Using Social Network Analysis to Analyze Eye-tracking Behavior Data in Education Science

Pingjing Yang

University of Illinois at Urbana-Champaign, Urbana IL 61820, USA
py2@illinois.edu

Abstract. This study aims to systematically evaluate the use of social network analysis (SNA) metrics to measure eye-tracking behavior to assess and predict student learning performance. We integrated 11 network metrics from published research and tested them on six eye-tracking datasets. Our preliminary results indicate that no consistent predictor variable can effectively predict student performance across different datasets. The number of nodes, edges, reciprocity, and entropy measures contribute differently to predicting students' performance. This work deepens our understanding of how different SNA metrics relate to eye-tracking data and advances the methodological framework to predict learning outcomes.

Keywords: Social Network Analysis; Eye-tracking Network; Education Science

1 Introduction

Social Network Analysis (SNA) has been widely adopted in areas including social science [10], business administration [1,5,22], communication research [6,14,21], and crisis informatics [7,8,13]. In the area of education science, researchers still struggle to adopt SNA techniques to analyze learning materials. One possible reason is the complexity of the education process and the challenges in modeling educational data in the network framework. In this work, we explore how to use social network analysis to investigate educational data, specifically eye-tracking data to predict students' learning outcomes.

Evaluating student learning outcomes provides opportunities to create interactive and customized learning experiences, which can facilitate better learning outcomes. Most prior approaches have focused on capturing language signals in written tests, such as syntax, lemmas, or word patterns [25]. Learning is a complex procedure, and written text or think-aloud data is not always collected as part of the learning process. Hence, we suggest a network analysis-based approach to learn rules to analyze behavioral data (eye-tracking activity) to evaluate student learning outcomes.

Researchers have found that eye-tracking data such as fixations and saccades correspond with the complexity of material and viewer engagement, examples of how eye-tracking technology gives empirical evidence on visual attention [15].

These areas of interest (AOI) movement patterns may imply the students' attention to learning materials if students have spent adequate efforts absorbing learning materials. However, these measures do not always capture contextual information in user behavior.

Network representations with AOIs as nodes and movements between AOIs as edges enable network-based analyses of the student's learning process. For example, by tracking how frequently one student moves from AOI-a to AOI-b allows us to find how much effort they spent on specific topics and to infer how well they master the knowledge. Other metrics such as centrality and structural measures may also evaluate different aspects of students' behavior.

It is not clear how accurately we can use these network metrics to approximate student learning outcomes. Hypothetically, if network metrics can effectively predict students' learning outcomes, educators can provide more adaptive educational responses to students to improve learning outcomes in the future. For example, educators can apply interventions when student eye-tracking data-predicted learning performance is below expectations.

This paper reports on an empirical study to exhaustively evaluate network analysis metrics on eye-tracking behavior data in the context of learning. The findings will mitigate the gap between SNA and the analysis of education data. This work sheds light on the development of novel computational solutions for analyzing eye-tracking data using network representations. We close with discussions for researchers who are interested in applying these metrics to eye gaze data to explain variance in learning-relevant outcomes.

2 Related Work

Social network analysis has been applied to study peer interactions and their impact on learning outcomes. Grunspan et al. [12] tested the association between network position and success on exams using in-classroom networks. Peeters [23] used online activities on Facebook to determine topics, challenges, and personal experiences that the students encountered. At the individual level, modeling the personal learning process in the framework of a network is still challenging. In this work, we use the eye-tracking-based scanpath network to represent a personal learning process.

Choosing appropriate network metrics is another challenge to adopting SNA. Centrality measures are often considered default measures in evaluating learning outcomes, such as the quiz / post-test score. More recent studies applied entropy measures to sequential networks but found possibly contradictory results. For example, [16] found a lower entropy associated with better performance in information seeking tasks, while [9] found higher entropy for engineers who perform better in manufacturing. Efforts need to be made to unify these different SNA metrics to understand why and how SNA metrics are related to a specific performance outcome under certain task conditions.

