Crowdsourcing of Environmental data using Mobile Sensor Cloud

Isha Patel, Priya Singara Velu, Saraswathi Thanu & Shraddha Kannav
San Jose State University
Prof. Chandrasekar Vuppalapati

Abstract

Sensors have become an integral part of today's technology. Today, the use of sensors is rapidly increasing than never before. This increase in sensors has led to the exponential increase in sensor data. The conventional methods of controlling and managing the sensors are through sensor networks and through Internet of things (IoT) web portals. Each of these methods has their issues when it comes to handling a large number of sensors and huge volumes of data. Mobiles will attract lot of applications from different domains like Environmental Monitoring, Smart city, Social Networking, Security and Safety and mainly Healthcare. Sensing as a service and providing the sensor data to the normal user which is real time using mobile sensor cloud and crowd sourcing. Using different mobiles as a sensor provider, we will get lot of sensor real time data to manage and provide to the users who need to access it. This will increase the accessibility in data sharing and scalability of the sensor networks. With the use of cloud we can provide this data anywhere anytime. In this research paper, we are considering the environmental related data to provide the users with an interactive Smart City map with real time information about the temperature, radiation intensity, humidity and moisture level and send alert messages to the user when there is any abnormal rise in temperature or radiation level. Encouraging users about green environment by planting trees or plants via messages.

Keywords*: Mobile Sensor Cloud, Crowd Sourcing, And Environmental Sensors*

Introduction

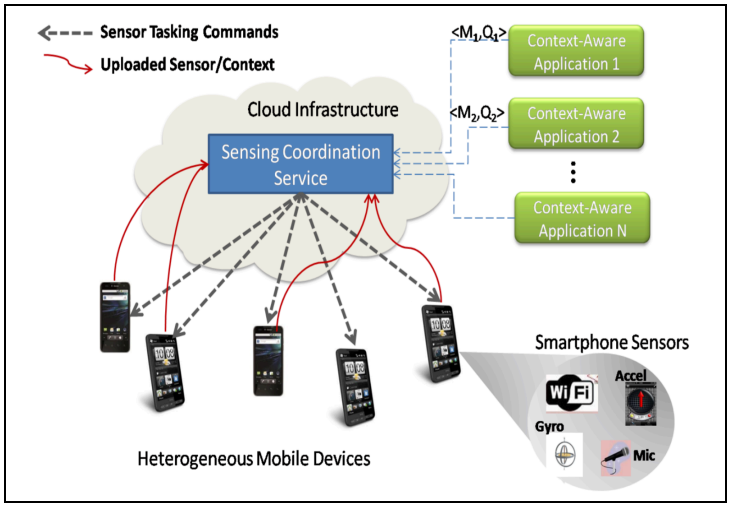
Now days, mobiles-Smart phones are widely used for many purposes. These smart phone applications provide cloud computational augmentation. Sensors from these smart phones detect and provide proper notifications to the person who is using it. Then why not change it globally.

Combination of the sensor provision and cloud technologies will create a mobile sensor cloud, which will be more scalable and accessible and inexpensive in-terms of deployment and management of the sensors. By using such data multiple context aware applications can be developed. So the applications will be used widely. Embedded sensors like Camera, GPS accelerometer, gyroscope, microphone, and network statistics and so on can be register to produce mobile sensor infrastructure as a service to the applications.

As every system need to satisfy certain requirements. It should be able to work from different mobile platforms with the help of the application. Secondly, it should be energy efficient, as data will be continuously uploaded to the cloud. Third, it should use less network resources. This Mobile sensor cloud will also need to attract user to participate in the in the sensing activities.

In this system, we will consider each mobile device as independent entity and the sensor data of each entity will be collected and computed accordingly. Continuous sensing and processing of real time data will be invaluable and provided to wide range of consumers and enterprise applications.

Here, comes the part of crowd sourcing. Many enterprise applications need similar kind of data from various sources to analyze the situation on proper scale.



