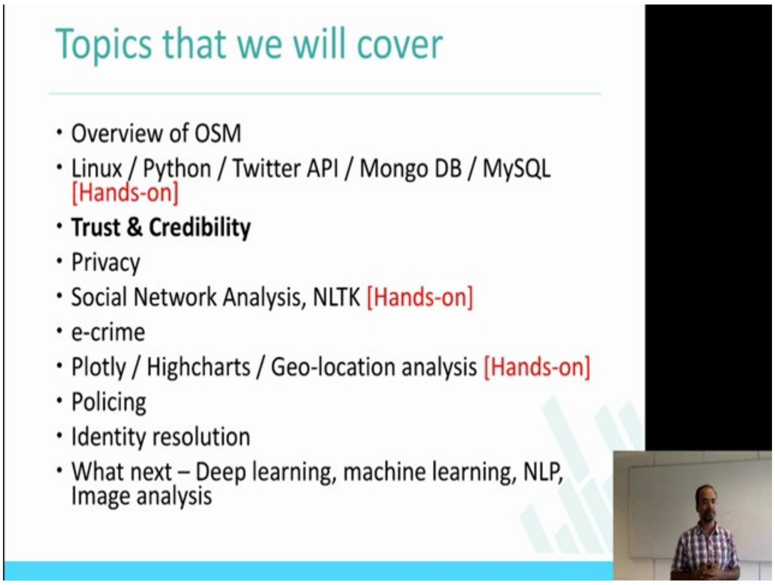
Privacy and Security in Online Social Media

Prof. Ponnurangam Kumaraguru (“PK”)
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Week – 2.2
Lecture – 09
Trust and Credibility on OSM

So now, Let us look at week 2.2 of Privacy and Security in Online Social Media course on NPTEL.

(Refer Slide Time: 00:12)

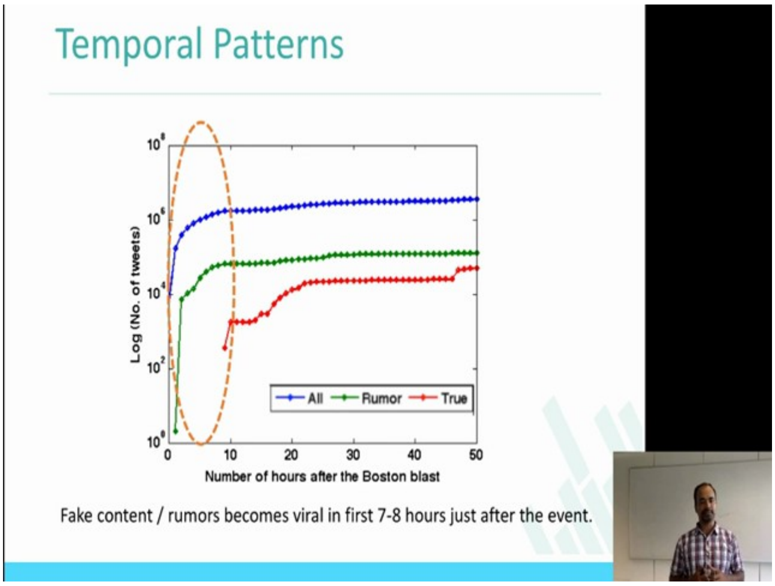


Topics that we will cover

- Overview of OSM
- Linux / Python / Twitter API / Mongo DB / MySQL [Hands-on]
- **Trust & Credibility**
- Privacy
- Social Network Analysis, NLTK [Hands-on]
- e-crime
- Plotly / Highcharts / Geo-location analysis [Hands-on]
- Policing
- Identity resolution
- What next – Deep learning, machine learning, NLP, Image analysis

In this part, what I am planning to cover is actually getting deeper into the topic called Trust and Credibility. This is the slide that I will probably tweet over the course multiple times just to tell you where we are and where we are going. We finished about overview of online social media and we in the lab sessions, we have done Linux and Python and you will actually get to see a little bit about Facebook, Twitter API’s and now we will actually look at the topic trust and credibility in detail, and later we cover some topics like privacy in social network analysis, e-crime, policing, using online social media and also identity resolution.