3 Methods

3.1 Datasets

We will reuse six previously collected eye-tracking datasets published in [3] and two unpublished education experiments. During these experiments, we recruited 30 undergraduate students per study who had completed at least one introductory biology course to participate individually in a 1-hour session in 2018 (2018 AH, 2018 TC) and 2019 (2019 AH, 2019 TC), and an online Zoom session in the 2020 study (2020 Body, 2020 Cells). In each study, the students studied two 4-page sets of illustrated passages about the immune system, intending to learn the material well enough to explain it to a peer. After reading each set, participants provided drawn and oral explanations. All responses were transcribed, and analyzed for recalled noun and verb phrases. A composite posttest score was calculated using Principal Components Analysis to objectively assess their explanations. Overall, we gathered six sets of think-aloud protocols, eye-tracking data, and PCA scores in the three studies. In this study, we will use eye-tracking data to construct network representations and PCA scores as performance evaluations.

3.2 Network Analysis

We create individual networks by converting students' eye-gaze movement data to network representations. The nodes represent the focus of the individual AOIs, and the edges represent the movement from one AOI to another. Considering learning processes often involve forth and back movements between nodes, we use multi-edge-directed graphs to model networks. Figure 1 illustrates network representations of individual eye gaze movement. Network A demonstrates that during the study a participant followed the order of AOIs in the learning materials and read the texts and diagrams. In contrast, network B exhibits a student who went forth and back across text and the corresponding diagrams. Network C shows a student following their own order to read the learning materials instead of following the designated linear order from t1 to d4.

We selected the following four categories of network analysis metrics from previous work: basic measures, centrality measures, structural measures, and entropy measures.

Basic measures refer to the number of nodes and edges in individual networks. The number of nodes represents the AOIs a student has explored; the number of edges represents the total movements across AOIs a student has explored.

Centrality measures refer to a collection of metrics measuring the importance of individual nodes or overall centrality (average centrality). Following existing research, we select degree centrality [4], betweenness centrality [17], closeness centrality [17], eigenvector centrality [24] and pagerank [18].

Structural measures include measures on the overall structure of networks. For example, density, a measure of edges as a proportion of all possible edges,

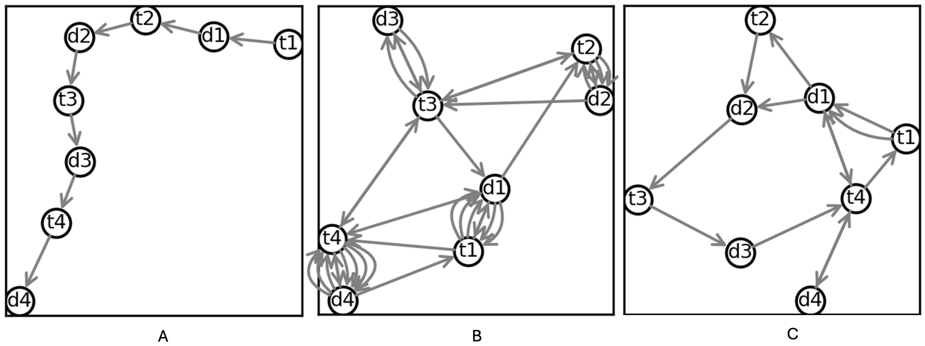

Fig. 1. Illustrations of eye gaze movement networks for three different participants. d refers to diagram and t to text, respectively.

was reported positively related to learning outcomes [26]. Reciprocity, a measure of the ratio of back-and-forth movement can be correlated with better outcomes when represented by integrating information from different sources or locations [20]. Node connectivity is equal to the minimum number of nodes that must be removed to disconnect a network or render it trivial [2, 11].

