SOCIAL SENSOR: AN ANALYSIS TOOL FOR SOCAIL MEDIA

Chun-Hsiao Wu,
National Chengchi University, Taiwan
shaldrem@gmail.com

Tsai-Yen Li,
National Chengchi University, Taiwan
li@nccu.edu.tw

ABSTRACT

In this research, we propose a new concept for social media analysis called *Social Sensor*, which is an innovative design attempting to transform the concept of a physical sensor in the real world into the world of social media with three design features: *manageability*, *modularity*, *reusability*. The system is a case-centered design that allows analysts to select the type of social media (such as Twitter), the target data sets, and appropriate social sensors for analysis. By adopting parameter templates, one can quickly apply the experience of other experts in the beginning of a new case or even create their own templates. We have also modularized the analysis tools into two social sensors: *Language Sensor* and *Text Sensor*.

A user evaluation has been conducted and the results show that *usefulness*, *modularity*, *reusability*, and *manageability* of the system are all very positive. The results also show that this tool can greatly reduce the time needed to perform data analysis, solve the problems encountered in traditional analysis process, and obtain useful results. The experimental results reveal that the concept of social sensor and the proposed system design are shown to be useful on the big data analysis of social media.

Keyword: Social Sensor, Social Media, Data Science, Big Data Analysis, Automated Analysis

1. INTRODUCTION

1.1 Research Motivation and Purpose

The rise of social media networks established a new style of network structure, and the policy of opening data led the data analysis of global social media frenzy. The rapid growth of a huge amount of social media data today, led many people and resources invested in the task to collect, filter, store, process, and manage a huge amount of complex raw data quickly. In addition, the growth rate has been much greater than the speed that human expert can analyze.

Ben Lorica (2013) [1] has pointed out that data scientists tend to use a variety of tools, often across different programming languages, to process big data. Workflows that involve many different tools require a lot of context-switching, which affects productivity and impedes reproducibility. Similarly, the analysis workflow in a data science team usually adopts a sub-role, sub-field approach to the entire analytical work. In the team, some engineers are responsible for crawl data, and then transfer the data to processing data operations performed by other engineers. Then the output data were explored and established by data analysts, who are then followed by data visualization staff to present analysis results to the domain expert for interpretation.

In this research, we studied the data analysis process of a research team in National Chengchi University in analyzing social media. We use the case study of "2012 Taiwan presidential election" [16] to illustrate the problems in a typical analysis process such as unmatched speed of data analysis with the speed of data collecting, independent and fragmented analysis steps, labor intensive manual analysis, manual file exchanges, lack of data and case management tools, difficulty to maintain domain expertise, high restart costs, and long waiting time, etc.

Therefore, in this research, we propose a new concept for social media analysis called "Social Sensor," which is an innovative design attempting to transform the concept of a physical sensor in the real world into the world of social media with three design features: *manageability*, *modularity*, *reusability*. This tool is used to help users collect a variety of social media data for rapid analysis and visual presentation. By developing the concept of social sensors for a massive amount of social media analyses, we aim to provide an analytical tool in the applications of social science studies for policy making, marketing, and design of intelligence interface.

1.2 Concept Explain and Design Goals

This research attempts to bring the concept of physical sensors to the virtual world of social media, shown in Figure 1. Physical sensors in real-world environment are like social sensors in the social media environment, and they all produce a great amount of environmental data. Physical sensors measure and produce physical or chemical data, such as light, heat, temperature, humidity. The equivalence collected from social media by social sensors may include tweets, users, locations, and so on. More specifically, examples of physical sensors include luminosity sensor, temperature sensor, etc., whereas social sensors may include language sensors, text sensors and the like. Social sensors also function similarly as for physical sensors because different sensors are designed to have their own sensing tasks and applications. A social sensor is implemented with various analytical methods and generates derived data for a specific analytical task.

