

Creating Reflections in Public Emotion Visualization: Prototype Exploration on Traffic Theme

Wu Jiayu

Independent author

Room519, Building15,
Xinzhaojiayuan, Chaoyang
district, Beijing
zoewoo1211@gmail.com

Fu Zhiyong

Tsinghua University

B466, Academy of Arts and Design,
Tsinghua University, Beijing
fuzhiyong@tsinghua.edu.cn

Liu Zhiyuan

Tsinghua University

Natural Language Processing
Group, Tsinghua University,
Beijing
liuliudong@gmail.com

Lin Xu

Tsinghua University
B466, Academy of Arts and
Design, Tsinghua University,
Beijing
linxu8891@yahoo.cn

Tang Jiayu

Tsinghua University

Natural Language Processing Group,
Tsinghua University, Beijing
tjy430@gmail.com

Pan Jajia

Tsinghua University

B466, Academy of Arts and
Design, Tsinghua University,
Beijing
kaluojajia@gmail.com

Zhao Chen

Microsoft Research Asia
MSRA, Haidian district, Beijing
chzha@microsoft.com

ABSTRACT

The paper presents a visualization prototype that tackles the issue of public emotion. It emphasizes the exploration process in the aim of creating reflections for the viewer to observe the whole social context as well as individual perspectives. Urban traffic conversation on SNS (Social Networking Site) is our current interest. The prototype we present is made on real post streams captured from Chinese most popular microblog Sina Weibo. After experimenting on two online prototypes we determine a visualizing flow to lead the viewer going through the insights from macro to micro view in three interaction frameworks: City Sentiment, Related Topics, and Post Content. Through showing the visualizing flow with interaction mode, data analysis, and prototype construction, the paper ends with discussing about design considerations in creating reflections on public traffic emotion in visualization prototype.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Author Keywords

Traffic emotion visualization; Social networks; Reflective mind flow creation

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI):
Prototyping.

General Terms

Design; languages.

INTRODUCTION

Card et al. define information visualization as to amplify cognition by using computers to provide interactive visual representations [1]. In our exploration on prototype making, amplify cognition means to embed insights in the visual representations for sense making rather than yield visual elements for decision making. Our way to amplify cognition in traffic emotion visualization is to build a delicate visualizing flow to generate reflections from knowing the general evaluations to contemplating individual insights. The research focus is to create a mind flow to guide the viewer to stroll between macro and micro level of understanding for public traffic emotion. The prototype exploration carries out by taking care of this research problem.

In the exploration of creating the visualizing flow, one

challenge emerges to us, that is, to what extent the visualization prototype should interpret the data. We aim to create a mind reflecting space for traffic issue, in which the viewer gets to see through the data and finds personal connections in the whole context that consists of others' opinions. We explore the way to involve the viewer as an information analyst instead of an information receiver. Therefore, we present the data a bit far from the literal translation by adding some social perspectives. In this way, the viewer gets to think of traffic issue in a wider social context. Meanwhile, to avoid driving viewers too far from their familiar surroundings, we make the interactive presentation by taking self-involvement into account. A mind flow is considered in the visualization prototype to guide the viewer observe, analyze, and reflect on the data interpretation. Through these steps the viewer forms and corrects his opinion by reflecting on the perspectives between general evaluation and personal attitude.

We will introduce the prototype by looking at the reflective mind flow creation in the next sections. Visual Presentation Exploration mainly focuses on interface presentation and interaction mode. Prototype Construction describes data construction and visualization front-end building. We finally discuss the design considerations about data interpretation that reflected in the experiments of building the visualization prototype on public traffic emotion.

VISUAL PRESENTATION EXPLORATION

In the prototype exploration, studying the connectivity between visualization presentation and data interpretation, and looking into the possibility that the traffic conversation data can contribute to urban public emotion visualization are the main tasks. We intend to create a reflecting process for citizens to observe, analyze, and reflect on the visualized facts that the prototype provides. We experiment a mind flow including three steps: getting to know the general traffic emotion across China in cities, analyzing the related topics around traffic and the emotion people talking about in each city, observing the individual post content by interacting with the self-filter options. In order to achieve this mind flow we create an interaction framework in three types of data interpretation: City Sentiment, Related Topics, and Post Content. Since public emotion is the base to inspire reflections in our prototype, we make emotion as the clue in every view of the visualization presentation.