We are going to discuss more about the applications where we can use the sensor data. There are multiple areas where we can think of using this data, HealthCare, Environmental Scenarios, Smart city, Smart Traffic solutions, Emergency vehicle alerts, and many more… These applications will have the sensor data along with the customers who are participating in providing such data.
In this article we will more talk about

- Environmental Sensors in Mobile phones
- Hybrid Crowdsourcing applications
- Design and Architecture of the System
- Big Data analytics of real time sensor data streams
- Threats in mobile crowdsourcing: Privacy, Security and Data Protection
- Energy efficient sensing task management
- Proposed cloud architecture and its issues

Environmental sensors in Mobile phones

According to ITU (International Telecommunication Union), most of the people by 2015 will start using mobiles and will be covered under some network. Every mobile has various sensors inbuilt. These sensors give us certain data of the person using it. In last decade many Wireless communication technologies have evolved like Bluetooth, Wi-Fi, Mobile Communication (GSM) etc.

All this prompts us to think about ways for collecting this huge massive data and use for our benefit. Sensor data collected for a huge crowd can be strong enough to recognize patterns, pollution at various places, help evacuating places etc. There are many powerful sensors attached in mobile phones today like camera, accelerometer, microphone etc. Applications in play store generally focus on particular sectors like entertainment, health, network, safety, news etc. They require different sensors for their operation. Based on this observation there are two categories Mobile Sensors can be divided. First ones are positioning sensors or orientation sensors. They just detect the position of the mobile and its orientation in terms of different parameters.

Second ones are environment sensors. These sensors give details about the surrounding environment of mobile. Examples of orientation sensors include accelerometer, magnetic compass, gyro meter etc. Environment sensors include humidity sensors, light sensors, proximity sensors etc. The use of sensors can be further enhanced to communicate with real time objects in the world, for some local analysis or to study some intangible factors of our environment.

Recent advancement in technology, energy emission has brought serious threats on our environment. People are getting more health conscious and try to live in a habitat, which nurtures their health. Taking into consideration the fact that almost all people now a day carry phones, we can make smart use of sensors to alert them about environmental hazards beforehand. Some of the in-depth descriptions of mobile sensors, which can be used to fulfill this purpose, are:

a. Accelerometer sensor:

This sensor is used to measure proper gravity force on the device. This can be further explained

as if your mobile is stationary it will give a value of 9.8 upwards. Whereas, if your mobile is falling freely it will give value 0. It is not restricted to have just one axis. It can have multiple axis which determines direction as well as power magnitude in that direction. It can also be used to measure differences in gradients of various altitudes. There is a concept to measure waves of gravity for in-depth analysis. It can be designed to calculate that as well. Sometimes air resistance can degrade the actual calculation but it's nearly accurate. Some practical applications for environmental protection is to use it to know the actual impact of surroundings on your jogging speed, rate at which vibration and energy expending occurs, photochemical monitoring etc. It can also monitor human activities due to its power to calculate motion and reaction to surrounding dynamic load exposure. Some activities like skipping, walking, jogging can be analyzed keeping in mind the surrounding environment exposure.

b. Light Sensors:

Light sensors detect the light sensitivity from the surroundings. Research has shown that excess amount of hazardous gases in the surroundings can affect the light intensity. Infrared rays in the light are highly affected by the excessive presence of gases like carbon dioxide, carbon monoxide and other Sulfuric gases. A drastic variation in some of these gases affects the intensity of infrared rays a lot. A thorough analysis done on the levels and thresholds of infrared levels can help us form a system to notify people when they approach such areas. By using location sensors we always have an idea about people's location. In very highly famous areas we generally can form a pattern. When a hazardous event occurs notification can be sent to these groups of people or a person can use this feature for their own benefit. They can make sure they stay in a clean environment by constantly checking variations in the

c. Temperature sensors:

Presence of harmful chemicals can affect the temperature to a great extent. When multiple sensors used collectively can enhance the performance to detect surroundings. Accuracy in this case is maximum because features of different sensors can be used collectively to enhance precision. Every chemical affects the electrical resistance in the air. If a person is hazardous to some chemicals, this set of rules can guide him to stay away from certain areas or notify when they are about to enter certain areas.

Hybrid Crowdsourcing Applications

In recent few years, many applications have started using crowd-sourced approach or participatory sensing approach to collect data from various entities and ordinary public. All this is made possible by rapidly growing affordability, availability and adoption rates in recent months and years of Internet-enabled and location-aware mobile devices, such as tablets and smartphones.

Google earth: This application by Google is widely used amongst the world. This is the world map. People contribute to it by providing photos directions and other content like videos and 3-D model of the location. It is the major crowd-sourced application, which is being widely used by many smartphone users as well.