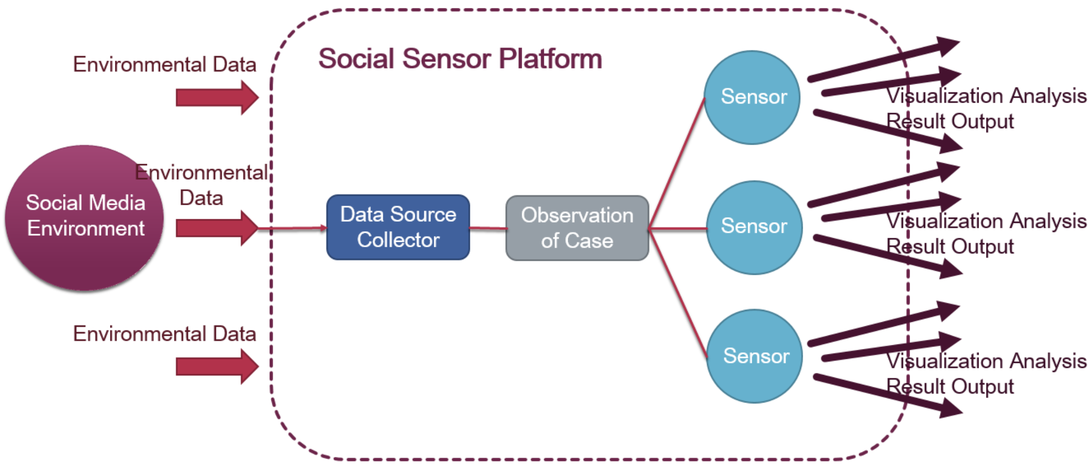

Figure 1. Social Sensor Conceptual Model

In addition, some physical sensors may contain parameter settings that can be used to adapt to different applications, and the settings of these parameters usually are from prior knowledge and know-how experience. Similarly, in the social sensor design, through parametric design, a same sensor can be customized to meet the needs of a specific application. The experience can be passed from one case to another similar one in a different research by copying the set parameters as a starting setting to save the time of trial and error from scratch.

2. RELATED RESEARCH

2.1. Social Media and Social Network Analysis

Social Media is defined in Kaplan and Haenlein's research [5] on the Internet for a group of applications featuring User Generated Contents (UCC) and media richness. Examples of Social Media include Twitter, Facebook, YouTube, Blog, Plurk, Flickr, and more. It can be viewed broadly as a service platform for people to share ideas, insights, experiences and perspectives through various forms of media with each other. These applications are built on the concepts and techniques of Web 2.0 and allow the User Generated Contents (UCC) to be established and exchanged. UCC is usually applied to describe the various forms of media content that are publicly available and created by end-users [7].

Social Network Analysis is a set of norms and methods to study the relationship between structure and properties. Mark Stelzner [8] think Social Media can be called an outlet for strategy and broadcasting while Social Networking is a tool and a utility for connecting with others. Yuan et al. [3] use semantic network analysis to calculate the frequency of words and phrases, the co-occurrence conditions, and the distance between words. The algorithm divides the data into different network clusters and uses the semantic network diagram to do the study. Kwak, et al. [6] use statistics, interactive link analysis, and separation index to investigate the Tweeter environment and its characteristics on information propagation.

In addition, there exist many tools for Social Media Analysis. For example, "yourTwapperkeeper" is a tool can be used to track more than one set of keywords or hashtag. Others include the statistical package of "Microsoft Office Excel", text processing software packages such as "Leximancer" and "Wordstat," and network analysis and visualization tools such as "Gephi." However, no single system or platform can serve as the best tool for all users since the needs in different application domains may also have great differences.

2.2. Application of Social Network Analysis

The techniques of Social Network Analysis (SNA) are quite extensive and its applications are also quite diverse. For example, some of the common techniques including data fusion and data mining, network communication modeling, user attributes and user behavior analysis, interaction analysis based on location, social sharing and filtering, recommendation systems, etc. These techniques can applied to business practices, such as customer interaction and analysis, marketing, and business intelligence. It can also be applied to information collection and consolidation, counter-intelligence, and law enforcement activities. There are also studies on food safety that make use of SNA techniques to build a query platform for plasticizer tests. It also be applied to "political elections." For example, Burgess and Bruns [4] use tweeter data to study the 2010 Australian Prime Minister Election. Some people use the SNA techniques to study catastrophic events. For example, Vieweg et al. [11] investigate how to capture relevant important messages by information extraction technologies. Shih [14] use machine learning techniques to process real-time new media contents for automatic classification.

2.3. Social Sensors

"Social Sensor" is a new concept derived from physical sensors. However, the concept has been defined by various researchers about social media but with different meanings. In this research, the innovative concepts and design of Social Sensor is quite different from the previously proposed ones.

Vasudevan et al. [12] applied the concept of social sensor to television programs. The "Social Sensor" design in this work is through the discussion and dissemination of real-time social media (such as Tweeter) to find out the hot peak point of social media in order to recommend the given TV program. Although this mechanism can potentially lead a large amount of social network users to participate in real-time TV show, it is only suitable for interstitial viewing.