Additionally, **entropy measures** [19] are rarely used to quantify scanpaths in traditional education research. The stationary entropy captures the distribution of eye gaze across different AOIs. Transitional entropy measures unpredictability and randomness in the sequence of transitions.

We use posttest scores as student performance evaluations. As mentioned in the dataset section, we applied Principal Component Analysis (PCA) to weigh the drawn elements, drawn relations, oral elements, and oral relations generated in studies to form posttest scores as proxi of students' performance. These post-test scores are necessary for a preliminary study to select effective network metrics and models to find possible correlations. We follow a two-step process: 1. using pairwise Pearson correlations to filter out negative coefficient metrics; 2. fitting ridge regression models with these positive coefficient metrics against the posttest score.

In formula 1, $\mathrm{cov}(x_i, \mathrm{y})$ represents the covariance between the network metric x_i and posttest score y, σ_{x_i} and σ_{y} are the standard deviations of x_i and y, respectively. After removing all network metrics with a negative Pearson correlation ρ, we vary parameter α and rely on mean square error to fit Ridge regression models following the formula 2 objective function. $\mathbf{y}$ is the vector of observed values for the posttest. $\mathbf{X}$ is the matrix of predictor variables that represent the metrics of the network. β is the vector of regression coefficients. α is the regularization parameter. We cross-validate α in the range of 0 to 100 with a step size of 0.001 to select the appropriate α to minimize the mean square error.

$$\rho_{x_i,\mathrm{y}} = \frac{\mathrm{cov}(x_i, \mathrm{y})}{\sigma_{x_i}\sigma_{\mathrm{y}}} \tag{1}$$

$$\hat{\beta}^{ridge} = \arg\min_{\beta} \left\{ \|\mathbf{y} - \mathbf{X}\beta\|^2 + \alpha\|\beta\|^2 \right\} \quad (2)$$

4 Results and Discussion

Table 1. Coefficients of Ridge regression models after removing negative coefficients (empty cells).

dataset	**2018 AH**	**2018 TC**	**2019 AH**	**2019 TC**	**2020 Body**	**2020 Cells**
alpha	0.010	0.100	0.010	0.001	0.010	0.010
mse	0.82	0.84	0.82	0.57	0.37	0.66
#nodes	0.00	0.30	0.20	0.09	0.00	0.26
#edges	0.00	0.10	0.01		0.23	0.00
deg centrality	0.00	0.01	0.00		0.24	0.00
closeness centrality	0.12				0.00	0.00
eigenvector centrality	0.00	0.13			0.02	
degree centralization	0.27	0.01			0.00	0.03
density	0.00	0.01	0.00		0.24	0.00
node connectivity	0.13		0.04			
reciprocity			0.00	0.34	0.00	0.28
stationary entropy	0.27	0.00	0.00	0.18		0.00
transition entropy	0.00		0.08		0.00	0.13

Table 1 shows an overview of the best model selection for each dataset. We used the mean squared error (MSE) to choose the best alpha score for Ridge regression models. There are no consistently effective network metrics to predict students' performance. Reciprocity is the most predictive metric in the 2019 TC datasets ($\beta = 0.34$) and 2020 Cells ($\beta = 0.28$) datasets. Centralization ($\beta = 0.27$) and stationary entropy ($\beta = 0.27$) have the highest correlations with students' performance in the 2018 AH dataset. In the 2018 TC and 2019 AH datasets, the number of nodes is most correlated ($\beta = 0.30$ and $\beta = 0.20$ respectively) with the performance of the students. In the 2020 Body dataset, the number of edges ($\beta = 0.23$), average degree centrality ($\beta = 0.24$), and density ($\beta = 0.24$) have coefficients similar to students' performance.

Figure 2 shows how the selected metrics fitted the ridge model perform against actual data points. These scatterplots show that our fitted models do not perform well on 2018 AH, 2018TC, and 2019AH datasets, as the MSE are 0.82, 0.84, and 0.82 respectively. Our regression model performs best on 2020 Body dataset with MSE equals 0.37. These results suggest for some datasets that there may exist more complex relations between network metrics and students' performance.

