Real Time Hasty Alert System in Online Social Network Over Distributed Processing Framework in Big Data

Mrs. A. Hepsiba[1]

[1]Associate Professor, Department of Computer Applications, Karpaga Vinayaga College of Engineering and Technology
[2]Research Scholar, Department of Computer Science, Mother Teresa Womens University.

ABSTRACT

In Online Social Network Event dispensation is a method of tracking and analyzing streams of information (data) about things that happen (events), and deriving a conclusion from Online Social Network in real time. Instagram, Twitter are a social network platform that consists of billions of users all over the world where people collaborate and Share information related to real world events. An important characteristic of Online Social Network is its real-time nature and also investigate the real-time interaction of events such as cyclones in Online Social Network and propose a framework to monitor messages to detect a target event. These large scales message data processing are done by placing those message events in a distributed system. The server processes the message queue and executes the operations based on it. An devise classifier of messages based on features such as the keywords in a message, the number of words, and their context. The status update which almost specifies what is happening around an individual and also around the individual's location. This small content with real world information when processed with some statistical tool may help us to predict a Real world event and regard each Online Social Network user as a sensor and apply particle filtering, which are widely used for location estimation. Message in the message queue is done by Apache Kafka which is a distributed publish-subscribe messaging system. These frameworks will parallelize our computations over a cluster of machines.

INTRODUCTION

Data mining has attracted a great deal of attention in the information industry and in society as a whole in recent years, due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. The information and knowledge gained can be used for applications ranging from market Analysis, fraud detection, and customer retention, to production control

and science exploration. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. A side from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post

processing of discovered structures, visualization, and online updating.

The main objective of the application is help people to pass message during natural disaster time using messages. People trapped in some location needs help by using our app they can easily pass the message to their friends in their account list. Suppose a friend is living in foreign nation and is native is India while India having a extreme whether condition like storm or earthquake or tsunami like that. The foreign friend may worried about his family and friends and other people. To establish a message medium we use messages so that people living in India can message saying "Heavy Strom but Staying Safe. Thank for all praying". This message will reflected to their friend timeline so each and everyone can be calm and relax without worrying. To handle tremendous message and a fraction of second we need to place the system in a clustering manner i.e distributed framework approach. By doing this we can easily load the balance of message both at uptime and average time. To fulfill this we use kafka a pub-sub message process .all those messageed messages are published and subscribed into kafka queue and process triggered. Storm pulls those message and store into DB. This wills parallelism the framework for clustering process.

In system crawl numerous messages related to target events and use a probabilistic model to extract events from those messages and estimate locations of event and finally developed an event notification or reporting system that extracts event from Online Social Network and sends a message to registered users. To obtain messages on the target event precisely and apply semantic analysis over a message to understand the concept of the posted messages. I also use this message information to forecast a particular event by applying statistical tools over the extracted data model.

Messages might include mention of the target event. If a message is referring to the target event, it might not be appropriate as an event report. For instance, a user makes messages such as the cyclone yesterday was scary. These messages are truly descriptions of the target event, but they are not real-time reports of the events. Therefore, it is necessary to clarify that a message is truly referring to an actual contemporaneous cyclone occurrence, which is denoted as a positive class. To classify a message as a positive class or a negative class and also use a support vector machine, which is a widely used machine-learning algorithm. By preparing positive and negative examples as a training set, and it can produce a model to classify messages automatically into positive and Negative categories.

There are three groups of features for each message as described below. Features a (statistical features): the number of words in a message message, and the position of the query Word within a message. Features B (keyword features): the words in a net.6. Features C (word context features): the words before and after the query word. It can search the message and classify it into a positive class if a user makes a message about a target event. The user functions as a sensor of the event. If she makes a message about a cyclone occurrence, then it can be considered, as a cyclone sensor returns a Positive value. A message can therefore be regarded as a sensor reading. This crucial assumption enables application of various methods related to sensory information. The motivations are the same for both cases; to detect a target event. Observation by sensors corresponds to an observation by Online Social Network users. They are converted into values using a classifier. The virtual sensors (or social sensors) have various characteristics of some sensors are activated (i.e., make messages) only by specific events, although others are activated by a wider range of events. The sensors are vastly numerous and there is more than 100 million "Online Social Network sensors" presented in the World Wide Web producing Message information around the clock.