Sakaki et al. [9] used the field characteristic of Twitter to study real-time interaction during an earthquake event. They used a keyword classifier for tweets, the number of words, and their contexts to build a model for computing the event location center and trace probability. They considered each Twitter user as a sensor and applied Kalman filtering and particle filtering for location estimation. Following this work, in 2013, Sakaki et al [10] proposed a method to extract real-time traffic information from social sensors in social media. Social sensors extracted information from social media through text-based classification, geographical coordinates, and important traffic events to provide real-time driving advices.

The aforementioned social sensors attempt to use the real-time characteristic of social media and to transform the given data into useful information in specific applications. The main body of social sensor can be referred to a data analysis method or to the user of social media. In contrast, social sensor in this research des not only refer to the design of a specific function of the sensor, but also include the features of pluggability and manageability for constructing a given observation of social media from a wide range of sensors. More specifically, in this research, our social sensors have three features: *manageability*, *modularity*, *reusability* to satisfy the needs of multidisciplinary analysis, plug-and-play, experience transfer through reuse, and case management.

3. SYSTEM CONCEPTS AND DESIGN

3.1. System Concept

In this research, we have used the research team on social media analysis at National Chengchi University as an example to understand the typical analysis process and find the current problems in the process. The process typically consists of three phases: data preparation, data analysis, data presentation. The data preparation phase typically includes the following steps: event occurrence, keyword collection, data storage and management, waiting for harvest. The data analysis phrase includes data preprocessing and data analysis while the data presentation phase includes data visualization and interpretation. We have used the concept of social sensor to streamline the steps in these three phases. Each sensor adopts a combination of the analytical methods in these steps and works independently. In addition, we try to separate data sets from observation cases for better manageability and reusability. In other words, the system uses a case-centered design, shown in Figure 2, which allows analysts to select the type of social media (such as Twitter), the target data sets, and appropriate social sensors for analysis. By adopting parameter templates, one can quickly apply the experience of other experts in the beginning of a new case or even create their own templates.

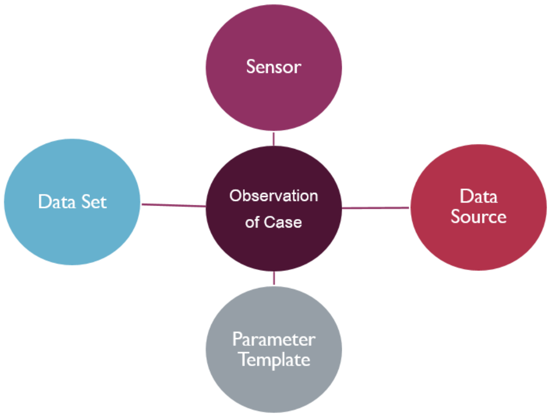

Figure 2. Social Sensor Design Concept

The design of our social sensor has three features: *modularity*, *reusability*, and *manageability*. First, modularity means that we hope each sensor is a reusable module derived from analytical methods. Second, in addition to the reusability resulting from modular design of social sensors, another aspect of reusability comes from the reuse of case study. Parameter settings in a typical case can be reused through the parameter template mechanism. Third, manageability provides an easy-to-operate analysis environment for establishing and managing the observed cases, parameter template, the selection of sensors, and the ability to save and publish the results.

3.2. System Architecture Design

The design of our system architecture is shown in Figure 3. We have dismantled a social sensor into components, each of which has its own task and goal. All components collaborate together to construct the whole social sensor architecture. In the following subsections, we will describe each of these components in more details.