'Geographic information System (GIS) by masses' is the phrase used by Kamel Boulos in 2005 [6].

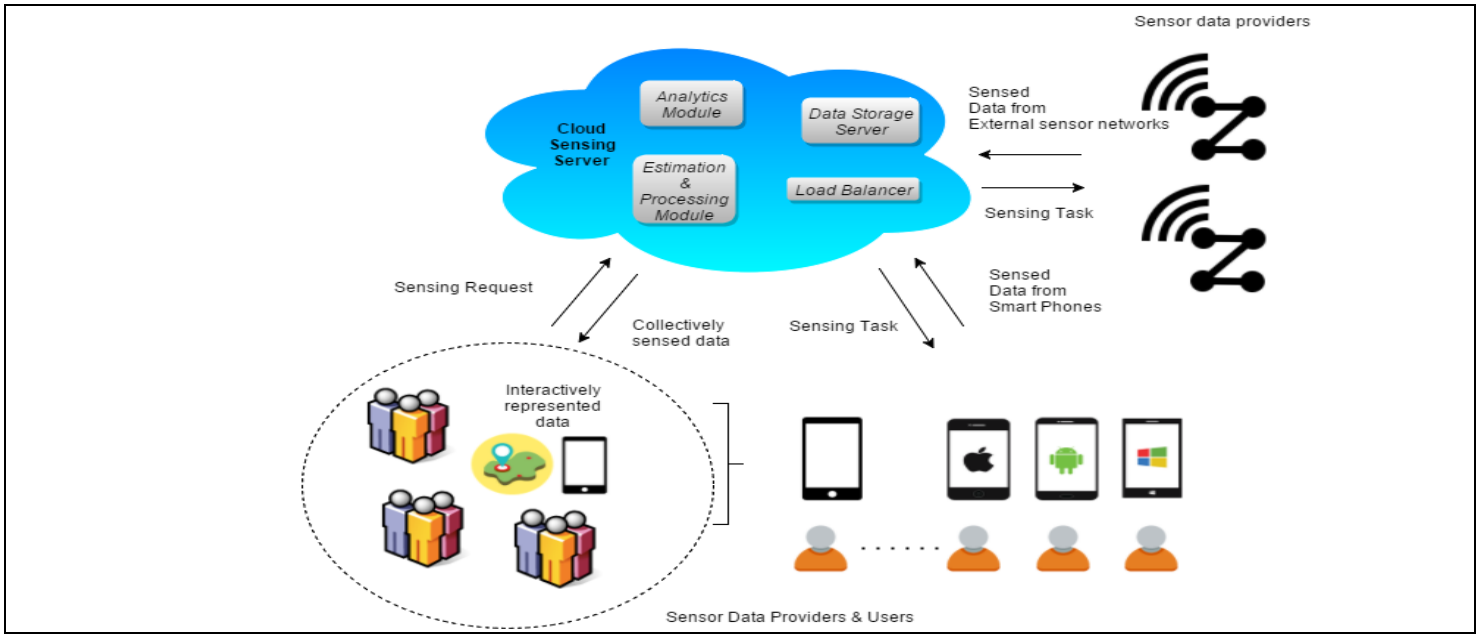
As this is the volunteered process that people allow using their data for crowd-sourcing applications and using the mass data for better analytics. Goodchild [7] came up with the term 'Volunteered Geographic Information (VGI)'

Global Positioning System: This application use travelers traces and upload it over Internet, which can be used for real time traffic information. Here Geo tagging and Context tagging is used.

Noise pollution detection: Using Smart phone of foot travelers and their location, the application is created using geo tagged and street level audio samples by uploaded data. This provides citywide noise pollution in the form of interactive graphical maps for the specific time of the day in a week.
So, to summarize there are lots of applications which are based on the sensor input and basic analysis of the data. The applications such as Love clean streets, Outbreaks near me, Med Watcher and Fire safety application for forests. Addition to these there is lifesaving apps like Citizen Responders, which will help Cardio pulmonary Resuscitation (CPR) as emergency services report to cardiac arrest

Design and Architecture

Infrastructure of the System:



As discussed in earlier topics, Sensor Infrastructure is used to mange all the sensors and the data. But to allow Mobile sensor cloud infrastructure to take over the overhead of deployment and management of the system. We have designed a simple infrastructure to access mobile sensors upload the sensor data to the cloud with user permission and provide is to group of users of industries with proper analytics.

