The Journal of Open Source Software

HINA: A Learning Analytics Tool for Heterogenous Interaction Network Analysis in Python

Shihui Feng [1,2¶], **Baiyue He** [2], **and Alec Kirkley** [2,3,4]

1 Faculty of Education, University of Hong Kong, Hong Kong, China **2** Institute of Data Science, University of Hong Kong, Hong Kong, China **3** Department of Urban Planning and Design, University of Hong Kong, Hong Kong, China **4** Urban Systems Institute, University of Hong Kong, Hong Kong, China
¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Learning analytics is a scientific field created in response to the need for deriving meaningful insights from learning process data, to understand how learning occurs and to optimize learning design (Gašević et al., 2015). **HINA** is a learning analytics tool that models and analyzes the relationships among different types of learning process data to uncover hidden learning strategies, provide learning performance indices, identify clusters, and generate dashboard visualizations.

Heterogenous interactions refer to the connections that occur between different types of entities during learning processes, such as students' interactions with learning objects (Feng et al., 2025) or students' affiliations with different coded behaviors (Feng et al., 2024). These heterogenous interactions can be modelled with **heterogenous interaction networks (HIN)** that consist of different sets of nodes, with edges only connecting nodes between different sets. Examples of heterogenous interaction networks in learning analytics are presented in Figure 1 below. HINA offers a set of flexible and adaptive methods to model a wide variety of interactions that can occur during learning processes in individual and collaborative learning contexts.

The Journal of Open Source Software

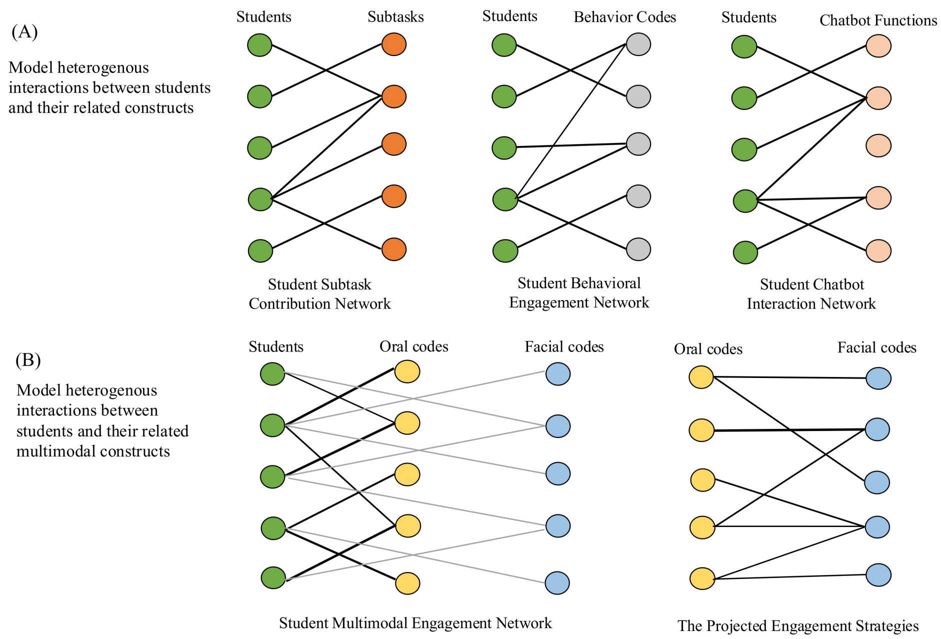

Figure 1: Examples of heterogenous interaction networks for learning in HINA.

Statement of need

Constructivism theory of learning emphasizes that learning occurs through students' active interactions with various aspects of their environment (Vygotsky, 1978) (Piaget, 1976). Students' interactions with their learning environment—for example, engagement with designed learning artefacts or with coded latent constructs—are inherently heterogenous, and can be captured using multimodal process data. HINA offers analytical modules to model these heterogenous interactions and address questions at multiple levels of interest, for example:

- How can we assess the quantity and diversity of individuals' interactions with their designed learning environments?
- How can we identify the interactions among pairs of nodes that are statistically significant under a suitable null model?
- How can we identify subgroups of individuals that share similar learning strategies indicated by their heterogenous interaction patterns?
- How can we visualize these heterogeneous interaction networks in an interactive and informative way, using different visualization formats that are tailored for learning analytics applications and implementations?

These analytical features offered by HINA can analyze multimodal process data to address a wide range of research questions in learning analytics. For example, studies can explore how to gauge individual contribution based on the interactions between students and learning artefacts (Feng et al., 2025), identify subgroups of students who share similar learning strategies based on their associations with behavioral and cognitive constructs during learning processes (Feng et al., 2024), uncover significant associations among behavioral engagement in different modalities (Feng et al., 2024), or design learning analytics dashboards for the visualization of heterogenous engagement to support teaching and learning practices (Feng et al., 2025).