(Refer Slide Time: 00:53)



Temporal Patterns

Fake content / rumors becomes viral in first 7-8 hours just after the event.

Let us take a look at this graph. In this graph just to read the graph at the x-axis, this is the number of hours after the Boston blast, the data basically from the Boston blast that happened in the US. The x-axis is number of hours after the blast and the y-axis is log of tweets basically what does it show , it shows that at any given point in time which is after the Boston blast, how much of tweets is being uploaded on Twitter.

So, there are 3 different colors in this graph which is blue, green and red. So, the one which is in the red is actually legitimate information, which you can call as the true information which is posted on Twitter. Green which is the rumor which is information that is not legitimate or untrustworthy, the non-credible content that was being posted like the example, like the crocodile example that I mentioned in earlier lecture and the blue one which is the sum of the rumor on the true information.

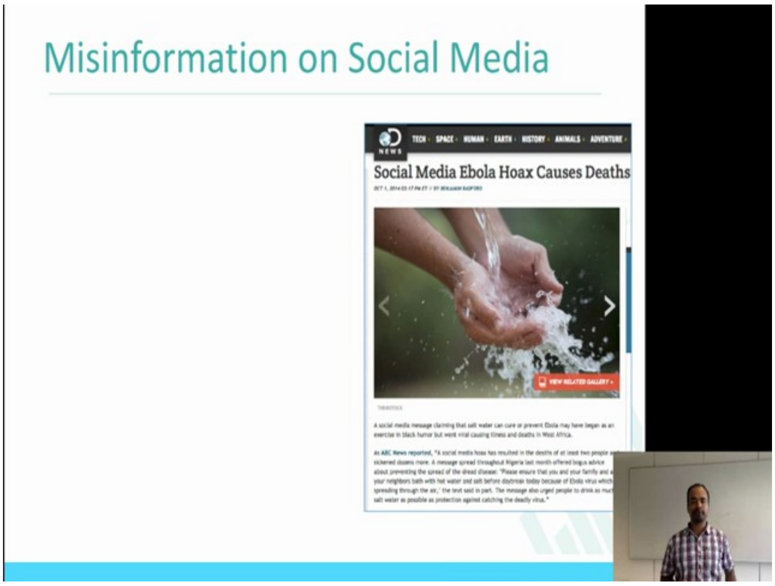
It clearly shows the messages, some implications from the step one that is the true information is actually coming later; it is taking much more time than the rumors that started. In this example, there was in this event Boston blast, there was actually multiple post which were related, which to this event was not actually legitimate, for example, one post which said that 8 year old kid was actually part of this Boston blast which when there was no kid involved in the Boston blast. There was also another tweet which said that please RT this tweet and we will actually pay 1 dollar to Boston marathon league which also was not true.

There are many examples like this and these tweets got retweeted for more than thousands of times when Boston blast happened. This actually shows that there is multiple things one can actually look at one; how do you actually reduce the time in which the true information is coming, which is from currently it is about 9 hours or so, how can you actually get this true information come on to the social network as early as possible.

The other solution that you could also think of this, how can you quickly reduce the false information that is going on social media from, to reduce, for example in this case the green one is actually peaking in couple of hours and then its actually higher than the true information, how can you actually quickly reduce the effect of or the flow or the information propagation of this particular rumor on social networks.

So, those are the two things that you could actually do, atleast do to reduce the effect of rumors you need to actually understand, what the rumor is? How can we actually identify these rumors on a Twitter that is what we basically look at in the section of this course. Which is to identify ways by which I will look at the tweets and identify whether they are legitimate or not.