SYSTEM ARCHITECTURE

System architecture is the conceptual design that defines the structure and behavior of a system. An architecture description is a formal description of a system, organized in a way that supports reasoning about the structural properties of the system. It defines the system components or building blocks and provides a plan from which products can be procured, and systems developed, that will work together to implement the overall system. This may enable one to manage investment in a way that meets business needs. Here we explain methods using a weather reporting as a target event. First obt ain messages on the target event precisely and apply semantic analysis of a message.

The fundamental organization of a system embodied in its components their relationships to each other and the environment and the principle governing its design and evolution. The composite of the design architectures for products and their life cycle processes. A representation of a system in which there is a mapping of functionality on to hardware and software components, a mapping of the software architecture on to the hardware architecture and human interaction with these components.

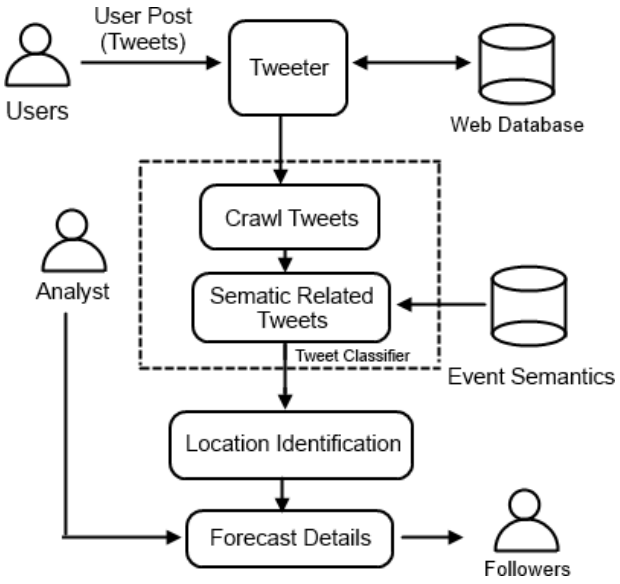


Figure 5.1 System architecture

An allocated arrangement of physical elements which provides the design solution for a consumer product or life-cycle process intended to satisfy the requirements of the functional architecture and the requirements baseline. Architecture is the most important, pervasive, top-level, strategic inventions, decisions and their associated rationales about the overall structure (i.e., essential elements and their relationships) and associated characteristics and behavior and proposes an event notification system that monitors messages and delivers notification promptly using knowledge from the investigation. In this work there are three steps: first, crawl numerous messages related to target events, second probabilistic models to extract events from those messages and estimate locations of events; finally, developed a weather reporting system that extracts information about weather from Online Social Network and sends a message to registered users.

The system architecture diagram will show how the process may be done. At first, the user can post the message in the messageer and the message can be stored in the b database. Through b data base the analyst can analyze the message. Generally, analyst can crawl the message from the messageer the information presented in b database and classify the message using event based on semantics messages related to the target event. At first generally can crawl the message related to event when using the semantic can get the information not only the related event and also various information about the event. By having this information the analyzer can analyze the message about the target event and also the same event can be occur before or not. If the same event may be occur in before by having that information like where the event may be occur and also how it will occur can be analyzed by the analyzer. After analyze these the analyzer can easily identify the location of the event and forecast the message to the followers. The following figure shows the system architecture diagram.

Particle filters of three kinds perform well compared to other baseline methods. Particle filter with weighting works sampling is similar to that of the particle filter with weighting when Ns ¼ 100(Ns, number of samples) and sampling by mean value method. The data are shown for comparison of the weighted particle filter and the sampled particle filter with each sampling method in performance and time complexity. Mean values work better than other sampling methods do. The performance of the sampled particle filter with mean value has a positive correlation with Ns; it converges Ns

¼ 300, which means that 5 percent of all sensors are sufficient for sampling.

cyclone that occurred is as shown in the figure3.8. Many messages originated from over a wide region. The estimated location of the cyclone (shown as estimation by weighed particle filter) is close to the actual weather report center, which shows the efficiency of the location estimation algorithm and compare results obtained using three particle filtering[51] methods with the weighted average and the median as a baseline. The weighted average simply takes the average of latitudes and longitude on all the positive messages; median simply takes their median.

cyclone that occurred is as shown in the figure3.8. Many messages originated from over a wide region. The estimated location of the cyclone (shown as estimation by weighed particle filter) is close to the actual weather report center, which shows the efficiency of the location estimation algorithm and compare results obtained using three particle filtering[51] methods with the weighted average and the median as a baseline. The weighted average simply takes the average of latitudes and longitude on all the positive messages; median simply takes their median.

