Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2018 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

6-26-2018

Meta-Information as a Service: A Big Social Data Analysis Framework

Kashif Ali
RMIT University, kashif.ali@rmit.edu.au

Margaret Hamilton
RMIT University, margaret.hamilton@rmit.edu.au

Charles Thevathayan
RMIT University, Charles.Thevathayan@rmit.edu.au

Xiuzhen Zhang
RMIT University, xiuzhen.zhang@rmit.edu.au

Meta-Information as a Service: A Big Social Data Analysis Framework

Completed Research Paper

Kashif Ali
School of Science, RMIT University
Melbourne, Victoria, Australia
Kashif.ali@rmit.edu.au

Margaret Hamilton
School of Science, RMIT University
Melbourne, Victoria, Australia
Margaret.hamilton@rmit.edu.au

Charles Thevathayan
School of Science, RMIT University
Melbourne, Victoria, Australia
Charles.thevathayan@rmit.edu.au

Xiuzhen Zhang
School of Science, RMIT University
Melbourne, Victoria, Australia
Xiuzhen.zhang@rmit.edu.au

Abstract

Social information services generate a large amount of data. Traditional social information service analysis techniques first require the large data to be stored, and afterwards processed and analyzed. However, as the size of the data grows the storing and processing cost increases. In this paper, we propose a 'Meta-Information as a Service' (MIaaS) framework that extracts the data from various social information services and transforms into useful information. The framework provides a new formal model to present the services required for social information service data analysis. An efficient data model to store and access the information. We also propose a new Quality of Service (QoS) model to capture the dynamic features of social information services. We use social information service based sentiment analysis as a motivating scenario. Experiments are conducted on real dataset. The preliminary results prove the feasibility of the proposed approach.

Keywords: Social Information Services, Big Social Data, Sentiment Analysis, Meta-Information, Service Composition

Introduction

The term big data is used to define the large amounts of data that are beyond the ability of traditional techniques and processes to collect, manage and process within a reasonable time-frame (Becker, King, and McMullen 2015). Big data is often defined on the basis of three characteristics (Zaslavsky, Perera, and Georgakopoulos 2013) or 3Vs: *Volume*, i.e., size of the data, *Variety*, data types e.g., sensor data, social data, etc., *Velocity*, i.e., speed of the data generation. Social information services, e.g., Facebook, Twitter, etc., are part of big data notion and retain an unique status. Social information services provide a free platform to its users for the data generation and sharing. As a result, these services generate a large amount of data defined as '*Big Social Data*'.

Big social data contains rich information such as public opinion and sentiment of social media users (Ali et al. 2017). Researchers have successfully used big social data in various domains such as business, marketing, politics, health surveillance and crisis management (K. Ravi and V. Ravi 2015). Despite the initial success of big social data, one major challenge is to efficiently collect, process and analyze the staggering amounts of data (Cuzzocrea 2017). Distributed technologies such as Hadoop, etc., have been used for the efficient and parallel processing. However, as the data size grows, the cost of infrastructure, i.e., storing and processing, also increases (Katal, Wazid, and Goudar 2013) (Labrinidis and Jagadish 2012). For instance, many applications such as customers' patterns detection, public political sentiment, etc., require to collect and store big social data for longer time periods, and later analyze for the information extraction. Consequently, these applications need larger storing and processing infrastructure. In addition, there are several online services for social information service analysis, however they only provide analysis for limited data size (Wan and Paris 2014).

The *Meta-Data* provides the information about the data (Singh et al. 2003). It contains the information about the creation, transformation, meaning, types and quality of the data. The meta-data provides the ability to effectively manipulate the large datasets. In contrast, the notion of *Meta-Information* is used to define the actual information derived from actual data processing and analysis (Michener 2006). In a typical meta-information application system, the required information is extracted from the large datasets and stored in meta-information database (B. Liu and Xiang 2016). Later, the information is composed and delivered to end users on-demand. As a results, the meta-information systems provide two main benefits: 1) Eliminate the need to store large amounts of the data. 2) Provide on-demand request and composition of the required information.

In this paper, we propose a 'Meta-Information as a Service' (MIaaS) framework that collects and transforms big social data into useful information, i.e., Meta-Information, and delivers as a service. Unlike existing social information service analysis techniques where the data is stored and later processed; our framework periodically captures the large streams of big social data, and instantly extracts and stores the information. Afterwards, the information is composed and delivered to end users as a service upon their request. Hence, the proposed framework eliminates the need to first store big social data and subsequently process it. The novelty of our approach is being first to utilize the concept of meta-information for big social data analysis.

