Modelling Network Dynamics in Social Annotation

Bodong Chen[1], Basel Hussein[1] and Oleksandra Poquet[2]

[1]University of Minnesota, 1954 Buford Ave, St Paul, MN 55108, USA
[2]University of South Australia, BH3-21 City West Campus, Adelaide 5000, South Australia, Australia

Abstract
Web annotation technology is used in education to facilitate individual learning and social interaction. Departing from a conceptual exploration of social interaction in web annotation as a mediated process, as well as a dissatisfaction with analytical methods applied to web annotation data, we analyzed student interaction data from a web annotation environment following the Relational Event Modelling approach. Included in our modelling were various annotation attributes, a contextual factor of student groups, and several social and spatiotemporal factors related to network formation. Results indicated that longer annotations were slightly more likely to attract replies, students in the same project group were not more likely to engage with each other, and several network factors such as student activity, reciprocity, annotation popularity, and annotation location played important roles in interaction dynamics. This study contributes empirical insights into web annotation and calls for future work to investigate mediated social interaction as a dynamic network phenomenon.

Keywords
web annotation, network analysis, collaborative learning, digital learning

1. Introduction

Computer-mediated communication tools [1] such as asynchronous online discussions and social network sites are often used to foster learner engagement and social interaction. Web annotation is one such tool for computer-mediated communication and learning. In a nutshell, a web annotation tool allows the user to highlight a target in a web document and post an annotation referring to that target [2]. The annotations make student thinking visible and encourages learners to interact with one another. This type of learning technology is not only well suited for active learning but could intensify the social nature of learning leading to improved motivation and meaningful social participation. Although the use of web annotation is widespread in education, research investigating spatial-temporal dynamics of individual participation and emergent social interaction in annotation environments is rare. A recent review on web annotation identified different ways to use it for learning [3]. However, previous studies do not attempt to reveal mechanisms driving social interaction in web annotation and offer limited practical guidance on how to facilitate student interaction.

This study aims to bridge this gap by examining social interaction in a web annotation environment used in an online class. The study had an overarching research question: *Which*

Proceedings of the NetSciLA21 workshop, April 12, 2021
chenbd@umn.edu (B. Chen); bhussein@umn.edu (B. Hussein); sspoquet@gmail.com (O. Poquet)
https://bodong.me (B. Chen)
0000-0003-4616-4353 (B. Chen); 0000-0001-9782-816X (O. Poquet)

dynamics describe the formation of social interactions from individual annotation behaviors in web annotation? In answering the question, we explicitly considered elements of course design, technology environment, and the spatiotemporal process of engaging with the environment. Following the Relational Event Modelling framework for the modelling of dynamic networks [4], we hypothesized that social interaction in the annotation environment was driven by three types of factors: emphactor attributes such as a learner's personal background, knowledge, and dispositions, emphhistorical relational events such as previous behaviors, and emphexogenous contextual factors such as being assigned to a group to collaborate on a course project. This study makes a major step in modelling mechanisms driving peer interaction within spatiotemporal and pedagogical constraints.

2. Related Work

Due to its unique affordances, web annotation can be used to support social reading, group sense-making, knowledge construction, and community building [5, 6, 7]. For example, *Hypothes.is*, a web annotation tool, creates layers of conversation on top of web documents. A student can highlight a piece of text and make an annotation, which can be responded to by other students in this group. Research shows such social annotation can support richer communication than other tools such as newsgroup and discussion boards, mostly because the annotated documents provide a context for engaged conversations [8, 3].

To advance analysis, and eventually, interventions in web annotation environments, research needs to consider simple mechanisms that move digital learning spaces from individual interaction with content to the formation of social structures. We argue that a methodological shift in the analysis of social web annotation is needed. First, actor–artifact relations need to be considered seriously [9]. Since social interactions are mediated by the annotation artifact, artifacts can have their own properties conducive, or not so much, to attracting others to engage. Further, to understand dynamics of social interaction, quantitative analysts need to model relational events that reflect the process, rather than collapsing a series of events into relational states between actors.