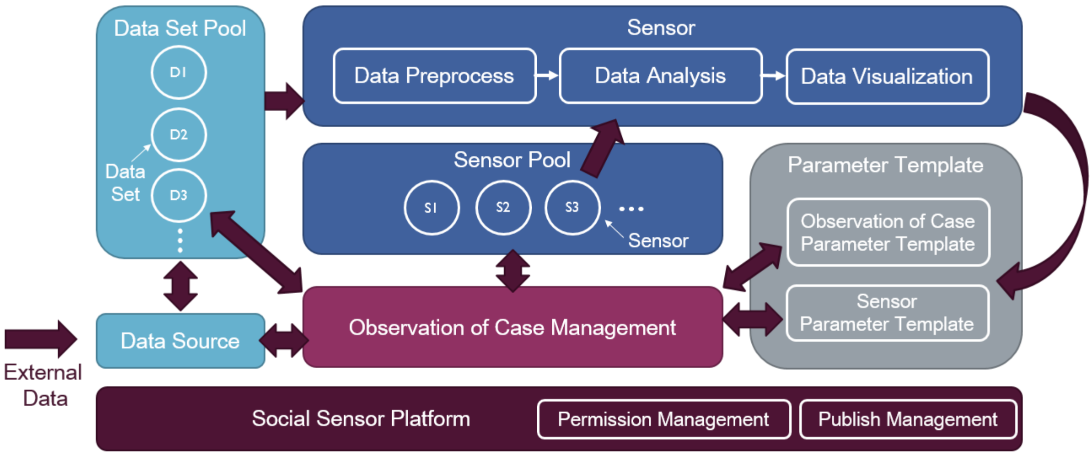

Figure 3. Social Sensor Architecture Design

3.3. Social Sensor Platform

Our social sensor platform uses a Web-based client-server architecture consisting of three tiers: *Presentation*, *Business Logic*, and *Data Services*. As in most client-server systems, the storage and computing resources are placed at the server side, and the client side usually operates through a browser. Some part of data visualization computation is handed over to the client side to speed up the operation. In term of software development methodology, we have adopted an object-oriented programming (OOP) paradigm to make the programs reusable, easy to maintain and extend.

In the platform design, we have adopted a role-based authorization mechanism with two types of users: system administrators and researchers. System administrators are responsible for managing system operations while researchers create observation cases, conduct research and publish results through analysis and visualization. In Figure 4, we show the snapshots of the implemented platform in action in various stages of the operations in a typical case analysis: the creation of an observation case, case management, sensors on-demand, chart analysis operation, parameter adjustment, sensor setup and tuning, setup of parameter template and other functions.

Figure 4. Snapshots of the functions provided in our social sensor platform

3.4. Observation of Case Management and Observation of Cases

In a typical observation case, the researchers usually choose an event as a research theme that they are interested in further investigation. Assume that a researcher is denoted as R, an observation case is denoted as p, and P_R is defined as the set of cases p‘s observed by R. In other words, $P_R = \{p_1, p_2, \cdots p_{n-1}, p_n\}$ represents the cases observed by R. An observation case p is defined with several parameters. For example, the data source is denoted by ds, the data set is denoted by d, the set of social sensors is denoted by s=($s_1, s_2, \cdots, s_{n-1}, s_n$), start and end times of the data set define the time interval for analysis are denoted by T_s and T_e respectively. While *ds* and *d* are single-value variables, *s* can have multiple choices. Researchers can ask “what-if” questions by performing experiments of data analysis repeatedly with different sets of parameter settings. What a desired analysis result is acquired, the researchers can save the current settings as a template to reproduce the same results in the future.

3.5. Data Sources

Today, a wide range of social media is available, and their data types and analysis goals are usually different. Data sources refer to the type of social media such as Twitter or Facebook. In this research, we have used Twitter as an example of data source to establish the observation cases. For a given data source, there could be multiple entries of data sets for analysis. For example, here are two examples of data sources with multiple data sets associated with: $ds_1\{d_1, d_2, d_3\}$ and $ds_2\{d_4, d_5\}$.

3.6. Data Set Pool and Data sets

A data set is acquired from a data source through the data collection process with specific query conditions such as hashtags, and time intervals. A *data set pool*, denoted by dp, is a common place for storing data sets in an observation case. In this research, we use database to store the data sets. In each observation case, one can change the data source by selecting an appropriate social media type and then the data sets from the given data set pool.

3.7. Sensor Pool and Sensors

Social sensor, denoted by s, is an analytical module designed for a specific analytical purpose. Each social sensor is designed based on a modular approach to use various analytical methods to satisfy the demands of social media studies. To be self-contained, though not required, a social sensor typically is designed with appropriate visualization to facilitate the analysis. Each sensor is independent and does not affect other sensors. Each sensor in the design must include three modules: *data preprocessing*, *data analysis*, and *data visualization*. In addition, the three modules need to be streamlined to avoid any manual external data exchanges.

A sensor s starts the analytical computation by first entering the data processing stage in which obtain d from dp and then acquiring the data in the intervals between T_S and T_e set in p. Then depending on type of analytical method used in the sensor, the desired data features, such as content texts in tweets, will be extracted automatically and sent to the next stage: *data analysis*. In the data analysis stage, based on the algorithmic design of the sensor, data will be further analyzed and save back to the independent storage for the sensor. In the third stage, the derived data after the analytical processing will be visualized according to the type of data and the design of the sensor. Usually, interactive queries and real-time display are preferred for the instant feedback. In addition, since the data analysis stage is the most time consuming stage, the computation in this stage is set in background by default such that the users can proceed to visualize previous or partial results while the new analytical results are under ways.