In this paper, we use social information service based sentiment analysis and opinion mining as an application scenario. However, our framework can be applied to other domains applicable to big social data analysis such as recommender system, prediction systems, etc. The main contributions of our work are as follows:

- A service oriented framework that efficiently transforms big social data into meta-information.
- A new formal model to present the component and composite services required for analysis. A data model to store and access the meta-information.
- A novel Quality of Service (QoS) model based on the dynamic features of big social data.

The rest of the paper is organized as follows. Section 2 describes the motivating scenario. Section 3 provides the related work. Section 4 provides the framework overview and Section 5 elaborates the details of the framework. Section 6 provides the details of experimental results and evaluation by real datasets. Section 7 concludes the paper and provides the direction for future work.

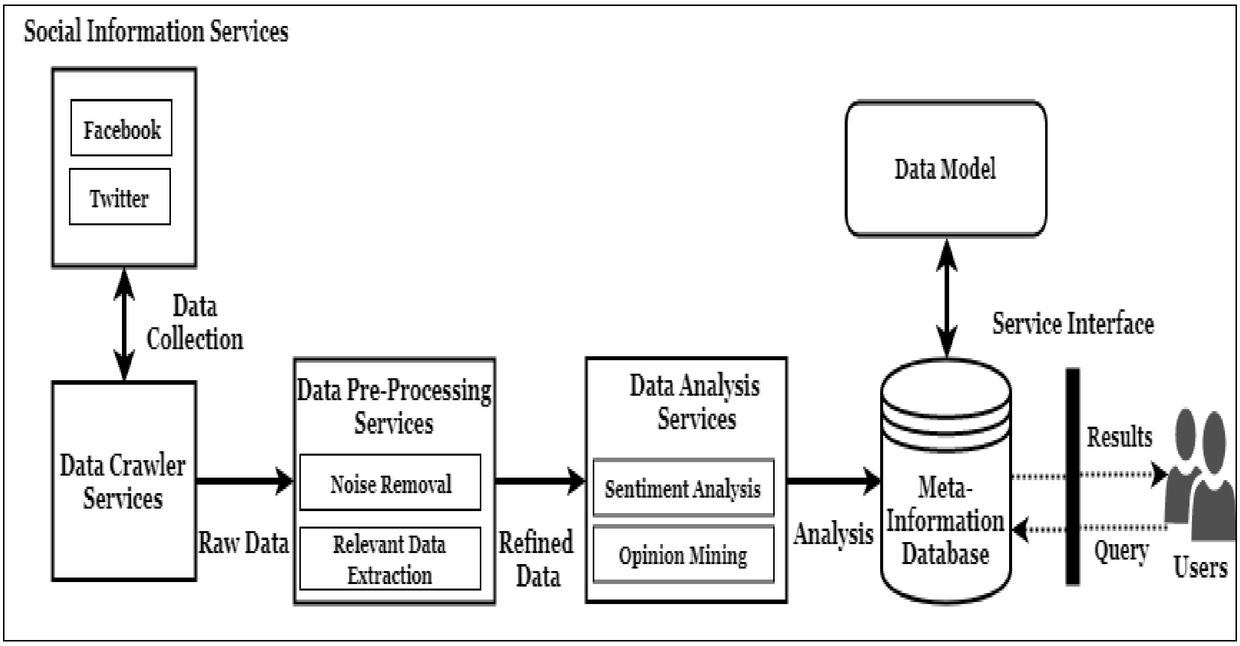


Figure 1. Political Sentiment Analysis Scenario

Motivating Scenario

We use a political sentiment analysis as our motivating scenario. Let us assume that Sara is a social information service analyst working for a news agency 'News First'. News First is interested in monitoring the political sentiment of general public in the all states of USA before the upcoming presidential election. Sara is assigned the task of developing a system to analyze the political sentiment from various social information services. Due to budgetary limitations, Sara is required to develop the system that utilizes limited computing resources, i.e., storage and processing. Moreover, the system needs to deliver the requested information to end users on-demand in different formats over the Internet.

