Towards a Visual Analytics System for Emotion Trajectories in Multiparty Conversations

Z. Huang[1], K. Kucher[1], and A. Kerren[1,2]

[1]Department of Science and Technology, Linköping University, Sweden
[2]Department of Computer Science and Media Technology, Linnaeus University, Sweden

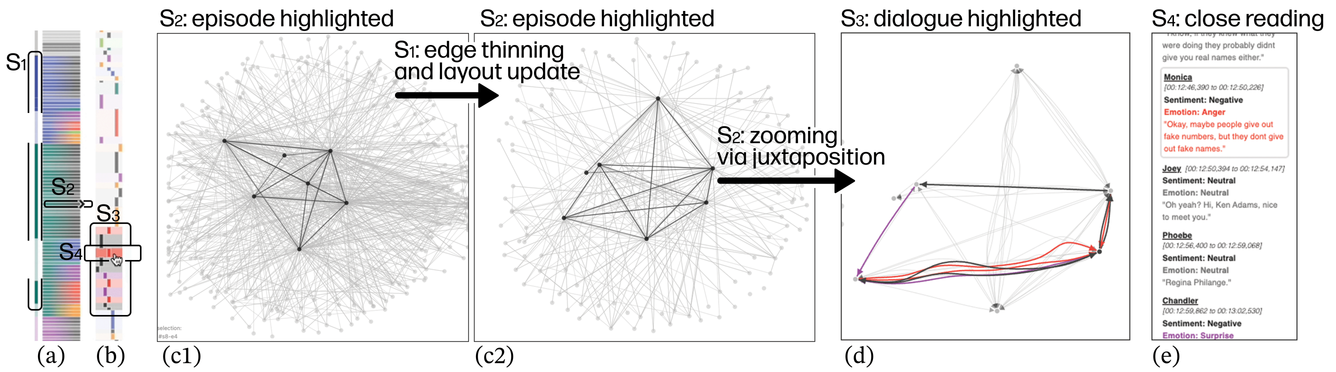

Figure 1: *Our tool enables users to explore speaker relationships in conversations through four selection strategies (S_1–S_4) across five views: (a) summaries of episodes, (b) a timeline of utterances on demand, (c) a speaker connectivity network with edge thinning, (d) a detailed mapping of emotions and utterances within a chosen episode, which is already highlighted in (c), and (e) close reading of utterances.*

Abstract

Visualizing sentiments in textual data has received growing interest; however, representing emotions within interlocutor relationships and associating them with the temporal progression of dialogues remains challenging. In this poster abstract, we describe the ongoing work on a visual analytics tool designed for analyzing emotion trajectories within dialogue collections composed of utterances from multiple speakers. The proposed tool provides exploration at different levels of detail to complex multigraphs, where edges represent direct responses between speakers through their utterances. Our approach includes several selection strategies for connecting different views: summaries of emotion transitions across dialogue groups, detailed analyses of individual utterances within specific dialogues of interest in interlocutor networks, and close reading. The tool aims to support model development in natural language processing by allowing users to explore text corpora interactively.

CCS Concepts

• ***Human-centered computing*** → *Visual analytics;*

1. Introduction

Visualization for natural language processing (NLP) has gained increasing interest that discovers patterns and provides insights into the semantics of language-related tasks. Among these, sentiment visualization, which emerges from opinion mining and affect analysis, has been one challenging research area. It involves identifying attitudinal content [MMSP14] in text and its associations with other linguistic features across different levels of detail [NSS21, HMJ*24]. Visualizing emotions in communication data, in particular, presents distinct challenges due to the sequential and interconnected nature of conversations [KPK18].

Many text visualization approaches often treat textual data as a unified block, focusing on summarizing statistics or analyzing feature embeddings [BBDW14, HWKK23]. These approaches may overlook the interaction sequences that produce the text. On the other hand, while some methods account for the temporal aspect by visualizing data as a series of exchanges between two parties, relatively less prior work addresses complex tree-like structures mining from multiparty dialogues, where responses can originate from any participant within the conversation. Moreover, visualizing emotion dynamics introduces further complexities as similar expressions

DOI: 10.2312/evp.20241092

can convey varied emotions depending on context, tone, and nonverbal cues, leading to uncertainties in analysis methods [CSZS23].

To address these challenges, various visual analytics approaches have been proposed that incorporate multimodal data. Video and audio information are visualized at various detail levels as panels or glyphs embedded in the text visualization and supporting main text analysis tasks [ZWW*20, WMW*22]. Furthermore, to capture the temporal dynamics of emotions, area charts linked with word frequencies are commonly used [SDB*15, HDP*21]. The use of network visualization techniques is often introduced to represent word clusters or sentence categories [SVS*23]. Most papers on emotion visualization focus on monologues or dyadic conversations. While there are many methods for network visualization of user interactions, they often do not support tracking down to individual utterances or provide multilevel navigation of the network based on textual categories of each node or link [HKK22, SLT*22].