Sensor Pool, denoted by sp, is a place to store available sensors $\{s_1, s_2, \cdots, s_{n-1}, s_n\}$, from which p selects the desired sensors for analysis. Each sensor will be assigned specific parameters that can be adjusted in a given p for its analytical objective.

3.8. Parameter Templates

Once we have the observation cases and sensors parameterizable, for a same case, different researchers may choose different analytical methods, and for the same analytical method, the researchers may even choose different parameter settings based on their experience and exploration. In our system, we hope to retain the sensors and parameter settings of a case conducted by a specific researcher in order to replicate the experience of experts and lower the barrier for obtaining meaningful analytical results when starting up a new case. In our design, the parameters of a case p can be built into a parameter template, as shown in Figure 5, which includes settings such as ds (data sources), d (data sets), T_S (start time for analysis), T_e (end time for analysis), the set of selected sensor groups, $\{s_1, s_2, \cdots, s_{n-1}, s_n\}$, and the parameter settings for each individual sensor. For the last part about the parameters of a sensor, one can choose to save the settings into a parameter template for the sensor for future reuse.

In our system, one can also set the access rights of each individual template (case or sensor) as private or public. The default access setting for a template is private but can be changed to public access when the users consider appropriate.


```json
{
    "name": "2012總統大選分析樣板_Shaw",
    "tag": "台灣選舉",
    "dataSourceId": "1",
    "dataSourceName": "歷史Twitter資料",
    "dataSetId": "1",
    "dataSetName": "2012總統大選資料集",
    "sensorsId": "1,2",
    "sensorsName": "語系分析,文字網絡分析",
    "sensors": [{
        "id": 1,
        "name": "語系分析",
        "createTime": "2014-04-26 16:22:19",
        "imageUrl": "images/lang.png",
        "defaultSetting": {
            "testId": "2",
            "addLang": "0,1,2,3,4,5,7"
        }
    },
    {
        "id": 2,
        "name": "文字網絡分析",
        "createTime": "2014-04-26 16:22:19",
        "imageUrl": "images/link.png",
        "defaultSetting": {
            "attentionWords": "馬英九,蔡英文,宋楚瑜,國民黨,民進黨,親民黨",
            "dicId": "1,2"
        }
    }],
    "createTime": "2014-04-26 16:22:19"
}
```


Figure 5. Example of parameter template of an observation case in JSON

3.9. Language Sensor and Text Sensor

We have implemented two types of sensors: *Language Sensor* and *Text Sensor* for demonstration purposes. The language sensor, as shown in Figure 6, was designed to detect the language used in a tweet. Language sensor can be used for the analysis of tweets about cross-regional discussions, such as Taiwan presidential election, umbrella revolution in Hong Kong, anti-China protest in Vietnam, etc. One can use the language sensor to automate the statistical analysis of tweets by different language communities and provides a comparative study. The detection is not trivial because several languages can be used in different parts of a tween. In the current implementation, the sensor supports the language identification of Traditional Chinese (zh-TW), Simplified Chinese (zh), Cantonese (gh), English (en), Japanese (ja), Korean (kr) and other languages. The detection of language used in a tweet can be based on different parts of a tweet such as entire body content, overall reply post, and @RT/@mention. The system allows a user to customize appropriate parts of source data in order to conduct language detection. The detection method has been implemented and shown to have an average accuracy of 0.94 for inter-coder agreement, and composite reliability of 0.98.

Figure 6. Sample visualization results of the language sensor

Text Sensor, as shown in Figure 7, is another social sensor that has been implemented in the current system. It contains several types of text mining methods, such as segmentation, word frequency statistics, and co-occurrence analysis, for analyzing the text contents in a tweet. These methods can be used to analyze the occurrence of keywords in the tweets about the changes of election situations. The sensor can also be used to do time-series analysis on various data items and display the co-occurrence relations in work network.