HINA tailors its methods—which include brand-new algorithms for pruning, clustering, and visualization in the HIN setting—specifically for learning analytics researchers and teachers working with HINs derived from learning process data. This makes HINA a unique contribution to the software space that provides a more specialized experience than existing packages for general network analysis (Hagberg et al., 2008) (Csardi & Nepusz, 2006) or network inference (Peixoto, 2014) (Kirkley & He, 2024).

Current Modules

- **Network construction** (hina.construction)
    - Provides functions to construct Heterogeneous Interaction Networks (HINs) (see examples in Figure 1A, 1B) directly from input learning process data. The methods in this module are designed to handle the typical data format encountered for learning process data traces, supporting seamless integration with learning analytics workflows.
- **Individual-level analysis** (hina.individual)
    - Provides functions to compute the node-level measures of (Feng et al., 2025) gauging the quantity and diversity of individuals' connections to different learning constructs. Students' group information and construct attributes can be flexibly manipulated for different applications.
- **Dyadic-level analysis** (hina.dyad)
    - Provides methods to identify statistically significant edges in the heterogeneous interaction network relative to different null models of interaction structure (Feng et al., 2025), which can be specified by the user.
- **Mesoscale clustering** (hina.mesoscale)
    - Provides methods for clustering nodes in a heterogeneous interaction network according to shared interaction structure (Feng et al., 2024), to automatically learn the number of clusters from heterogeneity in the interaction data to find a mesoscale representation. Utilizes a novel method based on data compression for parsimonious inference. If the input is a tripartite representation of heterogeneous interaction network, the function also returns the projected bipartite networks of the related constructs of individuals within each cluster.
- **Network visualization** (hina.visualization)
    - Provides network visualization functionalities for heterogeneous interaction networks. Users can generate a customizable network visualization using a specified layout, allowing for the pruning of insignificant edges, grouping of nodes based on engagement patterns, and customization of the graph's appearance. Users can also visualize HINs with a novel community-based layout, emphasizing the underlying bipartite community structure.
- **Dashboard deployment** (hina.app)
    - Provides functions to deploy a dashboard that includes a web-based interface serving multiple purposes.
        1. The dashboard serves as a web-based tool for conducting learning analytics with HINA using an intuitive user interface, enabling users to conduct the individual-, dyadic- and mesoscale-level analysis available in the package without any programming.
        2. The dashboard also allows teachers and students to visualize, interpret, and communicate HINA results effectively.

This dual functionality supports both data analysis and the sharing of actionable insights in an interactive and user-friendly manner, making it a versatile tool for both data analytics and teaching practice.

Acknowledgements

This work was supported by the Research Grants Council of Hong Kong under ECS Grant No. 27605223 (Shihui Feng) and HKU Institute of Data Science Seed Fund Grant No. IDS-RSF2023-0006 (Shihui Feng, Alec Kirkley).

References

Csardi, G., & Nepusz, T. (2006). The Igraph software package for complex network research. *InterJournal Complex Systems*, *1695*, 1–9. https://doi.org/10.5281/zenodo.3630268

Feng, S., Gibson, D., & Gasevic, D. (2025). Analyzing students' emerging roles based on quantity and heterogeneity of individual contributions in small group online collaborative learning using bipartite network analysis. *arXiv Preprint arXiv:2502.19112*. https://doi.org/10.18608/jla.2025.8431

Feng, S., Yan, L., Zhao, L., Maldonado, R. M., & Gašević, D. (2024). Heterogenous network analytics of small group teamwork: Using multimodal data to uncover individual behavioral engagement strategies. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 587–597. https://doi.org/10.1145/3636555.36369

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*, 64–71. https://doi.org/10.1007/s11528-014-0822-x

Hagberg, A., Swart, P. J., & Schult, D. A. (2008). *Exploring network structure, dynamics, and function using NetworkX*. Los Alamos National Laboratory (LANL), Los Alamos, NM (United States). https://doi.org/https://doi.org/10.25080/tcwv9851

Kirkley, A., & He, B. (2024). PANINIpy: Package of Algorithms for Nonparametric Inference with Networks In Python. *Journal of Open Source Software*, *9*(103), 7312. https://doi.org/10.21105/joss.07312

Peixoto, T. P. (2014). The graph-tool python library. *Figshare*. https://doi.org/10.6084/m9.figshare.1164194

Piaget, J. (1976). *Piaget's Theory*. Springer. https://doi.org/10.1007/978-3-642-46323-5_2

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (Vol. 86). Harvard University Press. https://doi.org/10.4236/psych.2012.35060