City Sentiment

City Sentiment locates in the first level of the interaction framework. In terms of the user feedback obtained from the first prototype experiment, we found that it's crucial to present the continuity between the general impression of cities' emotion across China and the related topics of each city. Otherwise the viewer tends to lose the clue when switching the view from City Sentiment to Related Topics. The feedback also indicates that the horizontal view directly showing the comparison of city's emotion is interesting to

the viewer. Considering the above findings, we design two views in this interaction level: Score View and Comparison View (see Figure 1 & 2).

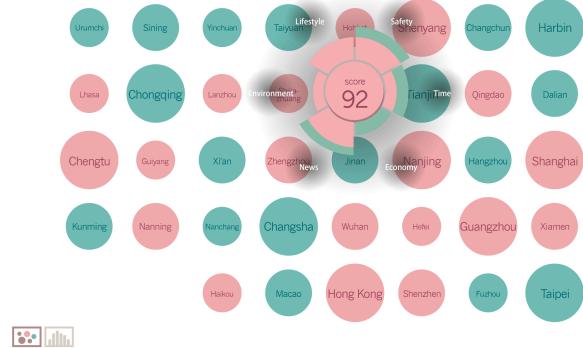


Figure 1. Score View

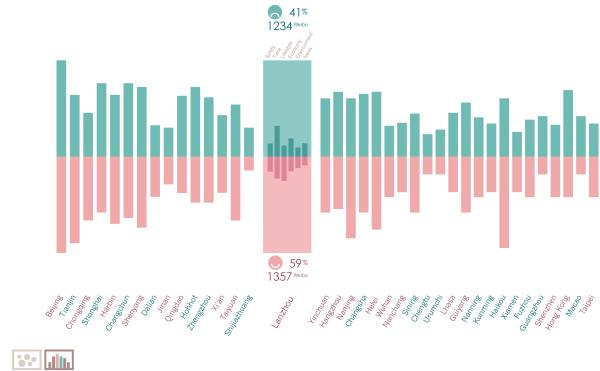


Figure 2. Comparison View

City Sentiment presents 38 major cities' general emotion in China. In Score View (Figure 1), the emotion of each city shows in colors. Warm hue indicates positive sentiment while cool hue negative, if above fifty percent people feel positive for urban traffic in a city, the color of the circle becomes warm hue, and vice versa. The total number of the posts in each city shows as the size of the circle, the bigger the circle, the more the posts. Score displays on top of the circle when the cursor hovers, representing the percentage of positive emotion in the total number of the posts in the city; at the same time, six related topics show up around the circle, they are presented in colored sectors implicating the proportion of the emotion for each topic. In this way, we hope to let the viewer understand how the general emotion for city traffic comes up, what forms the general emotion and how important each fragment of the traffic topic is in the whole traffic emotion. Comparison View (Figure 2) is another view to present cities' general emotion. It's designed for comparing cities' emotion horizontally. The positive and negative emotion displays in bar charts with a line in the middle separating the two poles of the emotion. The length of the bar indicates the amount of the posts in the city. When the cursor hovers over one bar, six related topics' emotion shows up in small bars inside the city bar,

descriptions such as how many posts are with positive mood and what percentage it holds in the whole posts are marked by the end of the bar.

Through City Sentiment we hope to make the viewer understand how the general sentiment comes up by observing the composition of city's emotion and analyzing the proportion of the fragments. This analysis paves the way for the viewer to step towards deeper reflections.

Related Topics

Related Topics view initially intends to present the most popular topics in each city. In the first version of visualization prototype we present ten top topics in each city by random, which is elaborated in another paper City Flow: Prototype Exploration for Visualizing Urban Traffic Conversations [2]. By observing that online prototype we found a variety of topics talked by people about urban traffic. We collect those topics and categorize them into six themes. According to the social impact the topics belong to, we determine the traffic related themes are: safety, lifestyle, environment, news, economy, time. These six topics basically encompass every aspect in citizen's traffic life. Safety theme includes traffic accident, rules, weather issues, etc. Lifestyle theme refines from citizen life such as culture, development, life quality, which are associated with urban traffic in a macro perspective. Environment theme comes from a word frequently appears in the first version of prototype: environment protection, which we assume has intimate relationship with urban traffic that deserves to be stressed in a single theme. News theme covers hot topics and emergent events that happen to be discussed currently. Economy theme is about costing, sustainability, price, and so on.