Starting with Smartphone owners, they will allow the application to upload the sensor data to the cloud, which is cloud-sensing server. This server has different modules, which will collect access and process the data.

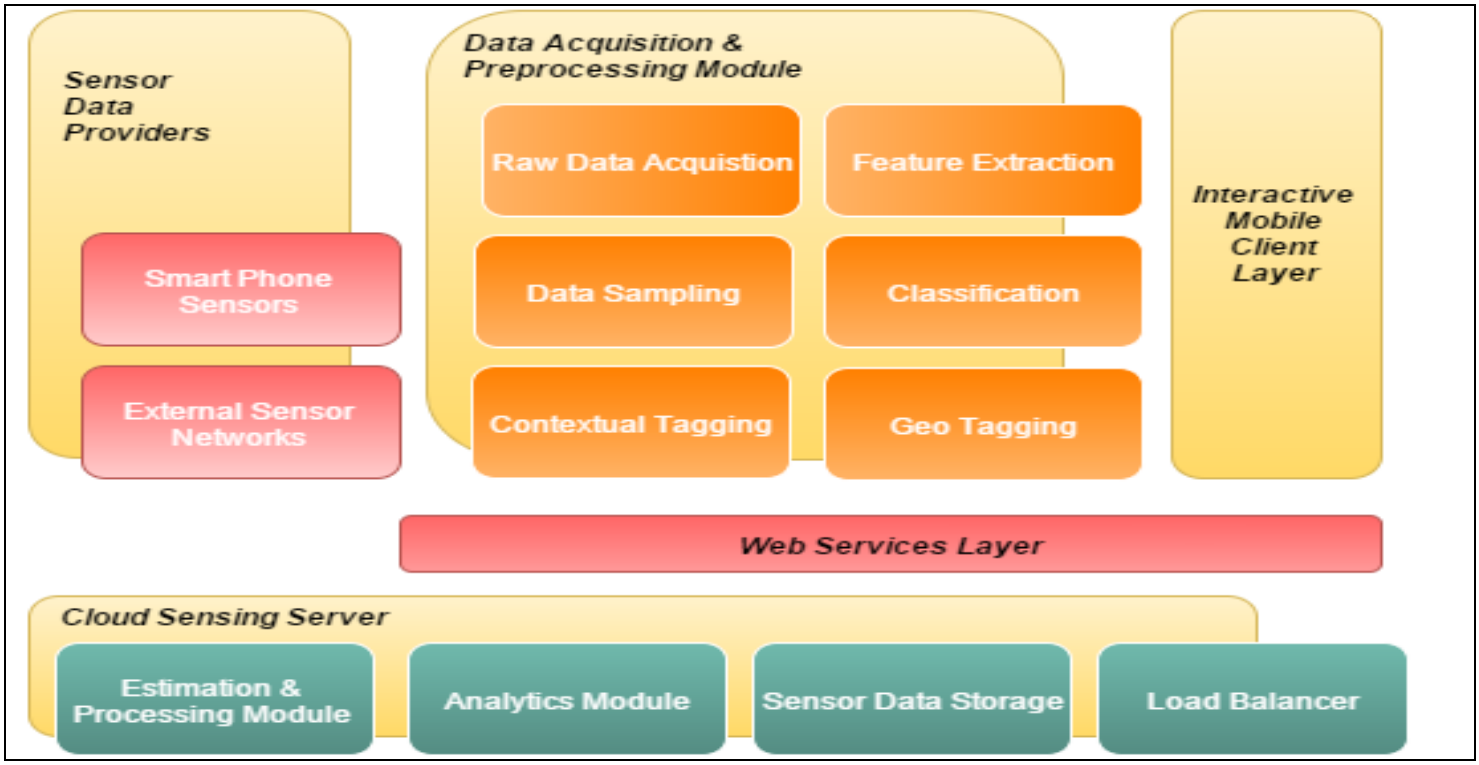
Data Storage Server: This will collect the data from various sensors of different mobile platforms and store in the cloud. Here we will need big data storage.
Load Balancer: The data, which is getting uploaded in the cloud, is huge. So, balancing which data should go where and how the access should be provided is the task of Load Balancer.