EVALUATION OF SPATIAL ESTIMATION

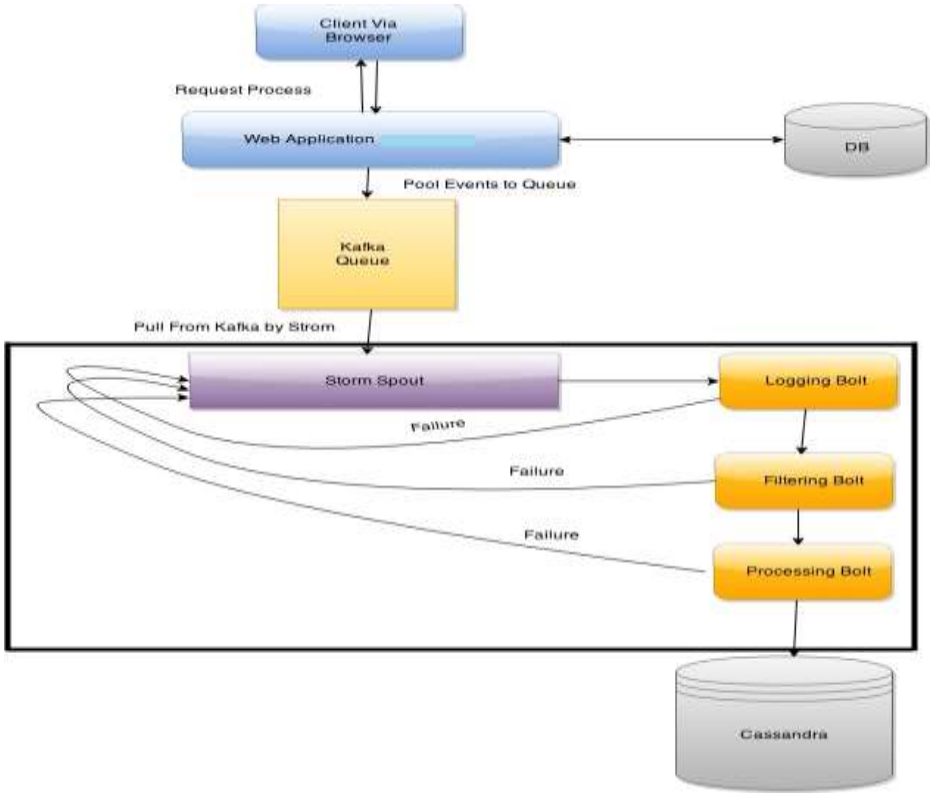
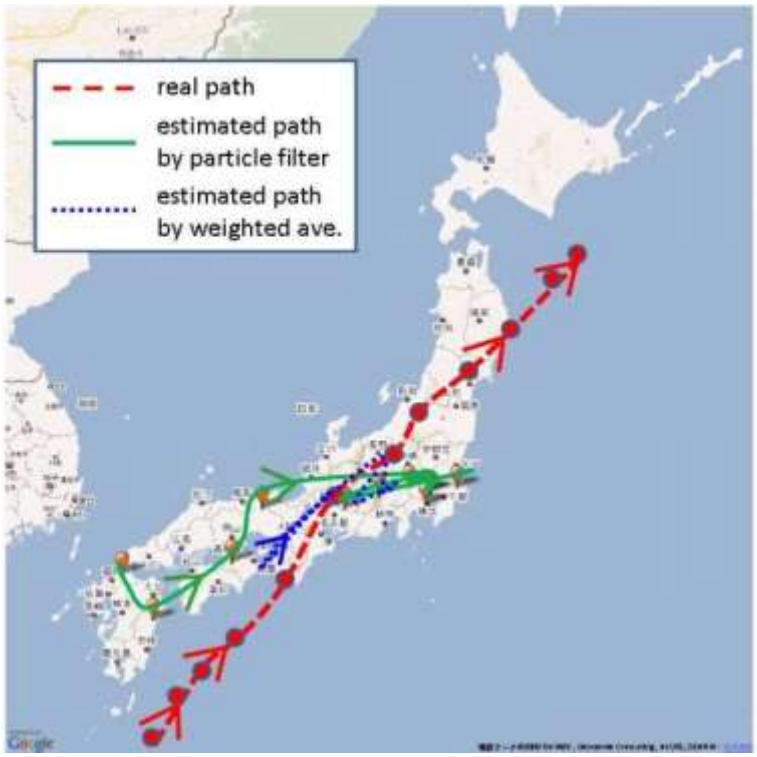