In this poster abstract, we present our ongoing work designed for visual analysis of emotion trajectories in multiparty conversations. Our approach currently supports the tasks **T1:** navigating through a complex multigraph where nodes represent speakers and edges denote utterances, and **T2:** identifying dialogues of interest, visualizing emotion flows between interlocutors, and offering a close reading of the textual details. This tool aims to explore large-scale dynamic multigraphs across different granularities while associating textual and structural aspects of conversations, for researchers working on model development in NLP.

2. Implementation

The dataset we use for the current application is the Multimodal Emotion Lines Dataset [PHM*18, CHK*18], which consists of multi-speaker dialogues from the television comedy show *Friends*. We constructed a network based on the sequential utterance data, such that a connection is drawn whenever one character speaks immediately after another.

T1: Navigation The dataset results in a dynamic multivariate multigraph to visualize speaker relationships. It can be further divided into subnetworks by criteria like episodes/scenes/meetings, allowing detailed exploration of conversations in each episode and utterances in each conversation. A primary challenge identified is the clutter and responsive interaction difficulties in visualizing such complex data in a web-based environment. To address this, a multilevel navigation workflow was introduced for efficient filtering and navigation through the network. Users can apply four filtering strategies: multi-selection of episodes for edge thinning (S_1 in Fig. 1(c1–c2)), selection of a specific episode to isolate the local network (S_2 in Fig. 1(c–d)), selection of a conversation to see emotion distributions (S_3 in Fig. 1d), and selection of an utterance within the conversation to read the associated texts (S_4 in Fig. 1e).

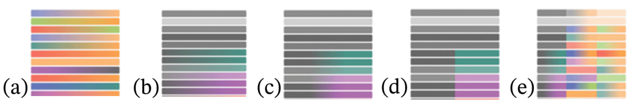

Figure 2: *The summary view (cf. Fig. 1a) interaction outcomes.*

Summary View To support the selection of one (S_2) or multiple episodes (S_1), our approach summarizes each episode to allow users to make decisions based on the summary (Fig. 1a), by calculating a Markov transition matrix of each episode's emotion sequence. The most probable emotional shift (e.g., from anger to joy) is the matrix cell with the highest probability value. This state change from one emotion to another is represented by a gradient bar consisting of two colors on the endpoint, following an emotion color scheme used throughout the entire visual interface with consideration of affective color guidelines [BPS17, AR22]. Additionally, the opacity of these bars reflects the confidence of transition occurrence. Users can organize the gradient bars either chronologically (Fig. 2a) or by specific emotion categories (Fig. 2b).

This feature directly links into the multi-selection interaction (S_1), where the brush size to perform edge thinning is determined by grouping episodes sharing the same dominant emotional transition (e.g., all episodes with the most probable transition being from anger to joy). Users also have the flexibility to fine-tune the gradient's smoothness (Fig. 2(b–d)) to enhance visual clarity or view multiple emotion transitions arranged by their likelihood (Fig. 2e).

Close Reading When a user selects an episode gradient bar (S_2), Fig. 1b displays a detailed timeline of all dialogues within that episode. In this timeline, each column signifies an individual speaker, and the length of the bars represents the duration of each utterance. The bars' color corresponds to the emotion, consistent with the established emotion color scheme. Additionally, the vertical positioning of each utterance bar corresponds to its occurring location within the episode's timeline. Hovering over a bar highlights all utterances from the same dialogue (S_3), which then leads to the colored links in Fig. 1d. Selecting an individual utterance bar (S_4) automatically scrolls the close reading panel to the textual content of that utterance for detailed analysis (Fig. 1e).

Network Representation The speaker network is drawn by a force-based radial graph layout, using one edge to represent multiple interactions between two speakers (Fig. 1c). Upon a multi-selection (S_1), only edges in selected episodes are shown. When there is a single selection (S_2), this currently selected episode subnetwork is pulled towards the middle and highlighted in black. Fig. 1(c–d) displays the network at different levels of detail side-by-side for comparative analysis, with (d) offering a zoomed-in view revealing every utterance as links between two nodes. A repulsion force is applied to each link to distinguish them better.

T2: Emotion Visualization Following the workflow outlined in T1, users can select dialogue from the episode group and then view the emotion trajectories in Fig. 1d. Additionally, Fig. 1a highlights the most likely emotion transitions per episode, Fig. 1b shows the emotions in a timeline view, and Fig. 1e displays the source text of interest.

3. Conclusions and Future Work

In this poster abstract, we have briefly introduced the ongoing work of creating a visual analytic tool for emotion trajectories in multiparty conversations. Future work includes training models to predict an emotion trajectory and visualizing the prediction results, performing sequential and structural pattern mining, and refining the layout further to minimize clutters and edge crossings.

Acknowledgements This work was supported through the ELLIIT environment for strategic research in Sweden.