(Refer Slide Time: 04:04)



Misinformation on Social Media

NEWS
TECH · SPACE · HUMAN · EARTH · HISTORY · ANIMALS · ADVENTURE ·
VIEW RELATED GALLERY »
THINKSTOCK

Social Media Ebola Hoax Causes Deaths

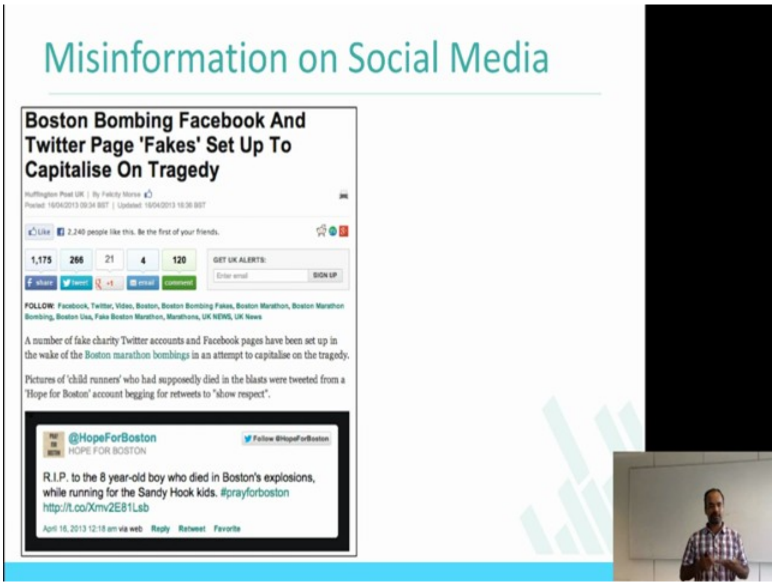
OCT 1, 2014 03:17 PM ET // BY BENJAMIN RADFORD

A social media message claiming that salt water can cure or prevent Ebola may have began as an exercise in black humor but went viral causing illness and deaths in West Africa.

As ABC News reported, "A social media hoax has resulted in the deaths of at least two people a sickened dozens more. A message spread throughout Nigeria last month offered bogus advice about preventing the spread of the dread disease: 'Please ensure that you and your family and a your neighbors bath with hot water and salt before daybreak today because of Ebola virus which spreading through the air,' the text said in part. The message also urged people to drink as muc salt water as possible as protection against catching the deadly virus."

So now, we will look at misinformation on social media, which are some examples of the misinformation that was on Twitter. Here is one example, which actually took a lot of effect in social media. When Ebola was going on there were a lot of messages saying Ebola hoax which causes deaths and there was also discussion on the post about how salt water could be used to actually reduce Ebola and things like that.

(Refer Slide Time: 04:36)



Misinformation on Social Media

Boston Bombing Facebook And Twitter Page 'Fakes' Set Up To Capitalise On Tragedy

Huffington Post UK | By Felicity Morse
Posted: 16/04/2013 09:34 BST | Updated: 16/04/2013 18:36 BST

FOLLOW: Facebook, Twitter, Video, Boston, Boston Bombing Fakes, Boston Marathon, Boston Marathon Bombing, Boston Usa, Fake Boston Marathon, Marathons, UK NEWS, UK News

A number of fake charity Twitter accounts and Facebook pages have been set up in the wake of the Boston marathon bombings in an attempt to capitalise on the tragedy.

Pictures of 'child runners' who had supposedly died in the blasts were tweeted from a 'Hope for Boston' account begging for retweets to "show respect".