Figure 5.1.1 System Architecture Using Cassandra



It takes 6:4 s for calculation by sampled particle filter with Ns ¼ 300; it takes 120 s for calculation using the weighted particle filter and perform computations 20 times faster than before with only a slight drop in performance. Results show that if the center of the cyclone is in a weather area, it is more difficult to locate it precisely from messages. Similarly, it becomes more difficult to produce well Estimations in less-populated areas. That result is reasonable: all other things being equal, the greater the number of sensors, the more precise the estimation will be. For a cyclone, the center is one location. However, for a typhoon, the center moves, producing a trajectory. The relative performance of everal methods is presented. The particle filter Works well and outputs a trajectory path resembling the actual path of the typhoon.

EVENT DETECTION SYSTEMS

Cyclone detection[1] systems are developed using a methodology of event detection system and explain the features.

CYCLONE REPORTING SYSTEM

Cyclone-reporting system can be developed using the event detection algorithm. Cyclone information is much more valuable if it is received in real time. It goes without saying that, for such a warning, earlier is better. Vast amounts of work have been done on intermediate term cyclone prediction in the Daisy Lab. Various attempts have also been undertaken to produce short-term forecasts to realize a cyclone warning system by observing electromagnetic emissions from ground-based sensors and satellites. In India, the government has allocated a considerable amount of its budget to mitigating cyclone damage. In fact, a cyclone early warning service has been operated. It provides advance announcements of the estimated seismic intensities and expected arrival times.

REPORT PREDICTION

The proposed system, has been operated a system screenshot is depicted. Users can see the detection of past cyclones they can register their e-mails to receive notices of future cyclone detection reports. It alerts users and urges them to prepare for the imminent cyclone. It is hoped that a user receives the e-mail before the cyclone actually affects that area and evaluate various conditions under which alarms might be

sent to choose better parameters for proposed system. I set alarm conditions as N message a positive message comes in 10 minute and evaluate those methods by Precision ¼ N cyclone Nalarms and Recall ¼ N cyclone All cyclone (N cyclone: Number of cyclone detected correctly, Nalarms: number of alarms, All cyclone: number of all cyclone that occurred) can be detected 93 percent of cyclones that re stronger than daisy lab. The precision is very low, which means the system produces many false-positive alarms in such cases ready investigated the reasons underlying errors of system. These errors are divided into errors of two types. The first type is the case of detecting one cyclone several times and designates such errors as "multiple detection." The second type includes cases other than "multiple detection. I designate this type as "incorrect detection." If people receive several alarms in short time span, they can understand that those alarm come from the same cyclone.

CONCLUSION

Online Social Network is a social network platform that consists of billions of users all over the world where people collaborate and Share information related to real world events. These large scales message data processing are done by placing those message events in a distributed system. Message in the message queue is done by Apache Kafka which is a distributed publish- subscribe messaging system. These frameworks will parallelize our computations over a cluster of machines.

FUTURE ENHANCEMENT

The future work is to show the claimed, assumption, while strong, is quite logical considering the information collective in the Online Social Network . Collecting these large data will impact processing to overcome it we can use Hadoop map-reduce concept for Big Data Processing via clustering. Developing a distributed framework will extremely high for reducing the cost in near future cloud computing can be used for low data consumption as well as high performance. Clustering the server also increase the number of working machines as a backup. A stable architecture will place it.

REFERENCES

[1]. Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo(2013) "Message Analysis for Real-

Time Event Detection and Earthquake Reporting System Development",IEEE

[2]. K Sundharivel, Udaya Suriya Rajkumar(2014) "Real Time Natural Disasters of Data Analaysis Using Message" .

[3]. R. Anitha, S. Selvakumar, V. Anusha Sowbarnika, and N. Varatharajan(2014) "Real-Time Event Detection and Earthquake".

[4]. Bala Kumari P1, Bercelin Rose Mary W2 and Devi Mareeswari(2014) Filtering Noisy Contents in Online Social Network by using Rule Based Filtering System.