The preliminary results do not directly suggest any specific rules for selecting network metrics in the task of predicting student learning outcomes. Researchers

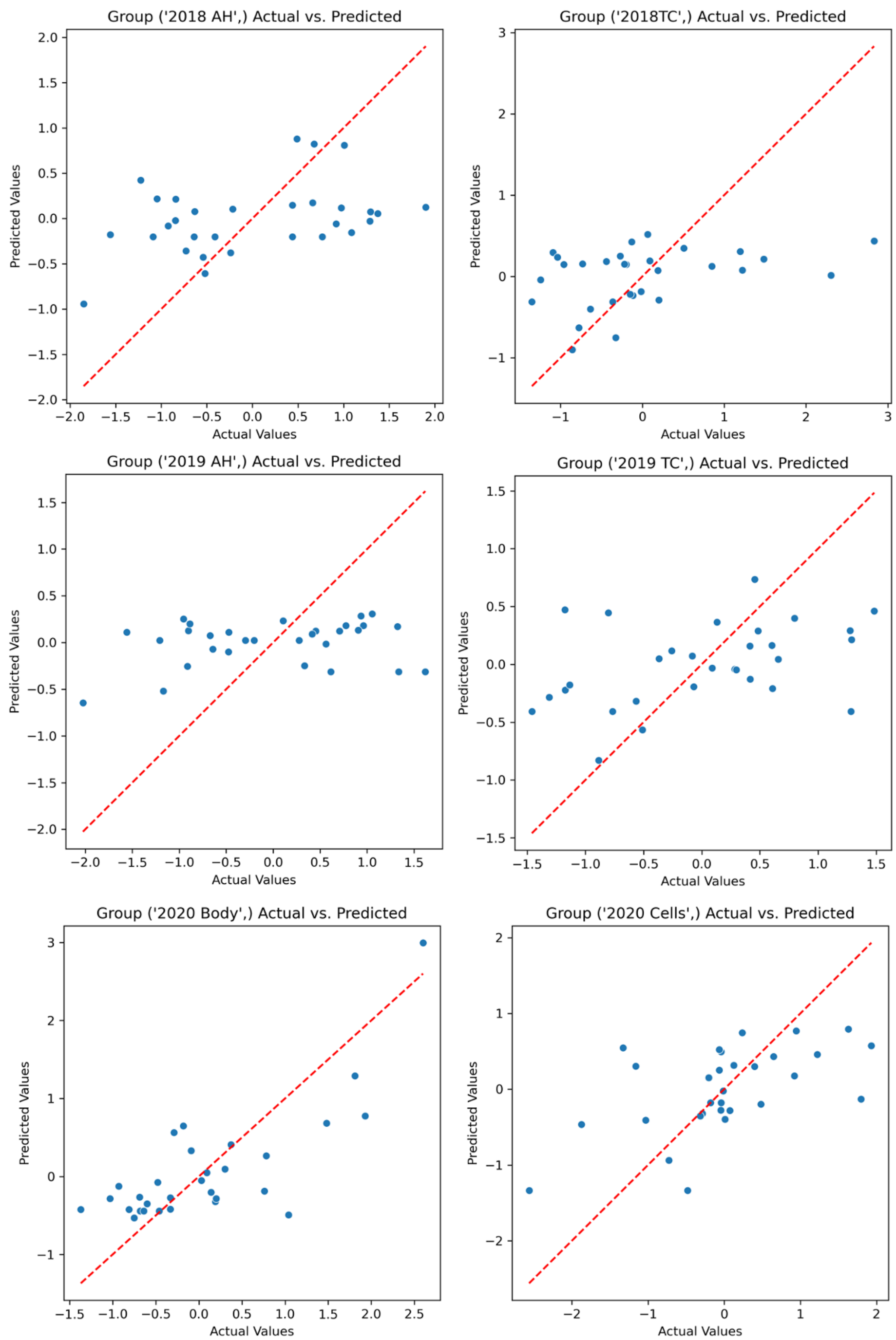

Fig. 2. Predicted posttest score against actual posttest score in six different datasets using Ridge Regression models.

still need to carefully select metrics to fit prediction models. Additionally, several metrics show less universality across datasets. For example, closeness centrality has a positive coefficient in only one model. More sophisticated models are needed to capture the correlations between various network metrics and student learning outcomes.