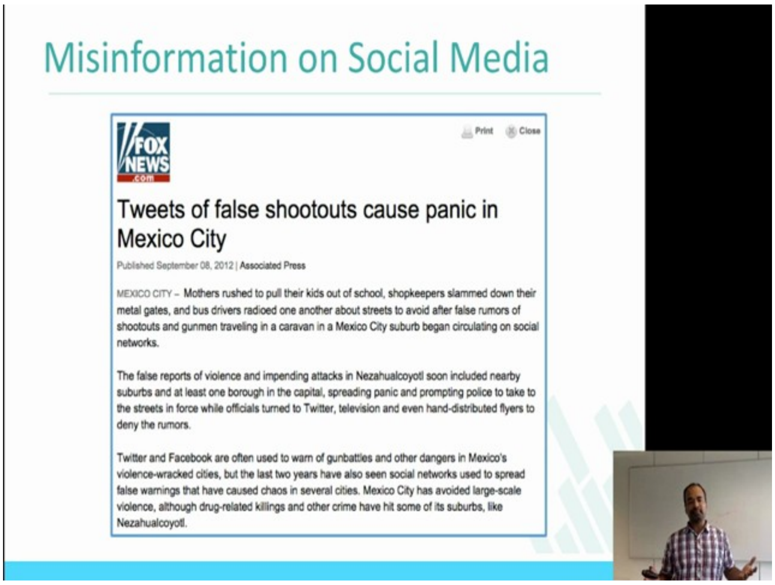
@HopeForBoston
HOPE FOR BOSTON

R.I.P. to the 8 year-old boy who died in Boston's explosions, while running for the Sandy Hook kids. #prayforboston http://t.co/Xmv2E81Lsb

April 16, 2013 12:18 am via web

Boston marathon, here is a specific tweet that I just now mentioned, which is R.I.P to the 8 year boy, who died in Boston explosions while running for the Sandy Hook kids and that was not true at all.

(Refer Slide Time: 04:50)



Misinformation on Social Media

Tweets of false shootouts cause panic in Mexico City

Published September 08, 2012 | Associated Press

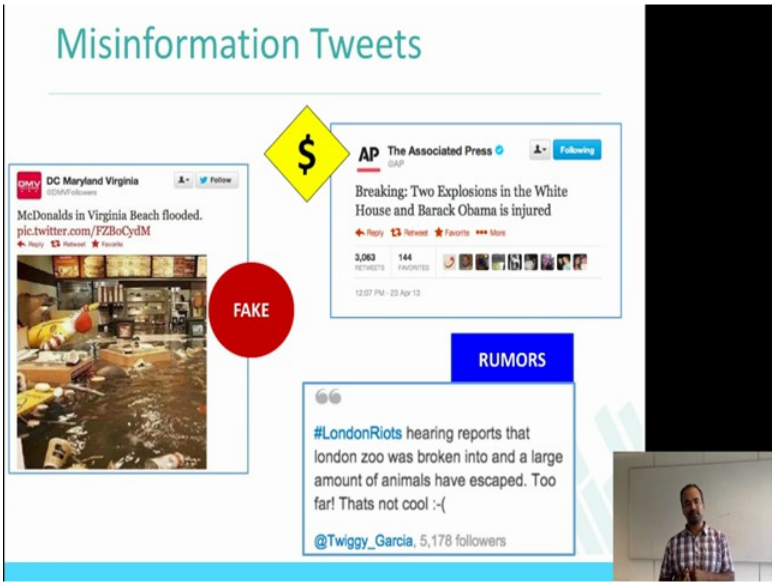
MEXICO CITY – Mothers rushed to pull their kids out of school, shopkeepers slammed down their metal gates, and bus drivers radioed one another about streets to avoid after false rumors of shootouts and gunmen traveling in a caravan in a Mexico City suburb began circulating on social networks.

The false reports of violence and impending attacks in Nezahualcoyotl soon included nearby suburbs and at least one borough in the capital, spreading panic and prompting police to take to the streets in force while officials turned to Twitter, television and even hand-distributed flyers to deny the rumors.

Twitter and Facebook are often used to warn of gunbattles and other dangers in Mexico's violence-wracked cities, but the last two years have also seen social networks used to spread false warnings that have caused chaos in several cities. Mexico City has avoided large-scale violence, although drug-related killings and other crime have hit some of its suburbs, like Nezahualcoyotl.

