**HEALTH WEATHER**

**Sai Kishore Bandaru**

**Jagadish Rao**

**Priyanka**

**Ebenezer Anand Arapally**

**Summary:**

As part of second increment we majorly focused on back end i.e. real time processing and collection of data. We are majorly focusing on integrating storm with HBase. Certain amount of work has been spend on developing certain Android pages.

**System Architecture Design:**

The system architecture basically consists of three parts.

In the first part we will be collecting data from the Twitter using STORM. Now using storm we will be collecting tweets and storing in a file. Now using the bolts in storm architecture we will be storing the tweets from file to HBase. Now we will be retrieving the real time data from Hbase either by web service or by HttpClient methodologies. Now to make recommendations to the user based on flu outbreaks, we use the static big data we have collected and analyze those data using the R java. Using R we shall use either naïve Bayes classification or K-means cluster to make recommendations for the users. The following snapshot gives the system architecture:

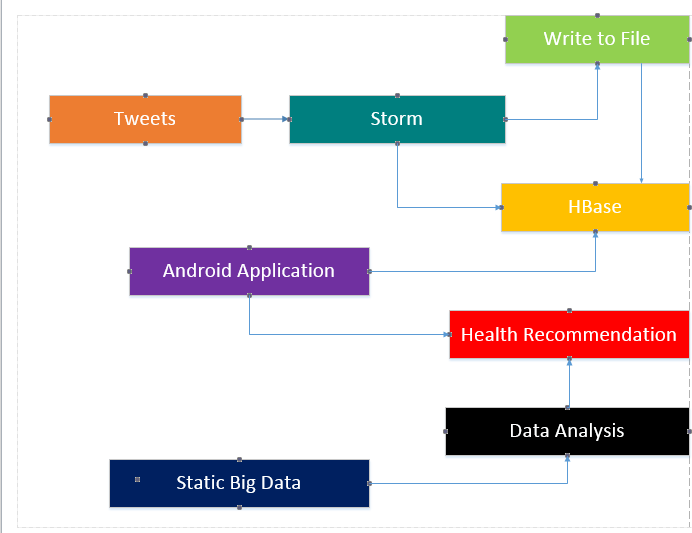


Figure 1System Architecture

**Domain Model:**

**Data Sources:**

For initial phases we will be collecting the flu outbreaks data from the twitter. Once we have the analysis mechanism ready we shall continue our scope of data collection to Image processing or other real time resources available. For recommendation system we have collected data from various health websites and this data consists of information about diseases, symptoms, nearby hospitals which provide treatment etc.

**Algorithms:**

Here to classify each disease based on the symptoms we will be using Naïve Bayes algorithm using either R or weka. Currently our focus is being made on R. The recommendation algorithm is typically an own algorithm. Here in the recommendation algorithm, we have two parts:

1. If the user is registered user we use context based recommendation and suggest the precautions location based on his profile.
2. If the user is a guest then we just suggest the users the precautions but the treatment availability and locations of hospitals respectively.

**Analytical Tools:**

The analytical tool which we will be using for our application initially is R. We will be using R to analyze the static data to classify each data based on the symptoms. Once we get desired results we either store the data on the disk for analysis or we will push that data to Solr or HBase or any other NoSQL tool. We will also use Hue to monitor the data inside the HBase.

**Analytical Tasks:**

The analytical tasks in our application include:

* Context profiling of the registered users for recommendations
* Analyzing twitter data based on the flu out breaks
* Classifying the medical data that is available in order to classify the diseases based on the symptoms.
* Future tasks include image processing.

**Application Specifications:**

System Specifications – Hbase, STORM.

Tools: R

OS: Android Froyo 4.2.0 – Android Kitkat 4.4.3

Analysis Algorithms: Naïve Bayes

**Service Specifications:**

We will be using glassfish web server to make transactions with the Hbase. The web service includes insertion of data into Hbase, Retrieval of data from Hbase. Typically using timestamp technique in HBase we shall collect data from the past one hour or past day.

**Design of Mobile Client:**

Our application is a native mobile application developed from android users. The user can be either guest or registered users. The registered users can receive recommendations from the system. But the guest user can only track of the flus.

The guest user can track the flus going on in each state in USA. The registered need not mention any location details as we will be making a context based recommendation we shall make recommendations based on his profile.