Sara develops the required system by using a service oriented approach (see Figure 1). She develops the required system by using various online services. Firstly, the system periodically collects the raw data from social information services by using data crawlers services. Secondly, each raw data batch is instantly processed with data preprocessing services (e.g., data cleansing). Afterwards, the refined data is analyzed by various data analysis services to extract multiple types of information such as user location, sentiment, etc. The analysis results are then stored in a meta-information database based on an extensible data model. The data model provides a structure to store, retrieve and compose the extracted information. Finally, end users can request the required information through service interface over the Internet. The information is on-demand composed from the meta-information database, and delivered in various formats, e.g., maps, graphs, etc.

In this scenario, Sara composes various available services to collect, process and analyze big social data, and stores the analysis results for later usage. Thus, the proposed approach first eliminates the need for large storage space to retain the data. Secondly, it saves the local computing resources by using the online services, and enables the on-demand delivery of information to end users.

Related Work

Social Information Services

Social information services have become an integral part of big data. Big social data generated by social information services have been used in various applications. For instance, in (Y. Liu, Han, and Tian 2013), extracted the requirements for developing new products from social media users' comments. Big social data provides the monitoring and surveillance capabilities for the epidemic monitoring. In (Padmanabhan et al. 2013), researchers used the blog posts to determine the trends of epidemic diseases among the users. The data obtained from social information services is also used for predicting the election results (Tumasjan et al. 2010). Social information services play an important part in emergency and crisis management. In (Gomide, Veloso, and Almeida 2011), a dengue surveillance model is proposed to identify the locations of flu victims by using the spatio-temporal data of Twitter. However, current techniques use typical data oriented approach. These approaches require to collect large datasets, and later manual labeling of test dataset and training algorithm for the information extraction. Moreover, current systems lack the capability of on-demand information analysis.

Sentiment Analysis

A significant portion of big social data is in unstructured textual format. One mechanism to extract the useful information from textual data is through Natural Language Processing (NLP) (Baldwin et al. 2013). Sentiment analysis and opinion mining are two subfields of NLP used for big social data analysis (Cambria et al. 2013). Sentiment analysis is used to automatically label text into three categories: positive, negative and neutral (Madhoushi, Hamdan, and Zainudin 2015). Whereas, opinion mining is considered as process to extract opinion, attitudes and emotions from the text (Med-hat, Hassan, and Korashy. 2014). Sentiment analysis or opinion mining process is comprised of three main steps: data collection, data preprocessing and cleansing, information extraction or text analysis (Guellil and Boukhalfa 2015). Social information service based sentiment analysis has been successfully utilized in various application such as financial market prediction, box office revenue prediction, marketing and business analytics, recommender system, and election results forecast (K. Ravi and V. Ravi 2015).

Service Oriented Paradigm

The service oriented paradigm utilizes services as building blocks to build new applications (Papazoglou 2003). Services provide a powerful abstraction that hides the underlying implementation information from end users (Neiat et al. 2014). As a result, end users are only concerned with the inputs and outputs of services and how to use them. The service oriented paradigm also provides the benefits of dynamic selection and composition of services to develop complex applications based on Quality of Service (e.g., price, response time) (Medjahed, Bouguettaya, and Elmagarmid 2003). Also, services are also used as a preferred mechanism for data delivery to various applications over the network (Neiat et al. 2015).

Big Social Data and Services

The big data service computing is an emerging research domain. Research work (Vu and Asal 2015) proposed a high level conceptual model for *Big Data as a Service* architecture. Similarly, (Demirkan and Delen 2013) provides an overview of three service computing models: *Data-as-a-Service*, *Information-as-a-Service*, *Analytics-as-a-Service*. However, these works only provide conceptual and descriptive models of big data as a service, and do not particularly focus on big social data. There are several online services available for social information service monitoring and analysis. For example, SentimentViz[1], SocialMention[2] and Brand24[3] provide the analysis of social data. However, these

[1] https://www.csc.ncsu.edu/faculty/healey/tweet viz/tweet app/

[2] http://www.socialmention.com/

[3] https://brand24.net/

services gather a limited amount of data and provide basic data insights, e.g., user clicks, likes, trending words, etc. Moreover, many of online social information services are used for customer engagement and content publishing. In addition, these tools lack the flexibility to retain and analyze historical data.