Estimation and Processing module: We have the data, which needs to be processed. This processing of data is done using this module.
Analytics Module: Analysis of the data is most important factor. As when we provide this data to users and industries it should be properly formatted and readable format. So, analytics module will analyze the sensors and data

Now, we only talked about sensors from Smartphones. But we can also consider the data from External Sensor networks. So as to use existing sensor networks in the cloud.

Important factor of the infrastructure is users or industries that will be using this data. An interactive real time data to such users by sending requests to the cloud and collectively sensed analytics of the data can be represented on the simple mobile app or providing APIs to access it.

Architecture of the System:



The architecture diagram shows the key components of the crowd sourced mobile sensor cloud system. It has five key components - Mobile Sensor cloud sources, Data Acquisition and Preprocessing Module, Cloud layer, Interactive Mobile Client and the Web Services layer.

Mobile Sensor Cloud Sources: As mentioned earlier, the data providers for the crowd sourced mobile sensor cloud system are the inbuilt sensors in smart phones and any virtually pluggable wireless external sensor networks.

Data Acquisition & Preprocessing Module: The raw data acquisition component involves collection of raw streaming sensor data from various types of sensors like Environmental, Motion, Position sensors. Further processing involves sampling of the data in order to identify patterns, classify the data and build inference model. For e.g. User activity recognition could be used to identify

the regularly visited areas by user and alert the user of pollution levels in these areas. Additionally, the data is geo tagged and contextually tagged. Geo tagging is the process of tagging the data with the geographical coordinates. Contextual Tagging is allowing the user to further describe the context of the data. For e.g., the pollution levels at morning, afternoon and night could vary. Hence, allowing such tagging enables to refine the analysis results.

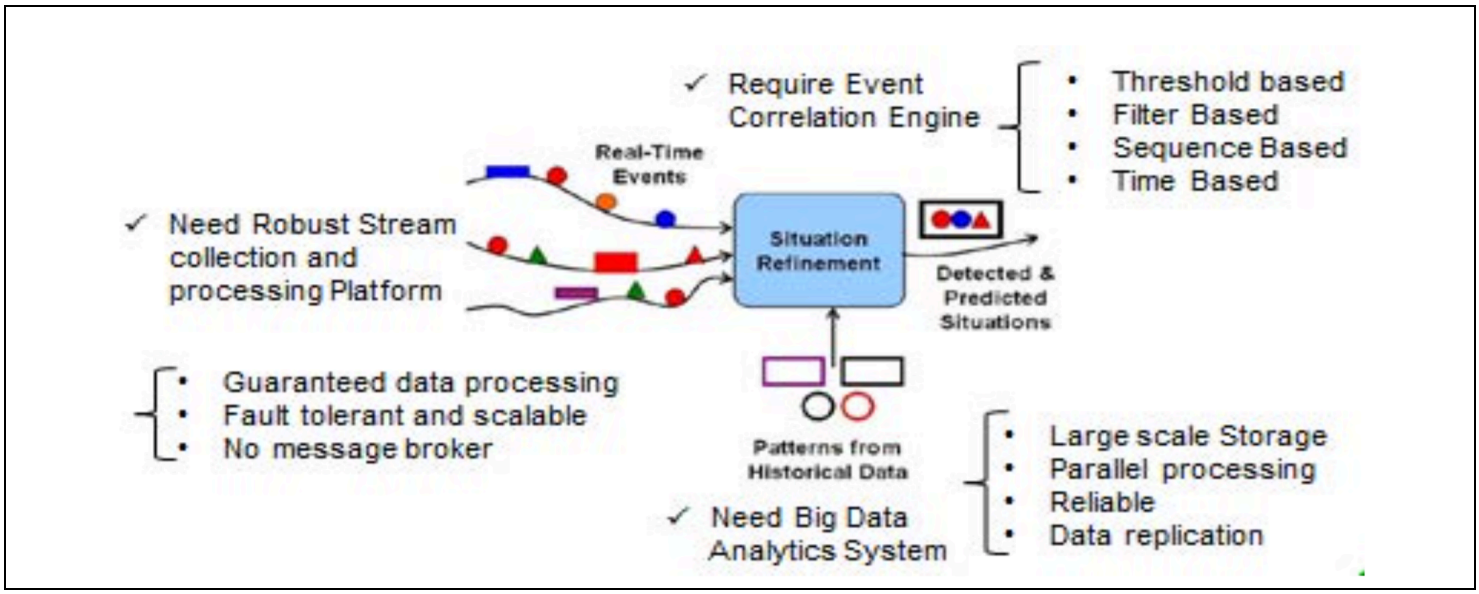
Web Services layer: The use of web services decouples the sensor data providers and the sensing server. This allows for establishing platform independent applications accessing the server and promotes interoperability.