Related Topics (Figure 3) is the second level of the interaction framework, viewers enter this view by clicking a city circle in Score View or a city bar in Comparison View. The blocks lining up on top of the page show city names. The viewer can switch the city by scrolling the city name blocks to the right and left. The current chosen city is highlighted in order to remind the viewer about where they are. Six themes are illustrated in circles with size differences indicating the sum of the posts people talking about this theme, and color distinctions indicating the emotion of the theme. The ring located outside of the circle shows the proportion of bad and good mood when people talking about this theme. When the cursor hovers over a theme circle the statistic information of the population talking about this theme appears in the space in below, such as gender, how many posts (s)he posted, how many followers (s)he has. Post Master shows the emotional trend for the posters in five levels: who have 0-200 posts, 200-500 posts, 500-1000 posts, 1000-2000 posts, and more than 2000 posts. Popular User shows the emotional trend for the posters who have followers in five levels of 0-500, 500-1000, 1000-5000, 5000-10000, and more than 10000. The

colored points below indicate the proportion of the emotion in each level.

Through Related Topics we hope to introduce the viewer an analyzing process in related social themes for urban traffic in each city. In the aim of creating reflective mind flow, we assume that the self-involvement is important when people participate in social topic. This view is designed for viewers to build a personal connection to the visualization prototype, and to prepare for the next step into a micro view of interaction framework: Post Content.

[Beijing](#) [Shanghai](#) [Nanjing](#) [Hefei](#) [Changsha](#) [Wuhan](#) [Sining](#) [Nanchang](#) [Jinan](#) [Chengdu](#) [Fuzhou](#) [Unmuchi](#) [Xian](#) [Taiyuan](#) [Lhasa](#)

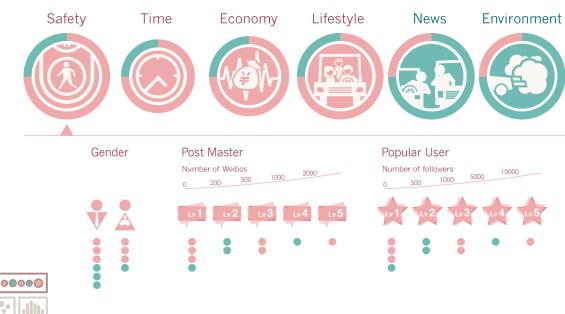


Figure 3. Related Topics

Post Content

Post Content (Figure 4) is designed to show personal view on how citizens really talk about traffic. Based on the feedback from the first prototype that viewers tend to look into this view with taking themselves into account, we determine that Post Content view should provide original post and personal filter mechanism.

On the right of this view there is a big space showing the original post contents. We use Dynamic Grid Layout to display the post with the content, the image, the poster's nickname, and the time of posting. The organization of the posts is in time order. On the left of this view it lists out the filter options: gender, the level of the poster regarding the posts (s)he put online, and the level of the poster regarding the followers (s)he has. The system provides multi-option. When the viewer chooses the options on the left the posts on the right will be filtered. The upper left corner places the six traffic themes in circles, with which the viewer gets to switch to other themes.

In this level of interaction framework it sets a close observation to traffic conversations by providing filter options. This design adds interactive pleasure to the viewer and involves the viewer into the individual attitudes by selecting the posts. In the visualizing flow this view shows the concrete facts, which encourage the viewer to compare, judge, think, and rethink on the previous abstract data interpretations. In this way, the viewer gets to form a new opinion.



Figure 4. Post Content

PROTOTYPE CONSTRUCTION

The prototype consists of three parts including data mining, sentiment analysis and data interaction between front-end and back-end. First, the posts talking about traffic are filtered out from all kinds of topics on Sina Weibo. Second, sentiment analysis is carried out on these traffic posts and keywords are generated for each city. All these organized data is maintained using an XML file to provide structural data for the visualization front-end. The data is captured and analyzed during the last 24 hours. The online prototype updates data every 10 minutes.

Data Mining

Data is obtained via APIs provided by Sina Weibo. Posts that contain one of the following keywords “traffic”, “transit” or “road condition” are filtered as the raw data. In our filters, posts that refer to real traffic conditions are distinguished from advertisements and non-traffic phrases. The first step is to filter out non-traffic phrases like “the Bank of Communications” or “the University of Communications”. The word “communication” in Chinese is identical to “traffic”. So these phrases contain the exact words of “traffic” but are in fact not related to the topics that we concern about for visualization. The second step is to cut off advertising posts with a user ID containing words like “real estate” or “group purchase”. With these filters, the posts that contain comments made by users about traffic are picked out for further analysis.