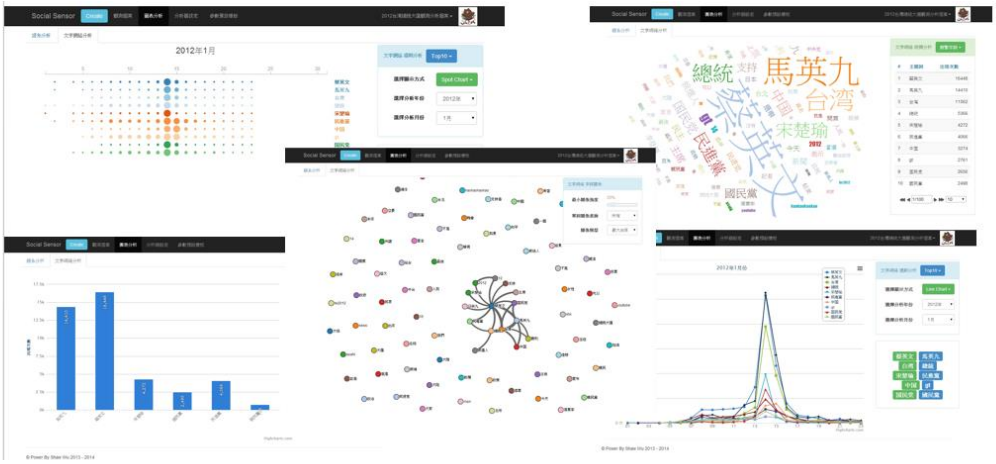

Figure 7. Sample visualization result of the text sensor

3.10. System Implementation

The social sensors in this work have been implemented as a web-based system allowing users to do the analysis through a web browser. The underlying language at the server side is Java and a Model-View-Controller (MVC) model based on SSH (Spring, Struts2, Hibernate) has been used to design the system as a component-based architecture. At the client side, we have heavily used JavaScript, JQuery and other front-end languages for interactive design, and adopted the AJAX technology to communicate between the client side and the server side. In terms of visualization design, we have used libraries such as HighCharts, d3.js, and SVG to implement various graphical displays according to the analysis needed in the social sensors.

4. EXPERIMENTAL EVALUATION

4.1. Experimental Design and Evaluation Method

We have designed experiments to evaluate the proposed system. The diary method [13] is adopted to do the evaluation and qualitative analysis. In addition to the recording of the whole operation process by the users, we have also designed unstructured open-ended questionnaire for the participants to reflect their user experience. Four users with domain knowledge have been invited to participate in the experiment. Users ID A and B are senior analysts, and users C and D are junior analysts. They are asked to use the system for three times spread in the period of about one week. The users record the steps and process they have used to analyze Taiwan presidential election. At the end of each trial run, the users are asked to fill in a survey of open-ended questions about usability. At the end of the overall experiment, another questionnaire about system usefulness is given to the users. In addition to the qualitative analysis of the recording and feedbacks by the users, we have also conducted a quantitative analysis based on the given textual feedbacks. The dual method is based on the concept of ladder of analytical abstraction from [2] [15].