This study provides an example for modelling social interaction in a web annotation environment. We explicitly include in the modelling spatial and content properties of the annotation artifacts, course design elements that may have influenced learner activity, and the temporal process of events that took place. Despite being a case study, our modelling approach offers a generalizable view on the individual to social dynamics in the web annotation environment.

Following the Relational Event Modelling framework [4], the study examined mediated social interactions as relational events. In particular, we asked: Can we predict the occurrences of relational events between actors (e.g., learners) and annotations with node-level, social, spatial, or temporal factors in this context? Informed by prior work about social dynamics in online interactions [10, 11, 12, 13], we asked the following sub-questions:

- RQ1: To what extent did the attributes of students and web annotations contribute to mediated social interaction?
- RQ2: To what extent did the course-design factor of student grouping contribute to mediated social interaction?

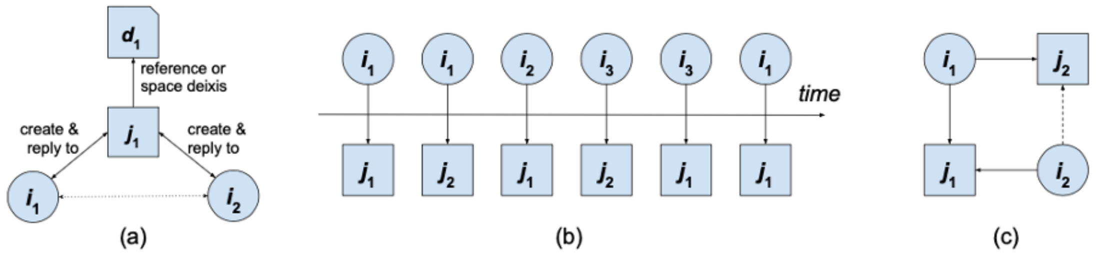

Figure 1: (a) Illustration of the mediated nature of social interaction in web annotation. An actor/author, i1, creates an annotation, j1, that references a document, d1. Another actor, i2, can indirectly communicate with i1 by replying to i2. (b) Illustration of the reply sequence. Actors/authors are represented as circles, annotations as squares. (c) The four-cycle closure is also illustrated, with the solid lines predicting the occurrence of the dotted line.

- RQ3: To what extent did the endogenous network factors contribute to mediated social interaction?

3. Methods

This study drew on a secondary dataset generated from an online course at a large public university in the United States. The course was a graduate-level seminar about Learning Analytics that involved 14 students. This course was designed to support collaborative knowledge building that demands extensive, emergent social interactions among students. A bulk of the class conversations took place on Hypothes.is, and students were asked to treat their annotations as a collective knowledge base for their course projects.

Using the Hypothes.is open API, we collected a total of 1,160 annotation events from the class, including 629 annotations and 531 replies. From the dataset, we generated an edge list for the two-mode, actor–annotation network; in each edge *<s, r, t> s* stands for the sender—an actor/author, *r* the receiver—an annotation, and *t* the timestamp. Using the *rem* R package, this dataset was then structured as a discrete ordered sequence of relational events. Besides the network data, information about the annotations (e.g., location and word count) and the actors (e.g., project groups) was extracted.

In our empirical setting, reply actions can be generated only by an actor (i.e., a student) and directed toward an annotation. This is a one-plex, two-mode network, which has two types of nodes—actors and annotations, and one type of edge—reply. Different from a one-mode reply network that is typically constructed between participants, this two-mode network is based on an argument that a reply event from an actor to an annotation in the Hypothes.is environment also depends on traits of the annotation, even more so than characteristics of its author.