There's been many, many examples I am just going to give you some examples as motivation for this section of this course, tweets of false shootouts cause panic in Mexico city, this is one of the incident.

(Refer Slide Time: 05:03)



Misinformation Tweets

DC Maryland Virginia
@DMVFollowers

McDonalds in Virginia Beach flooded.
pic.twitter.com/FZBoCydM

The Associated Press
@AP

Breaking: Two Explosions in the White House and Barack Obama is injured

3,063 RETWEETS 144 FAVORITES

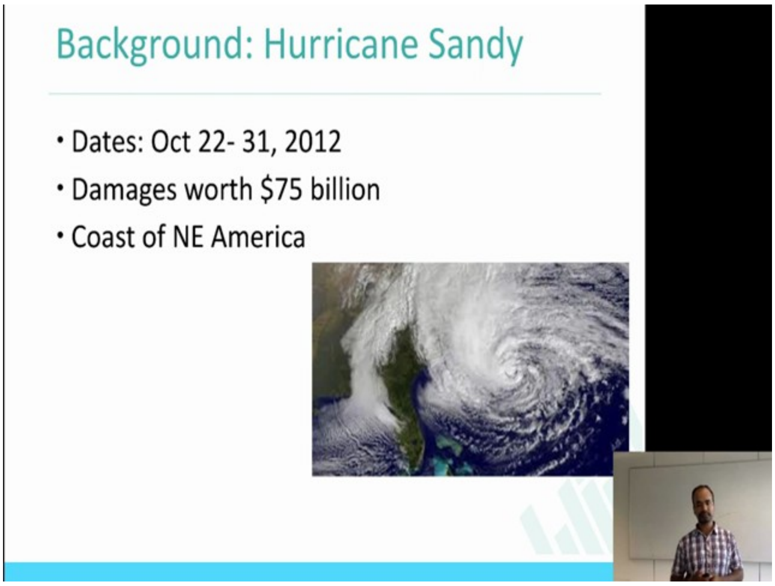
12:07 PM - 23 Apr 13

#LondonRiots hearing reports that london zoo was broken into and a large amount of animals have escaped. Too far! Thats not cool :-(

@Twiggy_Garcia, 5,178 followers

And some tweets, some images that I actually talked about even in my first lecture, which is McDonalds in Virginia Beach flooded, the image of the left where the image was actually the real image, but it was not taken during Virginia Beach flood, but it was actually taken many years before and they associated first and the past also. Here is a rumor in the right hand bottom, which is London riots, here it reports the London zoo was broken into and large amount of animals have escaped that is again a rumor. There have been many rumors like this in many events that have happened in the past.

(Refer Slide Time: 05:37)