Meta-Information Systems

The notion of meta-information system has been utilized in various data extensive applications. In (Alyami et al. 2017), researchers proposed a meta-data based health record system to organize and retrieve patients health records information in medical emergencies. Researchers have also explored the meta-data analysis of Wi-Fi traffic to estimate the demographics of users (Li, Zhu, and Ma 2017). In (Hoiles, Aprem, and Krishnamurthy 2017), researchers investigated the Youtube users engagement with video channels through various types of platform provided meta-information attributes. To the best of our knowledge, current meta-information systems mostly focus on meta-data properties for information retrieval and/or to incentivize the existing information. Moreover, there is limited work available for utilizing meta-information in big social data analysis.

In contrast, our framework specifically focuses on big social data. Instead of storing the big social data for later processing and analysis, it instantly gathers and transforms the big social data into meta-information. Afterwards, the information is delivered as a service to end users upon request.

The Solution Overview: Meta-Information as a Service

The proposed MIaaS framework is divided into two categories: 1) Service Model 2) Data Model. Service model describes the component and composite services required for big social data analysis. Also, it provides a novel QoS model for big social data analysis. The data model provides a baseline structure to store the analysis results as meta-information and on-demand composition of meta-information.

In service model, there are three types of component services: Data crawler, Data preprocessing, Data analysis. For each component service, we assume that there are several available candidate services provided by different service providers. The data crawler service is comprised of multiple data gathering APIs and services to collect the data from social information services. The data preprocessing services remove the noisy and irrelevant data from the collected data. The data analysis services extract various types of information from the data. The services communicate via data messages, e.g., XML, JSON, etc. The data model provides an eXtensible Markup Language (XML) based database architecture. It uses XML based documents to store and access the analysis results, i.e., meta-information.

The proposed framework utilizes the service oriented paradigm to compose available services. The service orientation provides two main benefits for big social data analysis: separation of concerns and abstraction. The separation of concerns provides the benefit of dividing a system into a set of loosely coupled layers. Each layer has a set of services provided by various service providers. Meanwhile, abstraction hides the implementation complexity from end users. As a result, end users are not concerned with the data collection, processing and storage processes.

Service Model

In this section, we first present the formal service model and propose a new QoS model. Second, we provide the details of our data model.

Meta-Information Service Model

We discuss key elements of our meta-information service model. We formally define the MIaaS as a composite service. Later, we formalize each component service. The service model is formally defined as follows:

Definition 1: Meta-Information as a Service (*MS*) is defined as a tuple of <*ID, SISi, T, Fi, Qi*>, where

- *ID* is the service id.
- *SISi* presents the set of social information services < *sis1, sis2, …., sis n* > from which the analysis (i.e., meta-information) is required.
- *T* is the topic of interest (e.g., entity) for analysis.
- *Fi* presents a set of functions (e.g., sentiment classification) offered by the service.
- *Qi* is a set of QoS properties < *q1, q2 ,…, qn* > offered by the service.

Definition 2: The data crawler service (*DS*) gathers the data from various social information services. *DS* is a tuple of <*D-id, sis, ki, int, qi*>, where

- *D-id* is the data crawler service id.
- *sis* is target social information service for the data collection.
- *ki* is the set of keywords or search terms to retrieve the data.
- *int* defines the time interval for the service to collect the data.
- *qi* is a set of QoS properties (e.g., throughput).

Definition 3: The data preprocessing service (*PS*) removes noise from the data. *PS* is a tuple of < *P-id, dt, fl, qi* >, where

- *P-id* is the data preprocessing service id.
- *dt* is data type (e.g., tweets, comments) to be processed.
- *fl* describes the appropriate types of noise removal filters.
- *qi* defines the set of non-functional properties (e.g., accuracy).

Definition 4: The data analysis service (*AS*) extracts information from the collected data. *AS* is a tuple of < *A-id, fn, qi* >, where

- *A-id* is the data analysis service id.
- *fn* defines the available functions (e.g., sentiment polarity) for analysis.
- *qi* determines the set of QoS (e.g., precision, accuracy).

Meta-Information Service Quality Model

Social information services provide various features (e.g., size, geo-tagging) for the data sharing (Ali et al. 2017). One challenge for end users is to compose information based on appropriate features. We propose following novel QoS model based on the dynamic features of big social data for meta-information service composition.