To address the specific research questions about factors driving social interaction in web annotation, we applied a novel network analysis approach named Relational Event Modelling (REM)[4]. Detailed explanations of REM can be found elsewhere [4]. Briefly, a relational event is defined as a "discrete event generated by a social actor and directed toward one or more targets"

[4]. The central goal of REM is to "understand how past interactions effect the emergence of future interactions" [4], based on a set of derived statistics about: (a) endogenous network factors reflected in past relational events (e.g., frequency of previous actor–annotation events), (b) node attributes (e.g., an annotation's length and location, an actor's gender, activity level), and (c) exogenous contextual factors (e.g., trust, student groups).

In this study, we used the rem R package to compute a variety of network statistics (elaborated below) in response to our research questions. Using these statistics, we trained multiple relational event models following a forward selection strategy and evaluated model adequacy based on the AIC (Akaike information criterion) score [4]. To model these computed statistics of relational events, we constructed a series of stratified Cox models using the *Survival* R package. We entered groups of variables in a stepwise manner and used the AIC score to evaluate whether the inclusion of new factors improved the models.

4. Findings

On average, students wrote about 80 total posts consisting of 45 annotations and 35 replies during the 14-week semester. Temporally, we observed posting behaviors skewed (about 64%) toward the window between Sunday morning and Monday evening, right before the class meeting on Monday evenings.

Outputs of the relational event models are reported in Table 1. As indicated by the AIC score, adding the contextual factor of student grouping in Model 2 did not improve the model whereas the exogeneous network factors added in Model 3 greatly improved the model.

The first research question inquired about the extent to which artifact attributes effected learner interaction with them. Specifically, we examined if an annotations' location, length, and inclusion of a question mark were predictive of the likelihood of being replied. We observed from the models that the relative location of an annotation was not predictive of the likelihood of being replied, neither was the inclusion of a question mark in the annotation. The word count was positively associated with the likelihood of being replied, indicating longer annotations were more likely to receive replies. However, this effect was small.

The second research question asked whether contextual factors, related to pedagogical design in this class, effected mediated social interaction. Results from Model 2 and 3 showed that the project group homophily (i.e., learners being in the same group) was non-significant. That is, learners in the same group have not interacted with one another more than with other peers.

The third research question inquired about the role of exogeneous factors, related to emergent activity and dynamics between learners. To this end, we added four factors in Model 3. Results showed that the learner activity level, prior popularity of annotation, and "four-cycle" reciprocity, had contributed to more future interactions. Location homophily factor showed a negative effect. These findings indicated that in this social annotation context, the more active a student was, the more popular an annotation was, and the more likely they were going to be involved in the next reply event. The positive and significant effect of the "four cycle" showed that students have an inclination to collaborate with prior collaborators (see Figure 1(c)). However, the negative and significant effect of the location homophily showed that when an annotation receives a reply, other annotations near this annotation are likely to receive fewer replies.

Table 1
Relational Event Models of Mediated Social Interaction in Web Annotation

	Model 1	Model 2	Model 3
Location of annotation	−.166	−.150	−.085
	(.169)	(.172)	(.180)
Question of annotation	.003	.002	.065
	(.108)	(.108)	(.112)
Length of annotation	.003	.003	**.003***
	(.001)	(.001)	(.002)
Project group homophily		.001	.004
		(.003)	(.003)
Actor activity			**.121*****
			(.017)
Location homophily			−**.010*****
			(.002)
Annotation popularity			**1.121*****
			(.075)
Reciprocity (four cycle)			**1.253*****
			(.160)
AIC	3677.903	3679.659	3269.597
R^2	.000	.000	.005
McFadden's pseudo R^2	.001	.001	.115

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

5. Discussion and Conclusions

The study was interested in examining mediated interaction and spatiotemporal dynamics in social web annotation. To foreground the mediated nature of social interaction in web annotation environments, we constructed a two-mode, bipartite network involving actors (students) and annotations. To examine network dynamics, we applied the Relational Event Modelling framework and modeled the impact of a number of factors on students' mediated social interaction.