[5]. Ch. Pavan Kumar, B. Raghu Ram, and B. Hanmanthu(2013) "A Machine Learning Approach to Riddle Abused Scraps from OSN User Space".

[6]. Sandhya Rani Vemparala, Sayeed Yasin(2012): "Preventing Attacks in Multi-path Wireless Routing Protocols".

[7]. Dharnasi Rama Devi, Ms. Naga Aswani Puduru(2014) "A Probabilistic Model of Real Time Event Detection and Reporting".

[8]. G.L. Danah Boyd and S. Golder, "Message, Message, Remessage:Conversational Aspects of Remessageing on Online Social Network ," Proc. 43rd Hawaii Int"l Conf. System Sciences (HICSS-43), 2010. [9]. Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo, "MessageAnalysis for Real-Time EventDetection and EarthquakeReportingSystem Development".

[10]. Vanithasri, K. and Tamilmani.G "Natural Disaster Recommendation Speedy Emergency Alert system using Online Social Network for social networks" [2012].

[11]. S. Anand, and K. Narayana(2013). "Real-time Interaction of Events Detection Earthquakes".

[12]. Wenwen Dou, Xiaoyu Wang, William Ribarsky,Michelle Zhou "Event Detection in Social Media Data".

[13]. Amol jadhao, SweetyKumari, Firdos Khan, Shekh Sultan, Ruchita Khandge.(2015). "Real-Time Detection of Traff ic From Online Social Network Stream Analysis".

[14]. Thi Zin, Member, IAENG, Pyke Tin, Hiromitsu Hama and Takashi Toriu(2013). "Knowledge based Social Network Applications to Disaster Event Analysis".

[15]. Naveed Ahmad,(2010). "Flood Prediction and Disaster Risk Analysis using GIS based Wireless Sensor Networks".

[16]. Sandeep Panem, Manish Gupta and Vasudeva Varma. "Entity Tracking in Real-Time using Sub-Topic Detection on Online Social Network ".

[17]. Mark Cameron, Bella Robinson, and Robert Power(2012). "Using Social Media to Enhance Emergency Situation Awareness Jie Yin, CSIRO ICT Centre Andrew Lampert, Palantir Technologies".

[18]. Mita K. Dalal and Mukesh A. Zaveri(2014). "Opinion Mining from Online User Reviews Using Fuzzy Linguistic Hedges".

[19]. M. Okazaki, and Y. Matsuo(2010). "Earthquake Shakes Online Social Network Users:Real-time Event Detection by Social Sensors".

[20]. Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Online Social Network : What 140 characters reveal about political sentiment. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington DC. [21]. G.L. Danah Boyd and S. Golder, (2010)"Message, Message, Remessage: Conversational Aspects of Remessageing on Online Social Network ," Proc. 43rd Hawaii Int"l Conf. System Sciences (HICSS- 43).

[22]. B. Huberman, D. Romero, and F. Wu,(2008) "Social Networks that Matter: Online Social Network Under the Microscope," ArXiv E-Prints, http://arxiv.org/abs/0812.1045.

[23]. A. Java, X. Song, T. Finin, and B. Tseng, "Why We Online Social Network : Understanding Microblogging Usage and Communities," Proc. Ninth WebKDD and First SNA-KDD Workshop Web Mining and Social Network Analysis (WebKDD/SNA-KDD "07)

[24]. Joachims, Baroper str. Dortmund "Text Categorization with Support Vector Machines: Learning with Many Relevant Features"Thorsten University at DortmundInformatik LS8,
Germany

[25]. Kerstin Borau, Carsten Ullrich, Jinjin Feng, and Ruimin Shen "Microblogging for language learning :Using Online Social Network to train communicative and cultural competence" Distance Education College of Shanghai Jiao Tong University.

[26]. H. Kwak, C. Lee, H. Park, and S. Moon, (2010)"What is Online Social Network , A Social Network or A News Media?" Proc. 19th Int"l Conf. World Wide Web (WWW "10), pp. 591 600,

[27]. Stefania Montani, Giorgio Leonardi, Alessio Bottrighi, Luigi Portinale, and Paolo Terenziani ,(2013)"Supporting Flexible, Efficient, and User-Interpretable Retrieval of Similar Time Series", IEEE transactions