- *Volume*: The volume *V* property determines the amount of the data (e.g., tweets, comments) from each social information service is analyzed. The volume is calculated as following function, where *N* is total number of data items per social information service *SIS*.

$$V = \sum_{i=0}^{n} SISi\ (N) \tag{1}$$

- *Freshness*: The freshness *F* property determines that the information *I* is recent. The freshness provides the temporal bounds for the meta-information composition. The freshness is calculated as follows, where *tl* is the lower time bound and *tu* is the upper time bound.

$$F = \int_{tl}^{tu} I \tag{2}$$

- *Coverage*: The coverage *C* property provides the ability to visualize the information *I* based on spatial (e.g., states) requirements. The coverage composes the meta-information based on geo-locations. The coverage property is determined as follows, where $GLi = \{gl1, gl2, \ldots, gln\}$ is the set of geo-locations.

$$C = I - (I \cap GLi) \tag{3}$$

```xml
<?xml version="1.0"?>
<Social_Information_Service Name="Twitter">
  <Document Dataset_ID="D1" Temp="Date1" Volume="500"/>
  <Document Dataset_ID="D2" Temp="Date2" Volume="450"/>
  <Document Dataset_ID="D2" Temp="Date2" Volume="450"/>
</Social_Information_Service>
```

(a)

```xml
<?xml version="1.0"?>
<Document ID="D1">
  <Subjective_Information type="Sentiment" Value="Positive">
    <Spatial Location="Loc1"/>
  </Subjective_Information>
  <Subjective_Information type="Sentiment" Value="Negative">
    <Spatial Location="Loc2"/>
  </Subjective_Information>
</Document>
```

(b)

Figure 2. (a) Master Data Document Structure (b) Summary Document Structure

Data Model

The data model is comprised of two components: Design Model, Information Composition Model. The design model provides the structure to store the meta-information. The information composition model outlines the process of composing the meta-information.

Design Model

The design model consists of two XML based components for data storing: Master Data Documents, Summary Documents.

Master Data Document contains the meta-data of each dataset processed and analyzed. For each social information service, there is only one master data document. The master data documents are iteratively updated after each dataset of a respective service is analyzed. Figure 2 (a) shows the details of the master data document. The properties of master data documents are defined as follows:

- *Dataset ID*: is the unique id assigned to each dataset analyzed.
- *Temp*: presents the temporal information (e.g., time-date) of the dataset.
- *Volume*: the number of items analyzed in the dataset.

Summary Document contains the simplified analysis results, i.e., meta-information, extracted from each dataset. Consequently, there are multiple summary documents for each master data document. Figure 2 (b) presents the details of the summary document structure. The properties of summary documents are defined as below:

- *Document ID*: is the unique name of the summary document which is inherited from master data document, i.e., Dataset ID.

- *Subjective Information Type*: defines a set of subjective information (e.g., sentiment, emotions, trending words) extracted from the dataset.
- *Value*: contains the values, i.e., results, of the subjective information (e.g., positive sentiment, negative sentiment).
- *Spatial-Info*: provides the names of geo-locations for each subjective information value.

Information Composition Model

The information composition model integrates the meta-information from different summary documents of various social information services. The composition is a two step process. First, the relevant summary documents are sorted by using master data document. Second, the information from summary documents is integrated.

For summary documents sorting, the freshness *F* and volume *V* requirements are used to search master data documents. The retrieve function $F_{Retrieve}$ successfully returns *k* number of *D* summary documents, if the documents within the lower *tl* and upper *tu* time bounds have minimum required number of analyzed data items *I*. The retrieve function is defined as follows:

$$F_{Retrieve}(F, V) = If \begin{cases} \int_{tl}^{tu} \sum_{i=0}^{k} (D|I) \geq V,\ D_k \\ Otherwise,\ \emptyset \end{cases} \tag{4}$$

In second step, the meta-information from the retrieved summary documents is integrated based on coverage *C*, i.e., geo-location requirements, by using integration function $F_{Integrate}$. It is possible that a summary document may not contain the meta-information based on the required geo-locations. In order to integrate the appropriate meta-information and comply with volume *V* requirement, the composer validates the geo-locations *GLi* from the retrieved documents with coverage parameters as follows:

$$F_{Integrate}(D_k, C, V) = \sum[(D_k|GLi) \cap C] \geq V \tag{5}$$

Finally, if the end user's requirements are validated, the meta-information from the collected summary documents is integrated as a data message and delivered into various formats.

