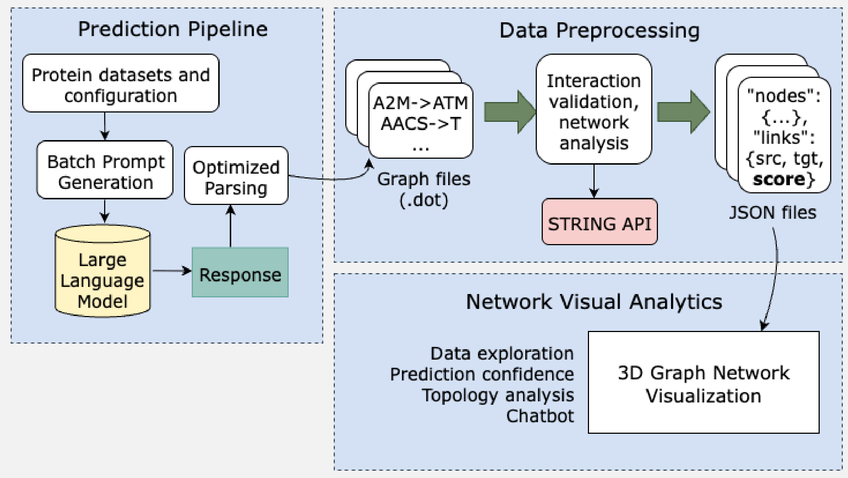
[10] M. Salim, T. D. Uram, J. T. Childers, P. Balaprakash, V. Vishwanath, and M. Papka, "Balsam: Automated scheduling and execution of dynamic, data-intensive HPC workflows," in *Proceedings of the 8th Workshop on Python for High-Performance and Scientific Computing (PyHPC)*. ACM Press, 2018.

[11] D. Szklarczyk *et al.*, "String api documentation," https://string-db.org/help/api/, 2024.

Methods

Comprehensive pipeline for generating PPIs using large language models:

- **Generation Layer**: vLLM [9] prompt creation using batch API processing.
- **Processing Layer**: Utilizes Balsam [10] to run batch prompts on Sophia. Llama 2 and Llama 4 models were used to generate the predictions.
- **Parsing Layer**: Response extraction and validation.
- **Output Layer**: Network visualization and analysis.



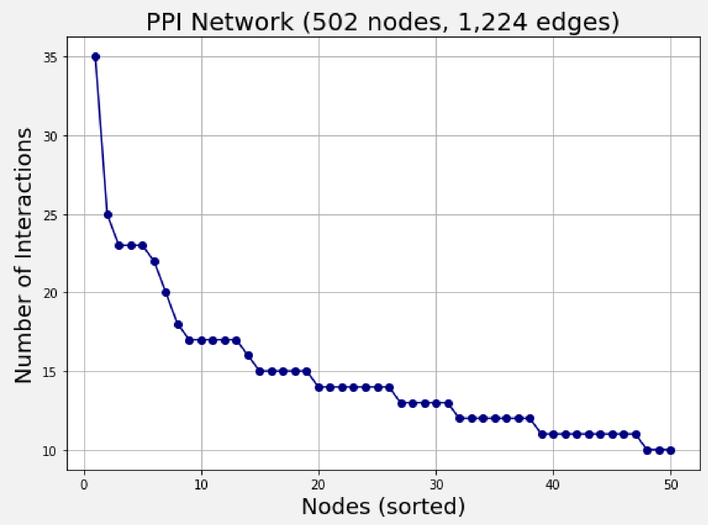
The source code for generating and parsing the LLM output is available at `https://github.com/atanikan/balsam_ppi_llama`.

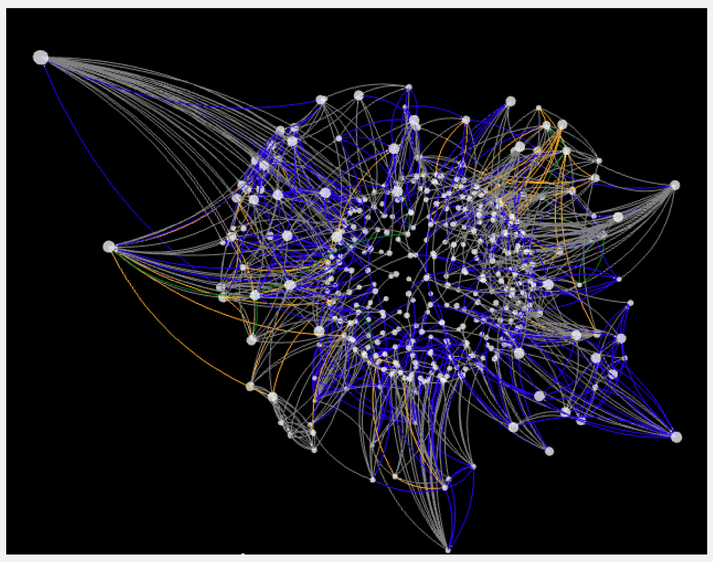
Prediction evaluation

We embedded scores for each interaction from the STRING API [1] by searching the predicted link across three different organisms (human, mouse, yeast):

Graph layout optimization

We improved the rendering and user interpretation of the PPI network by applying the concentric spherical layout [8]. This layout is optimal for *scale-free networks* like PPIs, as it allows the viewer to focus on the smaller group of proteins that have the most interactions.





Acknowledgments

This research used resources of the Argonne Leadership Computing Facility, a U.S. Department of Energy (DOE) Office of Science user facility at Argonne National Laboratory and is based on research supported by the U.S. DOE Office of Science-Advanced Scientific Computing Research Program, under Contract No. DE-AC02-06CH11357.