4.2. Evaluation of Usefulness

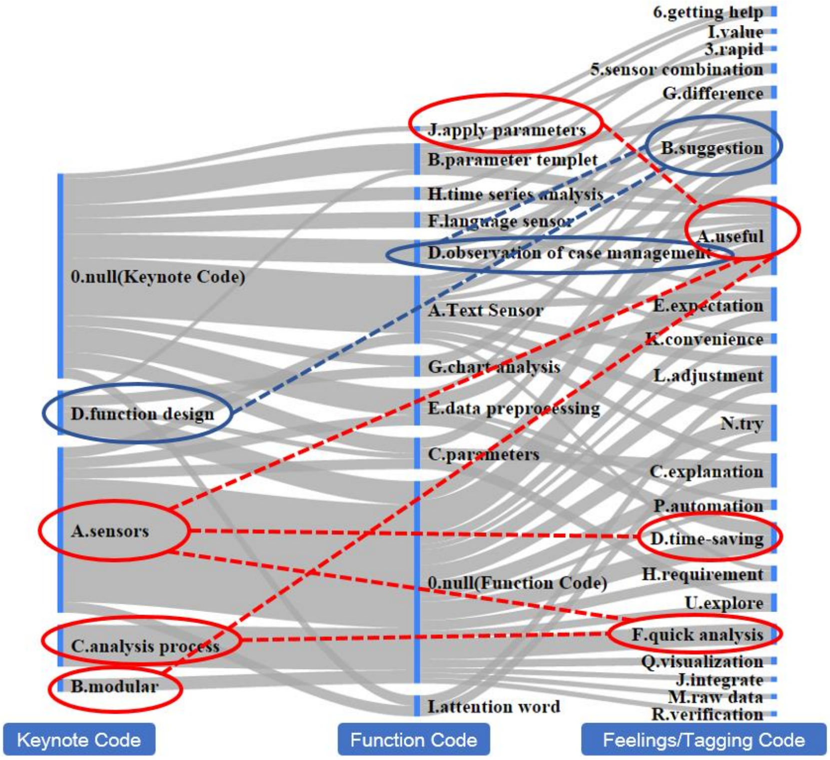

Figure 8. Coding System of System Usefulness

On the evaluation of the usefulness of the system, positive, fair, and negative comments are 60.32%, 31.75%, and 7.94%, respectively, showing that the participants consider most of the functions of the system useful. The descriptive codes of the positive comments are mostly “sensors”, “text sensor”, “parameter template”, “useful”, “time-saving”, “quick analysis”, etc. We have designed a coding scheme to further analyze the collected comments. The codes are classified into three categories: *keynote*, *function*, and *feeling* codes. By the coding path method as shown in Figure 8, each comment is coded with at least two codes out of three. For example, the coding path of “A0A” (sensors→null→useful) means that the sensors are considered useful. With this method, we can observe the large amount of feedback texts in a quantitative way. In Table 1, we can also find evidences about positive feedbacks about the design of system.

Table 1. The top tens of the coding path

Coding Path	Keynote Code Id	Keynote Code	Function Code Id	Function Code	Feelings/ Comment Code Id	Feelings/ Tagging Code	Path Count	Occurrence Frequency
A0A	A	sensors	0	null	A	useful	8	4.47%
B0A	B	modular	0	null	A	useful	5	2.79%
A0D	A	sensors	0	null	D	time-saving	5	2.79%
D0B	D	function design	0	null	B	suggestion	5	2.79%
A0F	A	sensors	0	null	F	quick analysis	5	2.79%
0DB	0	null	D	observation of case management	B	suggestion	4	2.23%
0BA	0	null	B	parameter template	A	useful	4	2.23%
A0N	A	sensors	0	null	N	try	4	2.23%
A0L	A	sensors	0	null	L	adjustment	4	2.23%
C0F	C	analysis process	0	null	F	quick analysis	3	1.68%

4.3. Evaluation of Modularity

In Table 1, we can find that the modularity of the system is well received and appreciated by the users. The frequency of coding path B0A, together with others, indicates that the sensor modularity is useful, time-saving and provides rapid analysis. Subject A mentioned that the language analysis sensor is important for analyzing tweets in mainland China, Taiwan and Hong Kong, but not necessarily applicable to other cases.” Providing a pluggable social sensor allows the users to select appropriate modules in different analysis cases. Subject B considered that modular design of the system can help users effectively learn the tools available and the controlling variables of these tools. Subject D considered the system as a time-saving tool for self-exploration and data processing. Subjects A and C mentioned their desires to see more sensors to be developed. In short, modular design of the system is considered to be useful, time-saving, and flexible for future integration of new sensors.

4.4. Evaluation of Reusability

The system attempts to use parameter template to facilitate the replication of the analysis experience from previous case studies. Basically, all four users consider the reusability through parameter template as a useful feature. The coding paths D0B and 0BA reveal that using the parameter template is indeed useful. Subject A said that “Using the default parameter values as a starter allows users to reduce the time of trial-and-error in the beginning.” Subject D mentioned that “The basis of corpus are rather common, and therefore it is a good idea to use the default parameters in a sensor to prevent the users from forgetting to adjust the parameters.” There are also some suggestions on the feature. Subject C suggests that the format of the parameter settings can be improved for better readability by the human. A better authoring interface is also more desirable.

4.5. Evaluation of Manageability

In the evaluation of manageability, all subjects agreed that the system indeed can provide effective management of case studies. The coding path of A0A, A0D, 0DB, A0N, A0L, and C0F are all examples of comments with positive feedback on case management. Nevertheless, there are also suggestions for improvement. For example, subject A suggest that the system should provide a data export function to allow further analysis in other advanced software packages. Subject C thought that saving analytical parameters for a case is more desirable than saving the analysis results. Subject C suggests to implement the function of personal analysis history in the future. In short, the manageability feature of the system can significantly reduce the time of preparing and managing a data analysis case. It is suggested to increase the interconnection interfaces with other analysis tools and to improve the management functions of case studies.

