Rendezview: An Interactive Visual Mining Tool for Discerning Flock Relationships in Social Media Data

Melissa Bica
Department of Computer Science
430 UCB
University of Colorado
Boulder, CO 80309-0430
melissa.bica@colorado.edu

Kyoung-Sook Kim
Artificial Intelligence Research Center
National Institute of Advanced Industrial Science and Technology
Tsukuba, Ibaraki 305-8568, Japan
ks.kim@aist.go.jp

ABSTRACT

Social media data provide insight into people's opinions, thoughts, and reactions about real-world events. However, this data is often analyzed at a shallow level with simple visual representations, making much of this insight undiscoverable. Our approach to this problem was to create a framework for visual data mining that enables users to find implicit patterns and relationships within their data, focusing particularly on flock phenomena in social media. Rendezview is an interactive visualization framework that consists of three visual components: a spatiotemporal 3D map, a word cloud, and a Sankey flow diagram. These components provide individual functions for data exploration and interoperate with each other based on user interaction. The current version of Rendezview can represent local topics and their co-occurrence relationships from geo-tagged Twitter messages.

CCS Concepts

•**Information systems** → **Spatial-temporal systems;** •**Human-centered computing** → **Information visualization;**

Keywords

Three dimensional visualization, spatiotemporal data analysis, flock phenomena, geosocial media

1. INTRODUCTION

Social media generates huge amounts of data, and researchers or organizations who analyze this data in search of social phenomena need tools to accomplish this efficiently and effectively. The purpose of this project is to create an interactive framework for visual data mining to enable finding patterns and relationships in complex social media data, particularly from Twitter.

International Conference for High Performance Computing, Networking, Storage and Analysis November 15–20, 2015, Austin, TX, USA

DOI: 10.475/123_4

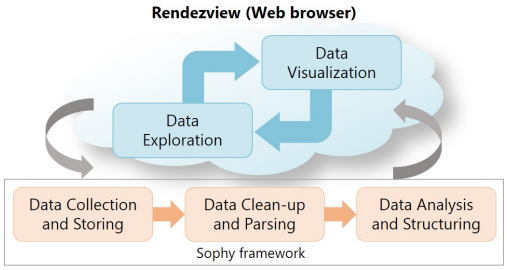

Figure 1: Rendezview overview.

The motivation behind Rendezview is to develop a web-based interface for visual data mining[3] to extend the functionality of the previous Sophy[4] framework which structures geo-tagged social media data (Figure 1). Our contributions are to create multiple, interactive data visualizations which represent geo-spatial, time, and topic data, and to improve user interactions with the data visualizations coupled with data mining processes. This makes the framework more usable and functional, allowing users to conduct more thorough data analysis.

2. DATA VISUALIZATIONS

The Rendezview interface is centered on three types of data visualizations: a spatiotemporal 3D map, a word cloud, and a Sankey flow diagram (Figure 2).

2.1 Spatiotemporal 3D Map

The 3D map is the primary visualization in the Rendezview interface. It is implemented using Three.js[5] and GeoJSON[2] data. Geospatial data is represented in the X-Y dimensions, while temporal data is represented in the Z dimension. Each of the cubes mapped on top represents a matching row from the database. Different colored boxes correspond to each keyword the user searches for, and different shades represent the measure of popularity. Boxes turn red when selected, and additional relevant information appears on the right side.

2.2 Word Cloud

The word cloud shows the frequency of keywords and Twitter hashtags used in conjunction with the search keyword. It updates to include the aggregate of all word frequency information when multiple boxes are selected. This was implemented using D3.js[1].