Cloud Sensing Server:

The cloud-sensing server is the pivotal part of the application. It acts as the heart of the system connecting the data providers to the data users. The decision to use cloud services is a natural one owing to the huge amount of continuously streaming data. The cloud-sensing server has components to store the streaming sensor data from the data providers, perform computations on the data, and perform data analysis and mining. The Estimation & Processing module does computations of various sensor data. For e.g., the actual calculation pollution levels from the collected sensor data is done in this module. Analytics module provides a comparison of pollution levels across various regions, various time phases, to get an understanding of the progression that has happened.

Big Data analytics of Real time sensor data streams

The Big Data [11] is not a problem now in terms of storing but when it comes to processing and manipulating data it is important to use big data analysis tools. Hadoop is widely used big data analytics. As we are considering Sensor data, which is Real time and Historic as well, we have to process it and provide predictions about the future using past data. Hadoop is batch-processing system; it also provides offline data processing platforms.



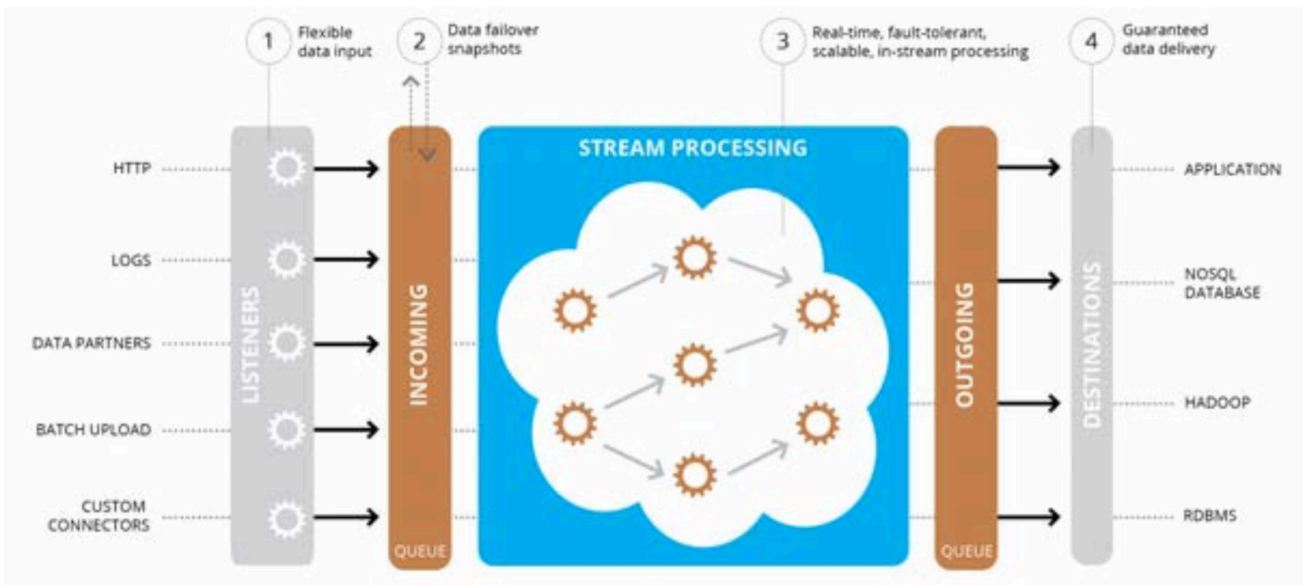
Shows the holistic view of the whole system: the Event Stream Collection system, Event Correlation Engine, and a Situation Refinement using models built by a Hadoop- based Big Data system [11].

We have application where we are using the concept of processing data real time as well as historic moments. But are we sure Hadoop can support this use case?

Real time systems need to have short time windows for analytics. Correlating and predicting events on the basis of streams generated before two three minutes earlier. Prediction needs perfect model to handle such huge amount of the data.