The way we create network representations may also affect model predictions. The current multi-edge directed network does not capture the fixation time on each node or the order of movement, meaning the varying amounts of effort students spend learning new material are not reflected in the network metrics. Future work will involve incorporating this rich node and edge information into models of the student learning process.

5 Conclusion

In this study, we explored the application of social network analysis (SNA) metrics to eye-tracking behavior data to evaluate and predict student learning performance. By integrating eleven network metrics from published research, we tested the metrics' effectiveness in six distinct eye-tracking datasets. Our findings revealed that there is no consistent predictor variable that can effectively predict student performance across different datasets. Metrics such as number of nodes, edges, reciprocity, and entropy measures exhibited varied contributions to predicting learning outcomes.

Despite the inconsistency in predictive power across different datasets, our study provides valuable insight into how different SNA metrics relate to eye-tracking data and their potential to predict educational outcomes. These findings highlight the complexity of using network metrics for educational data and underscore the need for further research to refine these approaches. Future work could focus on understanding the contextual factors that influence the predictive power of these metrics and developing more sophisticated models that capture the nuanced relationships between eye-tracking behaviors and learning performance. This research advances the methodological framework for predicting learning outcomes using eye-tracking data and opens avenues for more adaptive and personalized educational interventions.

Acknowledgment

Pingjing Yang is working in the Social Computing Lab, led by Professor Jana Diesner, at the University of Illinois at Urbana-Champaign. This work is conducted under the mentorship of Professor Jennifer Cromley in the College of Education at the University of Illinois at Urbana-Champaign. The research was supported with funds from the US National Science Foundation under Award No. 2225298 to the University of Illinois at Urbana-Champaign. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Bonchi, F., Castillo, C., Gionis, A., Jaimes, A.: Social network analysis and mining for business applications. ACM Transactions on Intelligent Systems and Technology (TIST) **2**(3), 1–37 (2011)
2. Cela, K.L., Sicilia, M.Á., Sánchez, S.: Social network analysis in e-learning environments: A preliminary systematic review. Educational psychology review **27**, 219–246 (2015)
3. Cromley, J.G., Kunze, A.J., Dane, A.P.: Multi-text multi-modal reading processes and comprehension. Learning and Instruction **71**, 101413 (2021)
4. Davalos, E., Vatral, C., Cohn, C., Horn Fonteles, J., Biswas, G., Mohammed, N., Lee, M., Levin, D.: Identifying gaze behavior evolution via temporal fully-weighted scanpath graphs. In: LAK23: 13th International Learning Analytics and Knowledge Conference. pp. 476–487 (2023)
5. Diesner, J., Carley, K.M.: A methodology for integrating network theory and topic modeling and its application to innovation diffusion. In: 2010 IEEE Second International Conference on Social Computing. pp. 687–692. IEEE (2010)
6. Diesner, J., Kumaraguru, P., Carley, K.M.: Mental models of data privacy and security extracted from interviews with indians. In: 55th Annual Conference of the International Communication Association (ICA), New York, NY (2005)
7. Dinh, L., Kulkarni, S., Yang, P., Diesner, J.: Reliability of methods for extracting collaboration networks from crisis-related situational reports and tweets. In: 2nd Information Systems for Crisis Response and Management Asia Pacific Conference, ISCRAM 2022. pp. 181–195. Information Systems for Crisis Response and Management, ISCRAM (2022)
8. Dinh, L., Yang, P., Diesner, J.: From plan to practice: Interorganizational crisis response networks from governmental guidelines and real-world collaborations during hurricane events. Journal of Contingencies and Crisis Management **32**(3), e12601 (2024). https://doi.org/https://doi.org/10.1111/1468-5973.12601, https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-5973.12601
9. Doellken, M., Zapata, J., Thomas, N., Matthiesen, S.: Implementing innovative gaze analytic methods in design for manufacturing: A study on eye movements in exploiting design guidelines. Procedia CIRP **100**, 415–420 (2021)
10. Edelmann, A., Wolff, T., Montagne, D., Bail, C.A.: Computational social science and sociology. Annual Review of Sociology **46**, 61–81 (2020)
11. Esfahanian, A.H.: Connectivity algorithms. Topics in structural graph theory pp. 268–281 (2013)
12. Grunspan, D.Z., Wiggins, B.L., Goodreau, S.M.: Understanding classrooms through social network analysis: A primer for social network analysis in education research. CBE—Life Sciences Education **13**(2), 167–178 (2014)
13. Hagen, L., Keller, T., Neely, S., DePaula, N., Robert-Cooperman, C.: Crisis communications in the age of social media: A network analysis of zika-related tweets. Social science computer review **36**(5), 523–541 (2018)
14. Hilbert, M., Barnett, G., Blumenstock, J., Contractor, N., Diesner, J., Frey, S., Gonzalez-Bailon, S., Lamberso, P., Pan, J., Tai-Quan, P., et al.: Computational communication science: A methodological catalyzer for a maturing discipline (2019)
15. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: Eye tracking: A comprehensive guide to methods and measures. oup Oxford (2011)