Experiments and Evaluation

To evaluate our proposed approach, we implemented a prototype by using our motivating scenario. We collected the approval ratings of USA president `Donald Trump' through sentiment analysis. We performed the experiments in three fold steps: 1) We collected the data from two social information services: Facebook and Twitter. 2) We evaluate the performance of data preprocessing component. 3) We analyze the collected data and extract the sentiment with respect to geo- locations and provide the detailed results. The prototype system is developed in *Visual Studio (2015)* by using *ASP.Net/C#*. The prototype is implemented on a 3.40 *GHz* Core *i7* processor and 8 GB RAM on Windows 7.

Data Collection

We collected the data from Facebook and Twitter for 7 days of interval between 19-March-2017 to 26-March-2017. The data crawler service is set to collect the data after 24 hours period. Social information services provide various mechanisms to collect the data. As a result, data crawler service is consisted of different data collection services. For Facebook, data crawler service used Facepager[4] to collect the data from Donald Trump's official Facebook page. Facepager uses 'Graph API' to collect Facebook posts and comments from public Facebook pages. For Twitter, data crawler service used 'Twitter Streaming API' to get the random tweets for given set of keyword(s). Table 1 provides the details of the data collected by data crawler service.

Table 1. Data Collection Results

	Facebook (Posts)	Twitter (Tweets)
Day 1	5606	12165
Day 2	11367	20025
Day 3	10840	20117
Day 4	10023	20098
Day 5	29242	20124
Day 6	26644	20361
Day 7	15034	20153
Total:	**108756**	**133043**

Data Preprocessing

The data preprocessing is used to remove irrelevant data (e.g., noise). For the data preprocessing, we used following three filters to exclude the irrelevant data:

- *Language Filter*: filters out the data based on a human language. We removed all the non-English data.
- *Search Term Filter*: includes the data based on the context matching. The collected data may contains the data items which are not related to the context. We used three keywords: 'US President', 'Trump', 'Donald Trump' retain a data item.
- *Stop Words Exclusion Filter*: excludes the a data item that is irrelevant to the context. For example, a tweet contains the term 'Jared Trump' holds the keyword 'Trump'. However, it is not relevant to the context. Therefore, 'Jared' is used as a stop word to discard such data items.

After the noise filtering, we filtered out the data without geo-location information. The Twitter service allows to geo-tag a tweet. Thus, all the tweets without geo-tags are discarded. In contrast, Facebook does not allow to gather the geo- locations of the data directly. As a result, the geo-locations are identified via data, i.e., text, parsing. We used Named Entity Recognizer (NER)[5] to extract geo-location names (e.g., cities) from data items. To evaluate the performance of our filters, we used *Signal-To-Noise Ratio (SNR)* as our evaluation measure.

$$SNR = \frac{Signal}{Noise{=}(Total-Signal)} \quad (6)$$

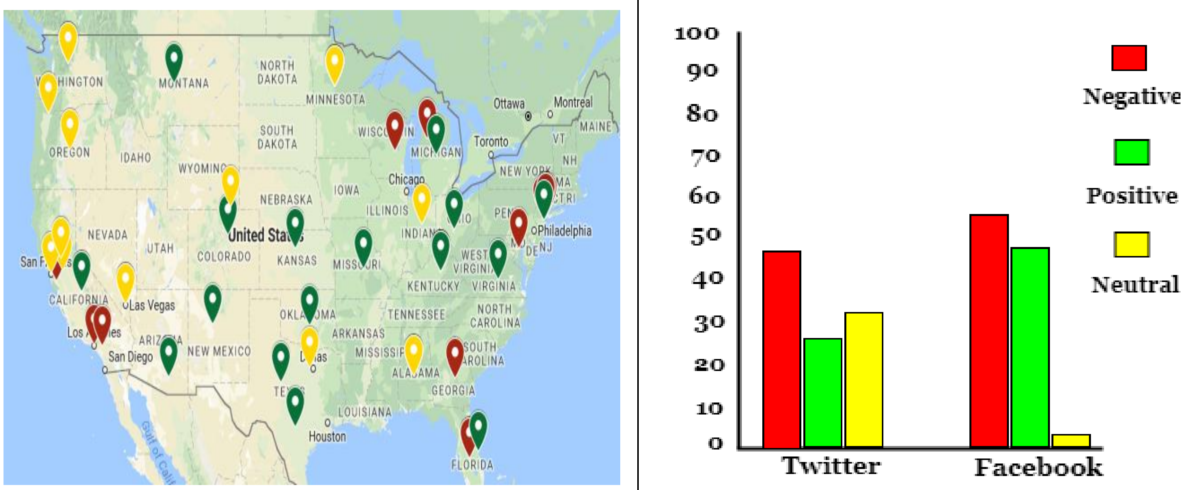
[4] https://github.com/strohne/Facepager