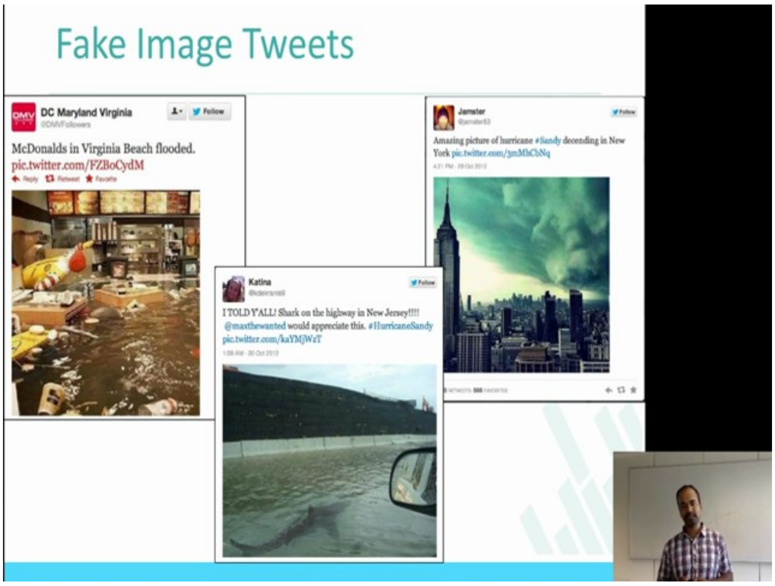
Background: Hurricane Sandy

- Dates: Oct 22- 31, 2012
- Damages worth $75 billion
- Coast of NE America

What we are going to do is we are going to actually take one specific example. We will actually do multiple examples over the course, over the entire course. Take up this event and look at the actually the topic of misinformation in this case and other topics in future to study how we can actually analyze this data and make some inferences out of it in the context of trust and credibility. We will do the similar way in the future also, for any topic we'll take an event, we will take some data that has been collected do some analysis on the content and make some inferences of the topic that we are interested in.

In this case we are going to take Hurricane Sandy. Hurricane Sandy happened in October 22-31, 2012 and the idea for using an event is that you will able to relate to it and most of the times analysis is done looking at the particular event, for example, now many people are interested in studying elections in the US and I know there are people also interested in studying elections in India when it happens. So, the damages that were totally worth for Hurricane Sandy was about 75 billion and the Hurricane Sandy basically in the north eastern part of the US.

(Refer Slide Time: 06:53)



Fake Image Tweets

DC Maryland Virginia
@DMVFollowers

McDonalds in Virginia Beach flooded.
pic.twitter.com/FZBoCydM

Katina
@kdeleranis9

I TOLD Y'ALL! Shark on the highway in New Jersey!!!! @maxthewanted would appreciate this. #HurricaneSandy pic.twitter.com/kaYMjWzT

1:08 AM - 30 Oct 2012

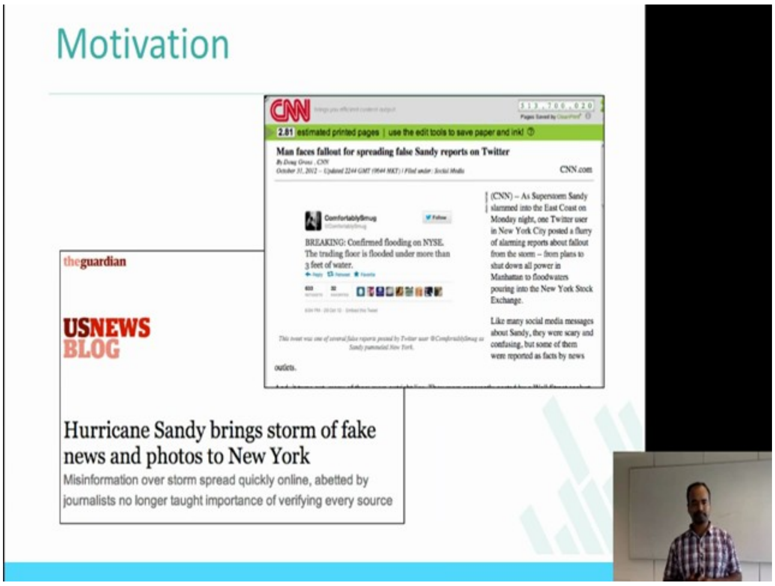
Jamster
@jamster83

Amazing picture of hurricane #Sandy decending in New Yoek pic.twitter.com/3mMhCbNq

4:21 PM - 29 Oct 2012

There has been many, many fake images that was floating during the Hurricane Sandy. The one McDonalds as I said before and one middle has shark in the water and people were actually, there was panic among a topic and the right hand topic which is also from Hurricane Sandy and which there was an image from a picture and it was actually posted on Twitter saying that is how it is looking in the US now.

(Refer Slide Time: 07:24)



Motivation

CNN

313,700,020
Pages Saved by CleanPrint

2.81 estimated printed pages | use the edit tools to save paper and ink!

Man faces fallout for spreading false Sandy reports on Twitter

By Doug Gross, CNN
October 31, 2012 -- Updated 2244 GMT (0644 HKT) | Filed under: Social Media

CNN.com

ComfortablySmug
@ComfortablySmug

BREAKING: Confirmed flooding on NYSE. The trading floor is flooded under more than 3 feet of water.