This study also draws attention to the spatiotemporal properties of mediated social interaction in web annotation environments. The relational event modelling incorporates temporal information by design. Results showed network factors such as actor activity and annotation/artifact popularity played a role in the temporal evolution of the network. The addition of spatial information about annotations further allowed the investigation of temporal and spatial dynamics in tandem. Even though the annotation's spatial location was not predictive of network formation, we found a reply to an annotation would suffocate potential replies to nearby annotations. This mechanism might reflect the temporally condensed activities from individuals and temporally distanced activities among them [14]. In other words, it is plausible in the web annotation context that some students would log on in a particular time when other students are unlikely to be active, scroll through a stack of peer annotations, and selectively reply to a few in different locations. The revealed spatiotemporal properties of social interaction in web annotation are worth considering if there is an interest in promoting peer interaction through instructional interventions.

In conclusion, this study contributes fresh insights into social interaction in web annotation, calls for attention to micro-level spatiotemporal patterns, and calls for future work to investigate mediated social interaction as a dynamic network phenomenon.

References

[1] D. H. Jonassen, M. Davidson, M. Collins, J. Campbell, B. B. Haag, Constructivism and computer-mediated communication in distance education, The American journal of distance education 9 (1995) 7–26. doi:10.1080/08923649509526885.

[2] R. Sanderson, P. Ciccarese, B. Young, Web Annotation Data Model, Technical Report, W3C, 2017.

[3] X. Zhu, B. Chen, R. M. Avadhanam, H. Shui, R. Z. Zhang, Reading and connecting: Using social annotation in online classes, Information and Learning Sciences 121 (2020) 261–271. doi:10.1108/ILS-04-2020-0117.

[4] C. T. Butts, A relational event framework for social action, Sociological methodology 38 (2008) 155–200. doi:10.1111/j.1467-9531.2008.00203.x.

[5] B. Chen, Designing for Networked Collaborative Discourse: An UnLMS Approach, TechTrends 63 (2019) 194–201. doi:10.1007/s11528-018-0284-7.

[6] J. H. Kalir, Social annotation enabling collaboration for open learning, Distance Education 41 (2020) 245–260.

[7] C. C. Marshall, Annotation: From paper books to the digital library, in: Proceedings of the Second ACM International Conference on Digital Libraries - DL '97, ACM Press, New York, New York, USA, 1997, pp. 131–140. doi:10.1145/263690.263806.

[8] A. Y. S. Su, S. J. H. Yang, W.-Y. Hwang, J. Zhang, A Web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments, Computers & education 55 (2010) 752–766. doi:10.1016/j.compedu.2010.03.008.

[9] H. U. Hoppe, Computational Methods for the Analysis of Learning and Knowledge Building Communities, in: C. Lang, G. Siemens, A. Wise, D. Gasevic (Eds.), Handbook of Learning Analytics, first ed., Society for Learning Analytics Research (SoLAR), 2017, pp. 23–33. doi:10.18608/hla17.002.

[10] B. Chen, O. Poquet, Socio-temporal dynamics in peer interaction events, in: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, ACM, Frankfurt Germany, 2020, pp. 203–208. doi:10.1145/3375462.3375535.

[11] T. Hecking, I.-A. Chounta, H. U. Hoppe, Role Modelling in MOOC Discussion Forums, Journal of Learning Analytics 4 (2017) 85–116. doi:10.18608/jla.2017.41.6.

[12] O. Poquet, J. Jovanovic, Intergroup and interpersonal forum positioning in shared-thread and post-reply networks, in: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, ACM, Frankfurt Germany, 2020, pp. 187–196. doi:10.1145/3375462.3375533.

[13] D. Vu, P. Pattison, G. Robins, Relational event models for social learning in MOOCs, Social Networks 43 (2015) 121–135. doi:10.1016/j.socnet.2015.05.001.

[14] B. Chen, T. Huang, It is about timing: Network prestige in asynchronous online discussions, Journal of Computer Assisted Learning 35 (2019) 503–515. doi:10.1111/jcal.12355.