**Implementation:**

**Implementation of Web service:**

We have completed the implementation of web service which serves for making transactions on the HBase. Basically we can make retrieval and input transactions on the HBase. We have ported the war file consisting of HBase transactions on our local glassfish.

**Implementation of Real Time Analysis:**

We have started writing implementation for extracting the tweets and posting them in HBase. We collect tweets in spout of storm and store them in local file. Then we use bolt in storm to forward the data from text file to HBase storage. We made sure that we run storm in local mode so as to make I/O operations.

**Implementation of User Interface:**

We haven’t made much progress on UI. We have implement subset of pages for the guest user like tracking the flus and tracking flus in respective or specific states.

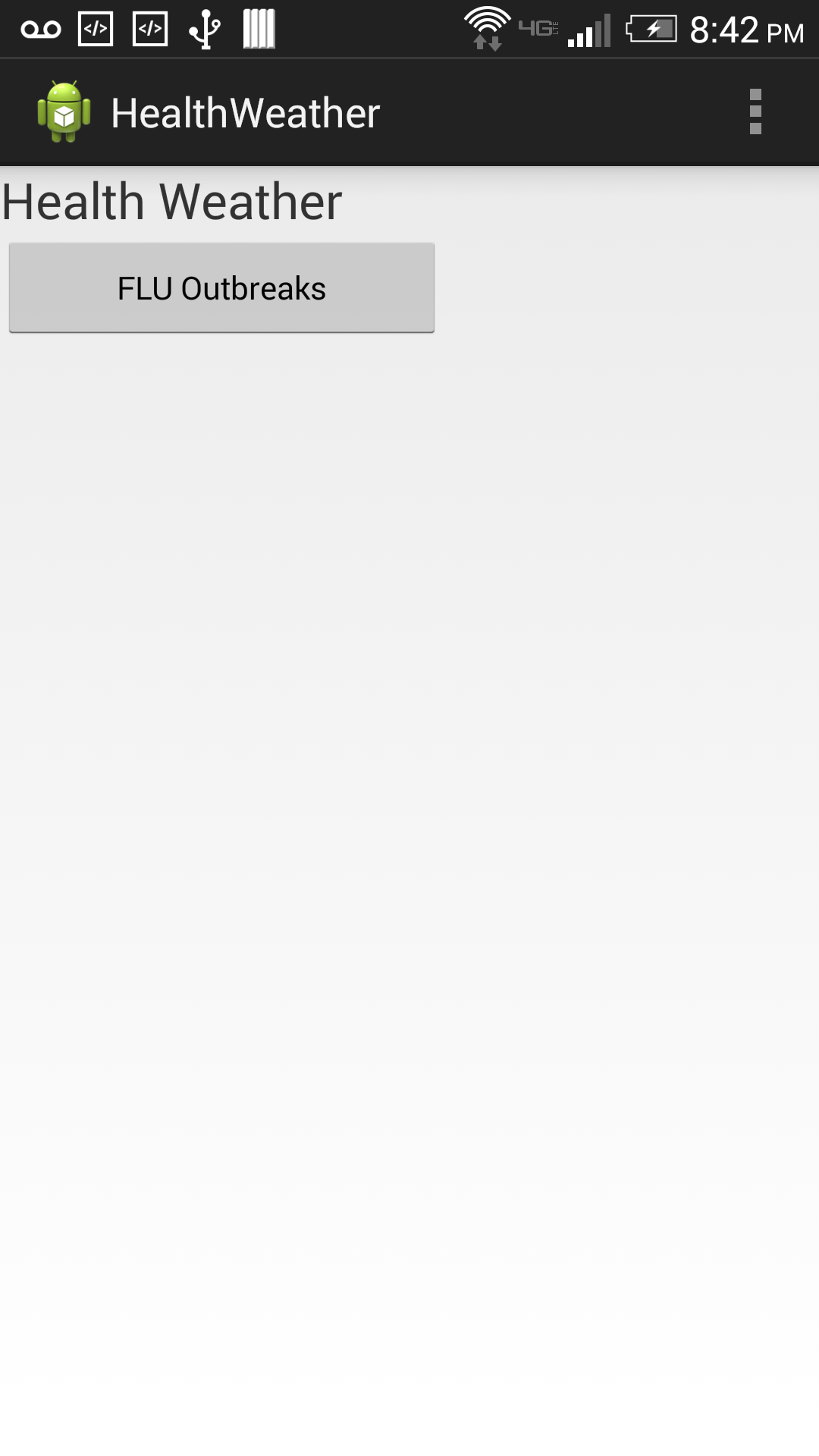


Figure 2Initial page

Welcome page consists of a button which on clicking gives you the total number of flu cases in USA.