[5] http://nlp.stanford.edu/software/CRF-NER.shtml

Where *Signal* is the number of relevant data items after each filtering step, and *Total* is the number of data items remaining. Table 2 provides the results of the data preprocessing.

Table 2. Data Preprocessing SNR Results

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Facebook Noise-Filter	1.96	4.19	3.1	3.63	4.22	3.66	3.37
Facebook Geo-Filter	0.39	1.12	0.45	0.49	0.25	0.3	0.35
Twitter Noise-Filter	0.52	0.76	0.65	0.69	0.71	1.06	0.93
Twitter Geo-Filter	0.009	0.012	0.013	0.01	0.014	0.01	0.008



(a) (b)

Figure 3. (a) Sentiment By Location (b) Composite Sentiment Analysis

Data Analysis and Composition Results

We analyzed the preprocessed datasets by two different sentiment analysis services: Sentistrength[6] and Alchemy API[7]. Sentistrength is specifically designed to analyze the twitter data. In contrast, Alchemy API is used to analyze text from longer text (e.g., blog reviews, news articles). The prototype offers two formats to view composition results: 1) Sentiment percentage bar chart. 2) Sentiment by location map. For sentiment visualization, we used USA as a spatial parameter. Figure 3 (a) shows the composite sentiment analysis results on Google Maps in time-line by random date parameters. Figure 3 (b) shows the sentiment analysis results in a percentage bar chart. It is observed that USA president has highest number of negative sentiment and lowest positive sentiment on Twitter. In comparison, Facebook shows almost equal amount of negative and positive sentiment. In comparison, a few number of users have neutral sentiment on Facebook, while a large number of users have neutral sentiment.

[6] http://sentistrength.wlv.ac.uk/

[7] https://alchemy-language-demo.mybluemix.net/

Conclusion

We proposed a service oriented framework ‘Meta-Information as a Service’ (MIaaS) that extracts the big social data from social information services, analyses and transforms into information, i.e., meta-information. Current analysis approaches first require to store the data, and later process and extract the required information. In comparison, MIaaS periodically collects the data and instantly analyzes and stores the extracted information. In addition, it enables the end users to on-demand compose the required information. The framework provides a service formal model to present the services required for big data social analysis. We proposed a data model that provides an XML based document structure to store and integrate the required meta-information. We also devised a new QoS model to analyze features of big social data for meta-information composition. We implemented a prototype and conducted experiments on the real-world data. We used sentiment analysis as an application scenario to collect the approval ratings of USA president. The preliminary results show the applicability of our proposed approach.

Limitations and Future Work: Currently, we have conducted our experiments with limited data from two social information services. In future, we plan to integrate a larger pool of social information services. We also aim to incorporate new features in the proposed QoS model. Finally, we are interested to compare the efficiency of our framework with existing big social data analysis approaches.

Acknowledgements

This research was supported under the Australian Research Council's Linkage Projects funding scheme (project number LP120200305) for the project titled "Integrated Smart Airport Services". The findings achieved herein are solely the responsibility of the authors.

References

Ali, K. et al. (2017). “Sentiment Analysis as a Service: A social media based sentiment analysis framework”. In: IEEE International Conference on Web Services, pp. 660-667.

Alyami, M. et al. (2017). “Managing personal health records using meta-data and cloud storage”. In: IEEE 16th International Conference on Computer and Information Science, pp. 265-271.

Baldwin, T. et al. (2013). How noisy social media text, how diffrnt social media sources?" In: International Joint Conference on Natural Language Processing, pp. 356-364.

Becker, D., T. King, and B. McMullen (2015). “Big data, big data quality problem”. In: IEEE International Conference on Big Data (Big Data), pp. 2644-2653.