This tweet was one of several false reports posted by Twitter user @ComfortablySmug as Sandy pummeled New York.

(CNN) -- As Superstorm Sandy slammed into the East Coast on Monday night, one Twitter user in New York City posted a flurry of alarming reports about fallout from the storm -- from plans to shut down all power in Manhattan to floodwaters pouring into the New York Stock Exchange.

Like many social media messages about Sandy, they were scary and confusing, but some of them were reported as facts by news outlets.

theguardian

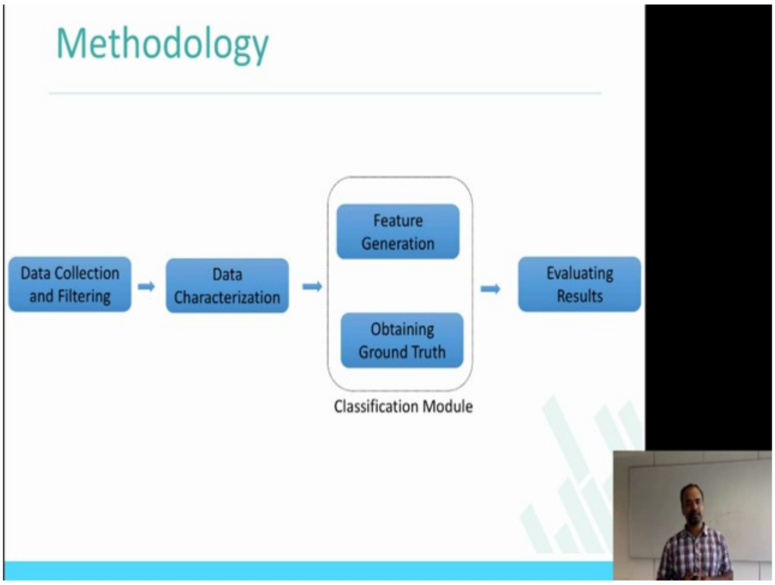
USNEWS BLOG

Hurricane Sandy brings storm of fake news and photos to New York

Misinformation over storm spread quickly online, abetted by journalists no longer taught importance of verifying every source

So, particularly for Hurricane Sandy, if you see that there is, we also know effect of these fake information that was going on Twitter.‘Hurricane Sandy brings storm of fake news and photos to New York’, ‘Man faces fallout for spreading a fake Sandy reports on Twitter’. These are some incidence which is happening around the event Hurricane Sandy.

(Refer Slide Time: 07:46)