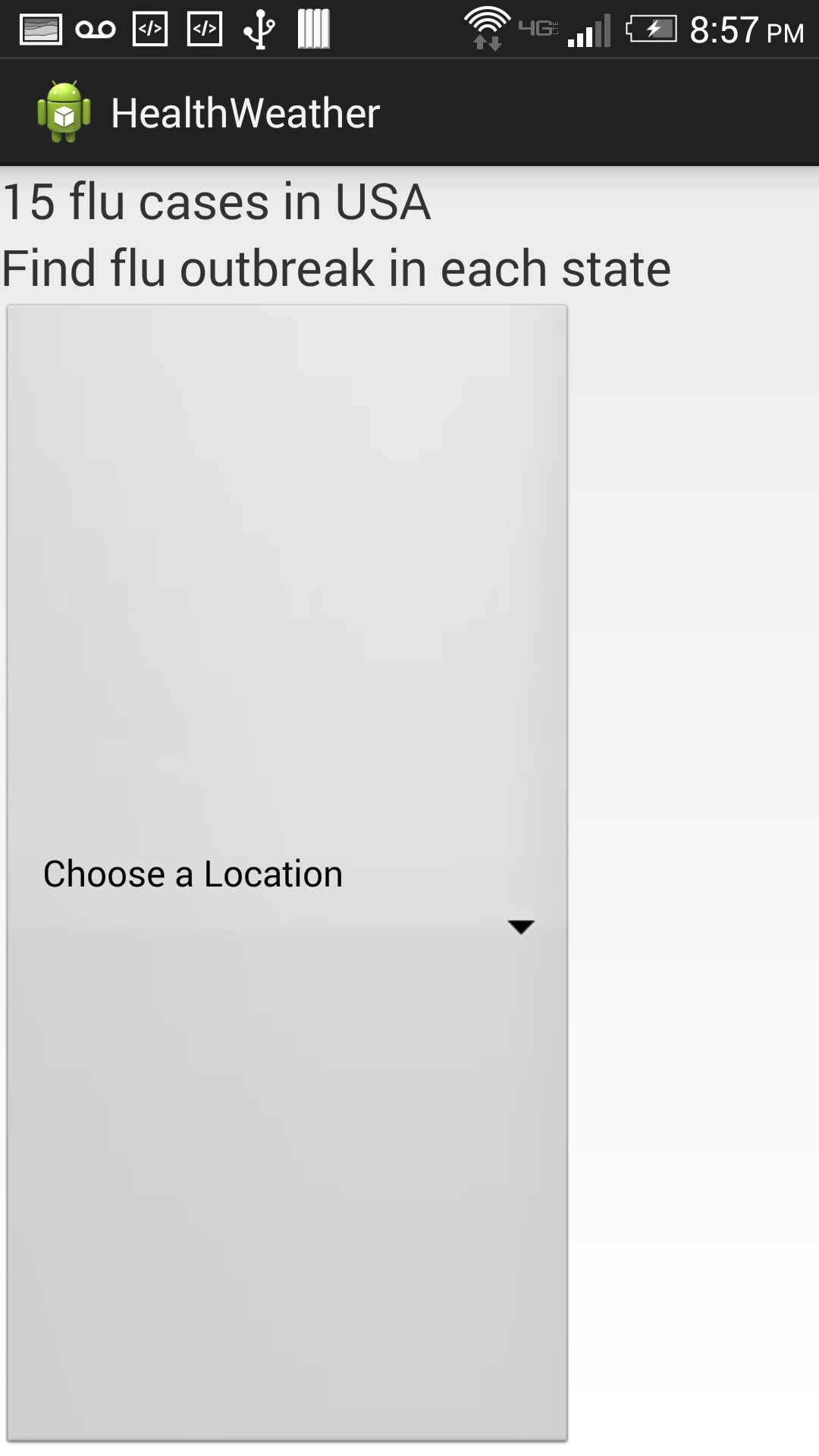


Figure 3Drop Down for guest user

The guest user can choose any state from the drop down to drill down the flu outbreaks.