Cambria, E. et al. (2013). “New avenues in opinion mining and sentiment analysis”. In: IEEE Intelligent Systems 28(2), pp. 15-21.

Cuzzocrea, A. (2017). “Multidimensional mining of big social data for supporting advanced big data analytics”. In: 40th IEEE International Convention on Information and Communication Technology, Electronics and Microelectronics, pp. 1337-1342.

Demirkan, H. and D. Delen (2013). “Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud”. In: Decision Support Systems 55(1), pp. 412-421.

Gomide, J., A. Veloso, and V. Almeida (2011). “Dengue surveillance based on a computational model of spatio-temporal locality of Twitter”. In: 3rd international web science conference. ACM.

Guellil, I. and K. Boukhalfa (2015). “Social big data mining: a survey focused on opinion mining and sentiments analysis”. In: 12th International Symposium on Programming and Systems, pp. 1-10.

Hoiles, W., A. Aprem, and V. Krishnamurthy (2017). “Engagement and Popu- larity Dynamics of YouTube Videos and Sensitivity to Meta-Data”. In: IEEE Transactions on Knowledge and Data Engineering 29(7), pp. 1426-1437.

Katal, A., M. Wazid, and R. Goudar (2013). “Big data: issues, challenges, tools and good practices”. In: Sixth IEEE International Conference on Contemporary Computing, pp. 404-409.

Labrinidis, A. and H. Jagadish (2012). “Challenges and opportunities with big data”. In: Proceedings of the VLDB Endowment 5, pp. 2032-2033.

Li, H., H. Zhu, and D. Ma (2017). "Demographic Information Inference through Meta-data Analysis of Wi-Fi Traffic". In: IEEE Transactions on Mobile Computing.

Liu, B. and J. Xiang (2016). "Extraction and management of meta information on the domain-oriented Deep Web". In: 7th IEEE International Conference on Software Engineering and Service Science, pp. 787-790.

Liu, Y., W. Han, and Y. Tian (2013). "Trending topic prediction on social network". In: 5th International Conference on Broadband Network and Multi-media Technology, pp. 149-154.

Madhoushi, Z., A. Hamdan, and S. Zainudin (2015). "Sentiment Analysis Techniques in Recent Works". In: Science and Information Conference (SAI).

Medhat, W., A. Hassan, and H. Korashy. (2014). "Sentiment analysis algorithms and applications: A survey". In: Ain Shams Engineering Journal 5, pp. 1093-1113.

Medjahed, B., A. Bouguettaya, and A. Elmagarmid (2003). "Composing web services on the semantic web". In: The VLDB Journal 12, pp. 333-351.

Michener, W. (2006). "Meta-information concepts for ecological data management". In: Ecological informatics 1, pp. 3-7.

Neiat, A. et al. (2014). "Failure-proof spatio-temporal composition of sensor cloud services". In: International Conference on Service-Oriented Computing, pp. 368-377.

Neiat, A. et al. (2015). "Spatio-Temporal Composition of Crowdsourced Services". In: Inter- national Conference on Service-Oriented Computing, pp. 373-382.

Padmanabhan, A. et al. (2013). "FluMapper: an interactive CyberGIS environment for massive location-based social media data analysis". In: Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery. ACM.

Papazoglou, P. (2003). "Service-oriented computing: Concepts, characteristics and directions". In: IEEE Proceedings of the Fourth International Conference on Web Information Systems, pp. 3-12.

Ravi, K. and V. Ravi (2015). "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications". In: Knowledge-Based Systems 89, pp. 14-46.

Singh, G. et al. (2003). "A metadata catalog service for data intensive applications". In: ACM/IEEE Conference on Supercomputing, pp. 33-33.

Tumasjan, A. et al. (2010). "Predicting elections with twitter: What 140 characters reveal about political sentiment". In: International Conference on Web and Social Media. AAAI, pp. 178-185.

Vu, H. and R. Asal (2015). "A framework for big data as a service". In: International Conference on Digital Signal Processing, pp. 492-496.

Wan, S. and C. Paris (2014). "Improving government services with social media feedback". In: 19th international conference on Intelligent User Interfaces, pp. 27-36.

Zaslavsky, A., C. Perera, and D. Georgakopoulos (2013). "Sensing as a service and big data". In: arXiv preprint arXiv, pp. 1301-0159.