Real time data processing challenges:

1. Volume, velocity and variety of the data three V's of the system affect the performance and processing of the data. Hadoop usually handles Volume and Variety of the data. But real time systems need to handle Velocity
2. Real-time systems should have proper data collections streams, as there are multiple events happening around the world, which is the data for the system.
3. Missing any of the data stream might mislead the prediction of the system and so success of the system.
4. Parallel processing is another most important factor and issue in real time data analysis systems.
5. It should perform event correlation using a Complex Event Processing engine to extract the meaningful information from this moving stream
6. Real time systems need strong network connection at the data source side as any delays in data streams also might lead to incorrect predictions



Streaming data can be collected from various sources, processed in the stream-processing engine, and then write the result to destination systems. In between, the Queues are used for storing/buffering the messages.

Threats in Mobile Crowdsourcing: Privacy, Security and Data Protection

In the past few years, different kinds of crowdsourcing have dramatically increased allowing the people to connect with one another, with governments to connect with public, in coordinating with natural calamities work or to map with the conflicts of political parties and share the data among others. This increased use of crowdsourcing prototypes keeps the public more active and updated with real time information. But, one main concern with respect to crowdsourcing is with the threats associated with it. For example, threats like security, privacy and data protection, which need to be considered at first place.

One such kind of crowdsourcing is mobile phone sensors. The sensors on the mobile phone can be used to attract many sensing applications in the area of environmental monitoring, social networking healthcare, transportation etc.

There are two types of crowdsourcing mobile applications:

1) Participatory systems
2) Opportunistic systems

In participatory systems, users provide the data to the system using the applications from their mobile phones. These applications are similar to the web based applications. In opportunistic systems, there is no user intervention. It has the capability to receive the data directly from sensors on the mobile phone. When these features are enabling in phone devices, sensors can capture the data such as location, image, text and movement without any intervention from the mobile users.

Irrespective of the type of crowdsourcing system, there are some level of privacy and security risks associated with the users. Smartphones play an important platform for crowdsourcing system because of many features like portability, easy access to Internet, environmental sensors like GPS, cameras and accelerometers. Though these features definitely are beneficial in this modern connected world, but they also expose the users with privacy and security threats. Below are some of the threats described:

A. Disclosure of User Identity

These crowdsourcing systems collect basic details about the user such as username, password, email address, contact numbers, which makes information visible to other registered users making the identification public. This may lead to many online attacks such as phishing and spams attacks. This threat can also be extended to the context of Smart city too.

B. Disclosure of User Location and Activity

All smartphones have inbuilt sensors which has the capacity to sense the applications continuously by automatically updating the user location and their movements providing the real time information of their current state. It will also capture the images, audio and text from the surrounding environment. Hence there is no sense of any privacy, confidentiality.

C. Combining crowd sourced data with other user data

With crowdsourcing the data is combined from different sources to result in a new sensitive information which is not in the comfort zone of the mobile user. The ordinary people have no knowledge of how their data is being used at downstream.

D. Lack of User privacy Awareness

In most of the cases, the mobile users are not aware of how their data is actually used and what are the risk factors associated with it. They simply agree to the condition believing that their data is safe and being protected by the application.

E. Vulnerability of mobile devices

The mobile phones are more vulnerable to security threats and breaches when compared to inter-

net security. Because of smaller limited computational power, speed and storage capacity and bypassing many traditional cryptographic techniques. There is a possibility of physical attack such as resetting the device if it falls in wrong hands.

F. Relying on info that may not be accurate.

The crowd-sourced data is combined to get a bigger picture of how the system will take decision on the kind of data being collected. The final outcome purely depends on the users data collected to make any critical conditions. So if there is any data inaccurate then it affects the decision making process. Hence we need to check if the information obtained from the user is legitimate before taking into account for decision-making.

G. Retaliation for Reporting Sensitive Information

Some crowdsourcing systems are used to report sensitive information's. Information's like violence activity or human rights abuses or criminal activities which are extremely important to responders, are publicly seen and can cause risks to victims and the people who report these activities. Hence, the public will trust the system only if it is secure and that they will face no public retaliation for making reports.

Conclusion and future work

The objective of this research paper is to provide every user in the Smart City an interactive city map smart phone application displaying the real time environmental details like temperature, radiation, humidity and moisture level. It will also list all the places where there is a lot of pollution and provide an alert to the people surrounding those places.