5. CONCLUSIONS AND FUTURE WORK

In this research, we have proposed the concept of social sensor for data analytics of social media and implemented a system to realize this concept with the design goals of *manageability*, *modularity*, *reusability*. In this system, a user can quickly perform data analysis and manage a case study by specifying data sources, adjusting case parameters, and visualize the result immediately. One can perform a case study by choosing appropriate social sensors that are extensible as plug-in. The experience of appropriate parameter settings for a case can also be saved into a parameter template for reuse in other similar case studies. In addition, in this system we feature a modularized design allowing different analytical modules to be reusable in different sensors and cases. We have also implemented the language sensor and the text sensor as examples. In the language sensor, we have proposed a method to distinguish the main language used in a tweet. The method is shown to have average accuracy of 0.94 for Intercoder Agreement (IA), and 0.98 for Composite Reliability (CR).

We have also conducted a system evaluation with qualitative analysis through the diary method. The evaluation result on usefulness, modularity, reusability, and manageability are all very positive. The results also show that this tool can greatly reduce the time needed to perform data analysis and solve the problems encountered in traditional analysis process. The experimental results reveal that the concept of social sensor and the proposed system design are shown to be useful in the data analysis process for social media.

In the future, we hope to improve the interface design and operational flow of the system. We would also like to improve the design of existing social sensors and increase the interactivity among the sensors. We also hope to handle more types of social media and accumulate more case experiences. We would also like to perform more case studies of major events such as umbrella revolution in Hong Kong, anti-China protest in Vietnam, related events such as China, Japan and South Korea, etc.

6. REFERENCES

[1] B. Lorica, Big Data Now: 2013 Edition Current Perspectives. *O'Reilly Media*, 2013.

[2] T. F. Carney, Collaborative inquiry methodology. *Windsor, Ontario, Canada: University of Windsor, Division for Instructional Development*, 1990.

[3] E.J. Yuan, M Feng, & J.A. Danowski, “Privacy” in Semantic Networks on Chinese Social Media: The Case of Sina Weibo. *Journal of Communication*,

2013.
[4] J. Burgess and A. Bruns, （Not） the Twitter Election: The Dynamics of the #ausvotes Conversation in Relation to the Australian Media Ecology. *Journalism Practice 20 Mar. 2012,* 2012.
[5] A. M. Kaplan, & M. Haenlein, Users of the world, unite! The Challenges and opportunities of social media, 2010.
[6] H. Kwak, C. Lee, & S. Moon. What is twitter, a social network or a news media? *International World Wide Web Conference Committee, North Carolina, USA*, 2010.
[7] OECD., Participative web and user-created content: Web 2.0, wikis, and social networking. *Paris: Organisation for Economic Co-operation and Development*, 2007.
[8] M. Stelzner, Social Media vs. Social Networking: What's the difference? *http://www.examiner.com/article/social-media-vs-social-networking-what-s-the-difference*, 2009.
[9] T. Sakaki, M. Okazaki, Y. Matsuo, Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. *WWW2010, April 26-30, 2010, Raleigh, North Carolina*, 2010.
[10] T. Sakaki, Y. Matsuo, T. Yanagihara, N. P. Chandrasiri, K. Nawa, Real-time Event Extraction for Driving Information. *Proceedings of the 2012 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, May 27-31, 2012, Bangkok, Thailand*, 2012.
[11] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, Microblogging during two natural hazards events: what twitter may contribute to situational awareness. *In Proc. CHI 2010, ACM Press, 1079-1088*, 2010.
[12] V. Vasudevan, J. Wickramasuriya, S. Zhao, L. Zhong, Is Twitter a Good Enough Social Sensor for Sports TV? *Pervasive Collaboration and Social Networking 2013 IEEE*, 2013.
[13] S.-M. Wang, ICTs and Everyday Rhythms: Looking into the Methods. *Journal of Communication Research and Practice, 23-43*, 2014.
[14] S.-F. Shih, Information Transmission Analysis and Automated Classification Design for New Media in a Disaster Event – Case Study of Typhoon Morakot. Master Thesis. *Department of Computer Science, National Chengchi University, Taiwan*, 2013.
[15] F.F. Chang, The Five Steps of Qualitative Data Analysis: Climbing up a Ladder of Abstraction. *Journal of Elementary Education, No. 35, 87-120*, 2010.
[16] Y.-C. Cheng, P.-L. Chen, Emerging Communities in Social Media During the 2012 Taiwanese Presidential Elections: A Big-Data Analysis Approach. *Mass Communication Research, No. 120, 121-165*, 2014.