References

[AR22] ANDERSON C. L., ROBINSON A. C.: Affective congruence in visualization design: Influences on reading categorical maps. *IEEE Transactions on Visualization and Computer Graphics 28*, 8 (Aug. 2022), 2867–2878. doi:10.1109/TVCG.2021.3050118. 2

[BBDW14] BECK F., BURCH M., DIEHL S., WEISKOPF D.: The state of the art in visualizing dynamic graphs. In *Eurographics Conference on Visualization (EuroVis) — STARs* (2014), The Eurographics Association. doi:10.2312/eurovisstar.20141174. 1

[BPS17] BARTRAM L., PATRA A., STONE M.: Affective color in visualization. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2017), CHI '17, ACM, pp. 1364–1374. doi:10.1145/3025453.3026041. 2

[CHK*18] CHEN S.-Y., HSU C.-C., KUO C.-C., KU L.-W., ET AL.: EmotionLines: An emotion corpus of multi-party conversations. *arXiv preprint arXiv:1802.08379* (2018). doi:10.48550/arXiv.1802.08379. 2

[CSZS23] CHEN F., SHAO J., ZHU S., SHEN H. T.: Multivariate, multi-frequency and multimodal: Rethinking graph neural networks for emotion recognition in conversation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), CVPR '23, IEEE, pp. 10761–10770. doi:10.1109/CVPR52729.2023.01036. 2

[HDP*21] HEALEY C. G., DINAKARAN G., PADIA K., NIE S., BENSON J. R., CAIRA D., SHAW D., CATALFU G., DEVARAJAN R.: Visual analytics of text conversation sentiment and semantics. *Computer Graphics Forum 40*, 6 (Sept. 2021), 484–499. doi:https://doi.org/10.1111/cgf.14391. 2

[HKK22] HUANG Z., KUCHER K., KERREN A.: Towards an exploratory visual analytics system for multivariate subnetworks in social media analysis. In *Poster Abstracts of the IEEE Visualization and Visual Analytics Conference* (2022), VIS '22. 2

[HMJ*24] HA H., MOON K., JOUNG H., KIM H., LEE K.: An exploration system to effectively analyze the visual metaphor used in sentiment visualization. *Information Visualization 23*, 2 (Apr. 2024), 157–196. doi:10.1177/14738716241228593. 1

[HWKK23] HUANG Z., WITSCHARD D., KUCHER K., KERREN A.: VA + Embeddings STAR: A state-of-the-art report on the use of embeddings in visual analytics. *Computer Graphics Forum 42*, 3 (June 2023), 539–571. doi:10.1111/cgf.14859. 1

[KPK18] KUCHER K., PARADIS C., KERREN A.: The state of the art in sentiment visualization. *Computer Graphics Forum 37*, 1 (2018), 71–96. doi:10.1111/cgf.13217. 1

[MMSP14] MUNEZERO M., MONTERO C. S., SUTINEN E., PAJUNEN J.: Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing 5*, 2 (Apr.–June 2014), 101–111. doi:10.1109/TAFFC.2014.2317187. 1

[NSS21] NARECHANIA A., SRINIVASAN A., STASKO J.: NL4DV: A toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (Feb. 2021), 369–379. doi:10.1109/TVCG.2020.3030378. 1

[PHM*18] PORIA S., HAZARIKA D., MAJUMDER N., NAIK G., CAMBRIA E., MIHALCEA R.: MELD: A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv:1810.02508* (2018). doi:10.48550/arXiv.1810.02508. 2

[SDB*15] STEED C. A., DROUHARD M., BEAVER J., PYLE J., BOGEN P. L.: Matisse: A visual analytics system for exploring emotion trends in social media text streams. In *Proceedings of the IEEE International Conference on Big Data* (2015), Big Data '15, IEEE, pp. 807–814. doi:10.1109/BigData.2015.7363826. 2

[SLT*22] SHI Y., LIU Y., TONG H., HE J., YAN G., CAO N.: Visual analytics of anomalous user behaviors: A survey. *IEEE Transactions on Big Data 8*, 2 (Apr. 2022), 377–396. doi:10.1109/TBDATA.2020.2964169. 2

[SVS*23] SEVASTJANOVA R., VOGELBACHER S., SPITZ A., KEIM D., EL-ASSADY M.: Visual comparison of text sequences generated by large language models. In *Proceedings of the IEEE Symposium on Visualization in Data Science* (2023), VDS '23, IEEE, pp. 11–20. doi:10.1109/VDS60365.2023.00007. 2

[WMW*22] WANG X., MING Y., WU T., ZENG H., WANG Y., QU H.: DeHumor: Visual analytics for decomposing humor. *IEEE Transactions on Visualization and Computer Graphics 28*, 12 (Dec. 2022), 4609–4623. doi:10.1109/TVCG.2021.3097709. 2

[ZWW*20] ZENG H., WANG X., WU A., WANG Y., LI Q., ENDERT A., QU H.: EmoCo: Visual analysis of emotion coherence in presentation videos. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (Jan. 2020), 927–937. doi:10.1109/TVCG.2019.2934656. 2

Proceedings published by Eurographics - The European Association for Computer Graphics.