We then clean these posts and only keep plain text data. Most plain text data is in Chinese. Hence, we will perform Chinese Word Segmentation (CWS) and part-of-speech (POS) tagging before sentiment analysis and keyword extraction. We perform CWS and POS tagging using THULAC, a practical system developed by natural language processing group at Tsinghua University [3].

Sentiment Analysis

Sentiment extraction defines the emotional aspect of each post. A list that composed of positive and negative words is constructed. Each word in this list has a value for measuring emotion. The post is split into single words

ready for emotional evaluation by the values of the word in the emotion dictionary. In this way, every post is assigned an emotion value and the sentiment of a city is delivered.

For each city, we predefine 6 aspects of the initial data, including “safety”, “lifestyle”, “environment”, “news”, “economy” and “time”. For each category we construct a word list. The posts containing the word in the list of these categories are labeled to deliver the sentiment of each aspect of a city.

Emotion dictionary construction is essential for measuring post’s sentiment. Existing emotion dictionaries are too small for traffic themes. For example, “lovely” is a positive word in general situation, but it seems to be appropriately considered as neutral word in traffic conversations. Therefore, it is important to improve the emotion dictionary in terms of traffic issues. The method employed in this project for improving the emotion dictionary is to measure the similarity and common information appears in a word regarding the existing words in the emotion dictionary and the traffic keywords. The data used in the experiment for improving the emotion dictionary is the total posts during a week on Sina Weibo. By using this method, the emotion value of the existing emotion words in terms of traffic themes can be defined.

Data Interaction Between Front-end and Back-end

In terms of the massive data the system has to deal with, to keep it light is the key consideration for data interaction between front-end and back-end. The data that front-end obtains from the back-end lies in two parts: a) the general city emotion and the emotion for the categories; b) the original post contents.

The general city emotion represents information in a macro view, it requires relatively low level of real-time data interaction, so that we update the results and format a XML file in the back-end every 10 minutes. The XML file is read directly by the front-end.

In Post Content view there are filter options and information loading when scrolling down the page, so that it requires a high level of real-time data interaction. And with the abundance of data, we consider that letting the front-end use Ajax to send requests to the back-end and acquiring the corresponding data in the database is an appropriate method. The interface we design for acquiring data from database is flexible. The parameters for selecting include city id, category id, filter options (gender, post master, popular user), and the number of the posts. In terms of these requests, the back-end retrieves the JSON formatted post contents in order to let the front-end to resolve.

DISCUSSION

Through experimenting on the prototypes to create reflections in public traffic emotion visualization, we found that the principal challenge on data interpretation is to

balance the presentation between the overall inclination of the city emotion and the massive pieces of personal attitudes. Two design considerations deserve to be discussed: one is building the mind flow between the general and individual emotions, another is creating self-involvement for the viewer to the visualization prototype.

In order to build a smooth mind flow to inspire reflections for the viewer, we believe that presenting the logic to connect general emotion and individual attitudes is the key. In this case, we set up analytic mind flow between views to remind the viewer the connectivity between collective and individual emotion. To create self-involvement, providing the statistic information as the filter options would be benefit for the viewer to take care of others' opinions by comparing with one's own. One of the design challenges in social topic visualization is to arouse viewers' personal feelings and to help them build their own views through observing the data.

ACKNOWLEDGMENTS

We would like to acknowledge Microsoft Research Asia as the project founder. Special thanks go to Long huiling, Ran Chenwei and Chen Xinxiang, who contribute a lot of efforts in the project.

REFERENCES

1. Card S. K., Mackinlay J., and Shneiderman B. Readings in Information Visualization, Using Vision to Think, chapter 1, pages 1–34. Morgan Kaufmann, 1999.
2. Jiayu Wu, Zhiyong Fu, Zhiyuan Liu, Jiajia Pan, Huiling Long, Xu Lin, Haoqing He, Xinxiang Chen, Jiayu Tang. City Flow: Prototype Exploration for Visualizing Urban Traffic Conversations. 2012 ASE/IEEE International Conference on Social Computing, Pages 481-489.
3. Zhang K and Sun M, A stacked model based on word lattice for Chinese word segmentation and part-of-speech tagging, <http://nlp.csai.tsinghua.edu.cn/thulac>