Apart from pollution detection, it also pin points where there is an exposure to chemical substances which might lead to serious health issues and does a reporting issue to the respective Government agencies to take necessary steps to control this. All this features is made possible only by the huge benefits of mobile sensor cloud using the crowdsourcing approach in a smart city where we collect data from the embedded mobile smart phone users and also from the Sensor network community and apply the necessary logistics and data mining technologies to provide the exact information in a presentable and user understandable format in the mobile app.

As the usage of crowdsourcing system is increasing in many domains, we need to take into account the level of control in security, privacy and data protection in crowdsourcing. We highlighted some of the
Common issues related to privacy and security threats, which are associated with mobile crowdsourcing systems. This can be helpful for potential contributors while assessing the security risks that they face with different types of crowdsourcing system and to keep them informed about decisions. This will also be considered by sensor network owners, government agencies, smart homeowner's community and administrators of crowdsourcing systems to help minimize risks to users.

Concerning future research in this area, we think there are several interesting directions to be included in our upcoming research paper. First, is to accommodate the features of crisis management system, second is waste management in smart city and providing recycling program to make Smart city, a clean green environment for a healthy and happier living. Lastly, we would also like to focus on disaster warning and response system (natural disaster or crime zone areas), which will provide adequate data for situation assessment purposes.

References

[1]. Sheng X, Xiao X, Tang J, Xue G: Sensing as a Service [978-1-4577-1767-3/12/$26.00 ©2012 IEEE]
http://optimization.asu.edu/papers/XUE/XUE-CNF-2012-SENSORS-SaaS.pdf

[2]. Stevens M. & D'Hondt E.: Crowdsourcing of Pollution Data using Smartphones
http://soft.vub.ac.be/Publications/2010/vub-tr-soft-10-15.pdf

[3]. Boulos M.N.K, Resch B, Crowley D.N., Breslin J.G, Sohn G., Burtner R, Pike W.A., Jezierski E., and Kuo-Yu Slayer Chuang: Crowdsourcing, citizen sensing and sensor web technologies
http://www.biomedcentral.com/content/pdf/1476-072X-10-67.pdf

[4]. Sougata Sen, Archan Misra, Rajesh Balan: The Case for Cloud-Enabled Mobile Sensing Services [ACM 978-1-4503-1519-7/12/08]
http://conferences.sigcomm.org/sigcomm/2012/paper/mcc/p53.pdf

[5]. E. T.-H. Chu, Y.-L. Chen, J. W. S. Liu and J. K. Zao: Strategies for Crowdsourcing for Disaster Situation Information
http://eos.yuntech.edu.tw/eoslab/images/Crowdsourcing%20Strategies.pdf

[6]. Boulos M.N.K.: Creating Interactive maps using Google APIs - International Journal of Health Geographics [doi:10.1186/1476-072X-4-22]
http://www.ij-healthgeographics.com/content/4/1/22

[7]. Goodchild M: Citizens as sensors: the world of volunteered geography. Geo-Journal [2007, 69(4): 211-221]
http://link.springer.com/article/10.1007%2Fs10708-007-9111-y

[8]. Kanhere SS: Participatory Sensing-Crowdsourcing Data from Mobile Smartphones in Urban Spaces. [6-9 June 2011; Luleå, Sweden 2011-3-6]
http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6068482

[9]. Heimbuch J.: Sensordrone Turns Your Phone Into an Environmental Sensor [June 13, 2012]
http://www.treehugger.com/gadgets/sensordrone-turns-phone-envirnomental-sensor.html

[10]. Saenz A: Japan's Nuclear Woes Give Rise to Crowd-Sourced Radiation Maps In Asia and US [24 March 2011]
http://singularityhub.com/2011/03/24/japans-nuclear-woes-give-rise-to-crowd-sourced-radiation-maps-in-asia-and-us/

[11]. Bhattacharya D.: Analytics on big fast data using real time stream data processing architecture [2013 EMC]
http://www.happiestminds.com/sites/all/themes/happiestminds/pdf/ANALYTICS_ON_BIG_FAST_DATA_USING_A_REALTIME_STREAM_DATA_PROCESSING_ARCHITECTURE.pdf

[12]. PollMap - Lunch Break Edition.
http://nutrition.esri.com/lunchbreak/

[13]. GeoChat by InSTEDD (Innovative Support to Emergencies, Diseases and Disasters).
http://instedd.org/technologies/geochat/

[14]. The Ushahidi Platform.
http://ushahidi.com/products/ushahidi-platform

[15]. Crowdmap: http://crowdmap.com/