16. Iaconis, F.R., Jiménez Gandica, A.A., Del Punta, J.A., Delrieux, C.A., Gasaneo, G.: Information-theoretic characterization of eye-tracking signals with relation to cognitive tasks. Chaos: An Interdisciplinary Journal of Nonlinear Science **31**(3) (2021)
17. Jayasinghe, A., Sano, K., Rattanaporn, K.: Application for developing countries: Estimating trip attraction in urban zones based on centrality. Journal of Traffic and Transportation Engineering (English Edition) **4**(5), 464–476 (2017)
18. Kakatkar, C., Spann, M.: Marketing analytics using anonymized and fragmented tracking data. International Journal of Research in Marketing **36**(1), 117–136 (2019)
19. Krejtz, K., Szmidt, T., Duchowski, A.T., Krejtz, I.: Entropy-based statistical analysis of eye movement transitions. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 159–166 (2014)
20. Li, S., Pöysä-Tarhonen, J., Häkkinen, P.: Patterns of action transitions in online collaborative problem solving: A network analysis approach. International Journal of Computer-Supported Collaborative Learning **17**(2), 191–223 (2022)
21. Monge, P.R., Contractor, N.S.: Theories of communication networks. Oxford University Press, USA (2003)
22. Pak, S.J.: Gentlemen bankers: The world of JP Morgan. Harvard University Press (2013)
23. Peeters, W.: The peer interaction process on facebook: A social network analysis of learners' online conversations. Education and Information Technologies **24**(5), 3177–3204 (2019)
24. Ryabinin, K., Erofeeva, E., Guseva, K.: Eye tracking data mining based on fuzzy sets of fixations. In: Fuzzy Systems and Data Mining IX, pp. 11–19. IOS Press (2023)
25. Sonbol, R., Rebdawi, G., Ghneim, N.: The use of nlp-based text representation techniques to support requirement engineering tasks: A systematic mapping review. Ieee Access **10**, 62811–62830 (2022)
26. Starke, S.D., Baber, C., Cooke, N.J., Howes, A.: Workflows and individual differences during visually guided routine tasks in a road traffic management control room. Applied ergonomics **61**, 79–89 (2017)