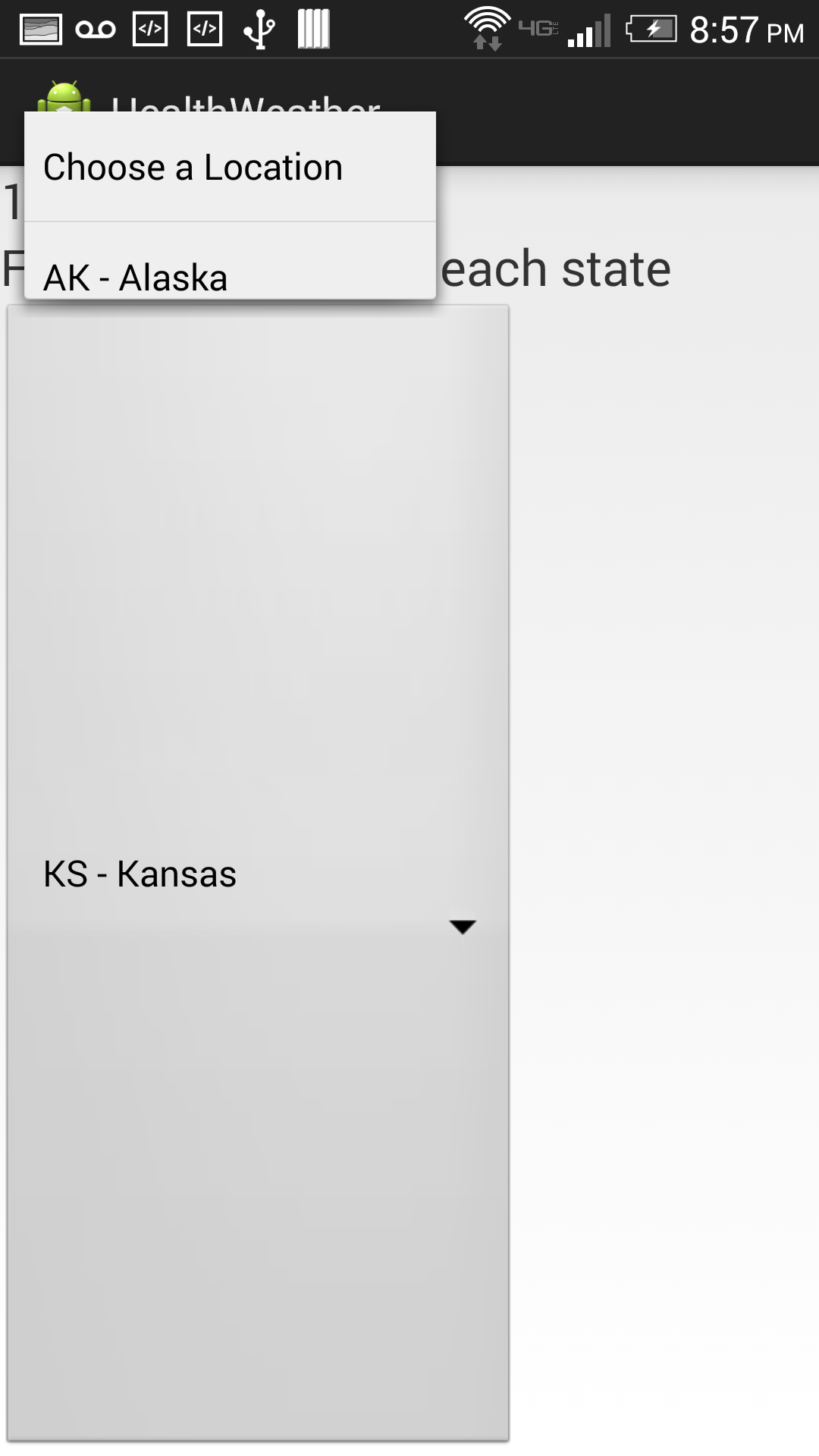


Figure 4Drop Down List

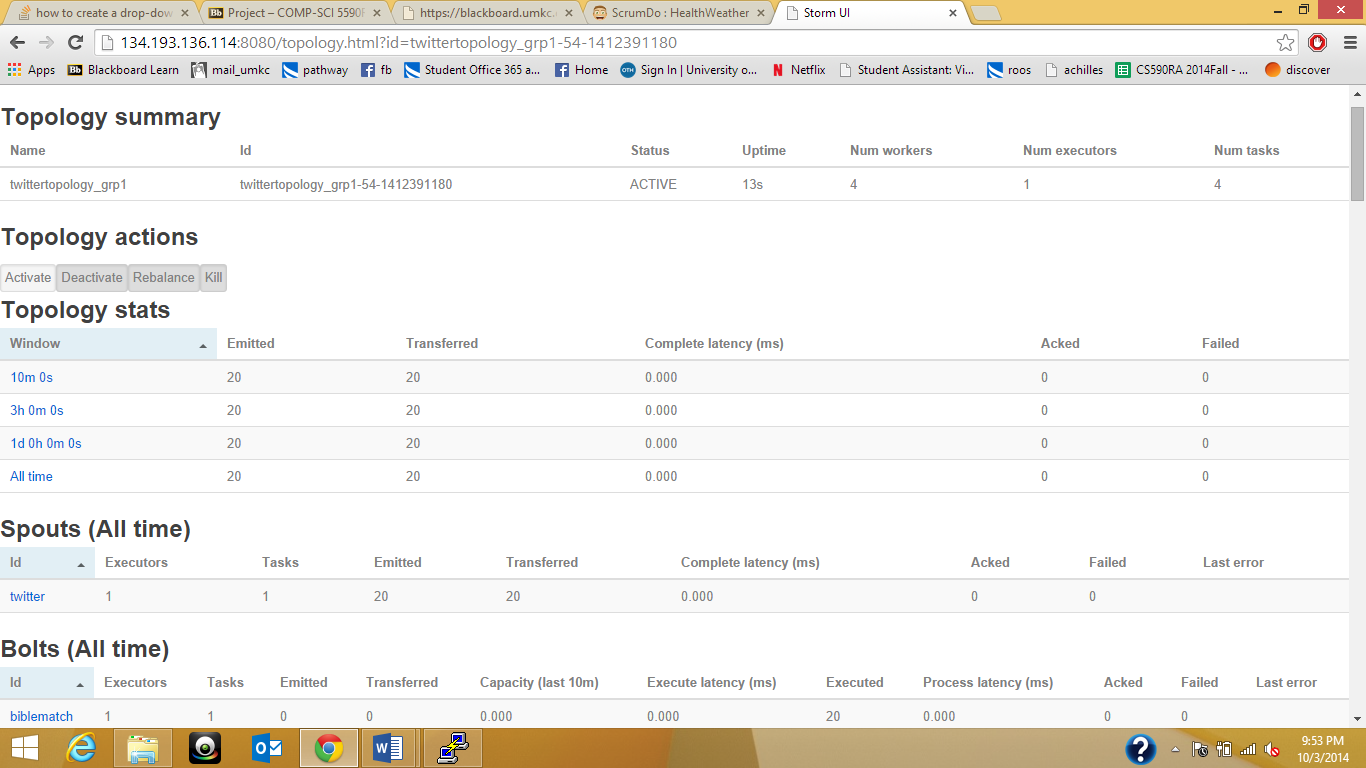


Figure 5Storm processing

The above snapshot gives the details of our running storm topology which is used to extract real time data.

**Project Management:**

In this phase we majorly focused on hbase storm integration, which is yet to be completed. But we have completed the implementation of web service in this phase. And also collecting tweets using STORM.

Tasks scheduled for second increment:

* Completion of collecting tweets and integrating storm with hbase – Jagadish, Priyanka
* Investigation on image processing or collection of data from YouTube – Anand
* Working on analysis of static data classification using R – Sai Kishore
* Integration of mobile application with web server glassfish for hbase transactions – Anand
* Completion of user and guest login pages – Jagadish , Priyanka, Sai Kishore , Anand

All the tasks and their increments are being updated in the scrumdo.

* <https://www.scrumdo.com/projects/project/healthweather/summary>

**Deployment:**

All the code and related brief data has been uploaded into the following github link:

* <https://github.com/SaiKishoreBandaru/RA_HealthWeather>