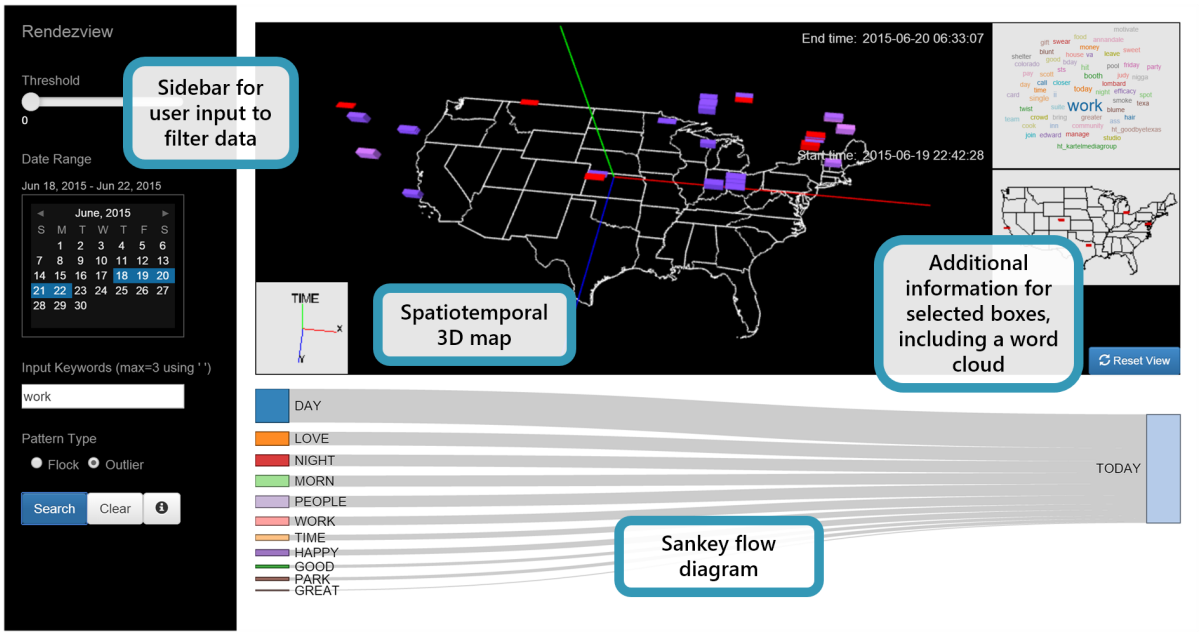

Figure 2: Rendezview UI layout, design, and functionalities.

2.3 Sankey Flow Diagram

The Sankey diagram consists of nodes and links, where nodes are keywords and links are connections between keywords. The width of the links shows the flow quantity, which is the value of the geospatial-temporal intersection of the two keywords. This was implemented using a D3 Sankey plugin.

3. INTERACTIVE DATA EXPLORATION

Visual data mining is the process of detecting patterns within big data using visualizations. The value of Rendezview comes from its ability to support *interactive* visual data mining. As shown in Figure 3, the data visualizations are not only user interactive individually, but interoperate with each other. The filtered data input is first visualized in the 3D map. When the user selects boxes in the map, the word cloud appears and represents data from the user-selected spatiotemporal area. The word cloud shows the relative frequency of words, which indicates to the user which keywords to filter the data on in subsequent searches. This also determines which keyword links the user should select from the Sankey diagram in order to display the most common keyword co-occurrences on the 3D map. These interactions allow the user to conduct a thorough analysis and find patterns hidden deep within the data.

4. CONCLUSIONS

We successfully implemented an interactive data visualization framework which allows users to view implicit flock patterns and relationships in social media. Future work includes making the system dynamic by connecting the interface to a live database. Additionally, improved interaction between the Sankey diagram and 3D map will be implemented to improve the user interaction and depth of visualization. This includes adding an option to view link widths according to various aggregation types and changing the Sankey nodes and links based on selected boxes on the map.

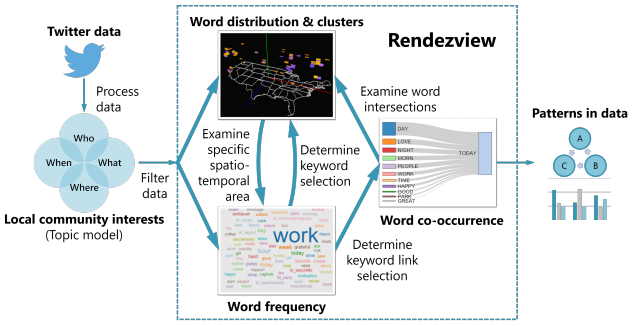

Figure 3: Interactive visual mining process.

5. ACKNOWLEDGMENTS

The authors of this project thank the Partnership in International Research and Education (PIRE) program, the Open Science Data Cloud (OSDC), the National Institute for Advanced Industrial Science and Technology (AIST), and the NSF for their support. Additionally, we thank Dr. Bob Grossman, Dr. Maria Patterson, and Dr. Jason Haga for their guidance.

6. REFERENCES

[1] D3.js. http://d3js.org/.
[2] GeoJSON. http://geojson.org/.
[3] D. Keim et al. Information visualization and visual data mining. *Visualization and Computer Graphics, IEEE Transactions on*, 8(1):1–8, 2002.
[4] K.-S. Kim, H. Ogawa, A. Nakamura, and I. Kojima. Sophy: a morphological framework for structuring geo-referenced social media. 2014.
[5] Three.js. http://threejs.org/.