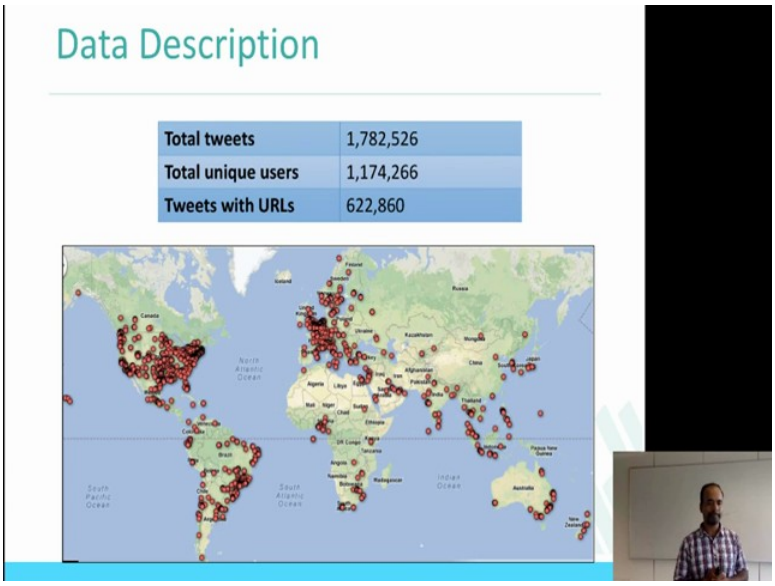


Methodology

So, what we are going to look at is again some methodology things that I will be talking about in this course is generic. I will try to actually emphasize on this methodology. So, that we would takes this methodology and apply it in any scenario that we were interested in, sometimes even in the homework and assignments that we will do as part of this course which is first we start collecting data from Twitter about the Hurricane Sandy and then some kind of data characterization which is understanding, how much of data is come? What data is come and things like? That future generation obtaining the ground truth and then evaluating the results.

This is a very high level probably 30-40000ft high level view of what the majority of the analysis on social media data would be going on. We ourselves in the course will look at different levels of view of this slide, which is later in that course we will also look at something more detail in terms of actually this the whole process. The simple process is collect some data and do some characterization, understand some features, use those features to create a model, use that model to actually study the larger amount of data and evaluate the results that is the general. If you have taken any machine learning or an information retrieval course that is the kind of a simple process that people follow.

(Refer Slide Time: 09:06)



Data Description

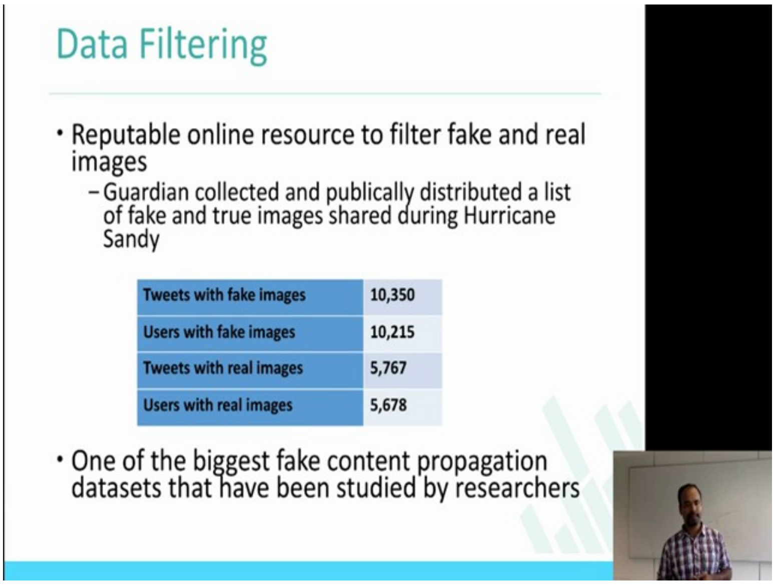
Total tweets	1,782,526
Total unique users	1,174,266
Tweets with URLs	622,860

So, the data we are talking about and this is one the most exciting thing that I feel about the studying and researching in the area of online social media, is the size of the data that we are talking about. In this case, in the Hurricane Sandy thing we are talking about 1,782,526 tweets that were collected, while Hurricane Sandy happened. Total unique users were about 1 million users and tweets with URLs was about 622,000. So, that gives you a sense of how much of data was collected in terms of the hurricane sandy.

So, again please keep this as a template when you are doing some kind of analysis of events. These kinds of attempt, these kinds of analysis, and these kinds of data description will help actually to look at the data to understand what the data is and in other sense it will also help for somebody who's going to do it again. These kinds of the same analysis or something similar they would be able to actually take away some points from the data that you would describe.

Also, in this case the map in the bottom also shows that where the tweet is come from of course, these tweets have geotagged the information therefore we are able to actually mark it on the whole map on where the tweet has come from.

(Refer Slide Time: 10:31)



Data Filtering

- Reputable online resource to filter fake and real images
  - Guardian collected and publically distributed a list of fake and true images shared during Hurricane Sandy

Tweets with fake images	10,350
Users with fake images	10,215
Tweets with real images	5,767
Users with real images	5,678

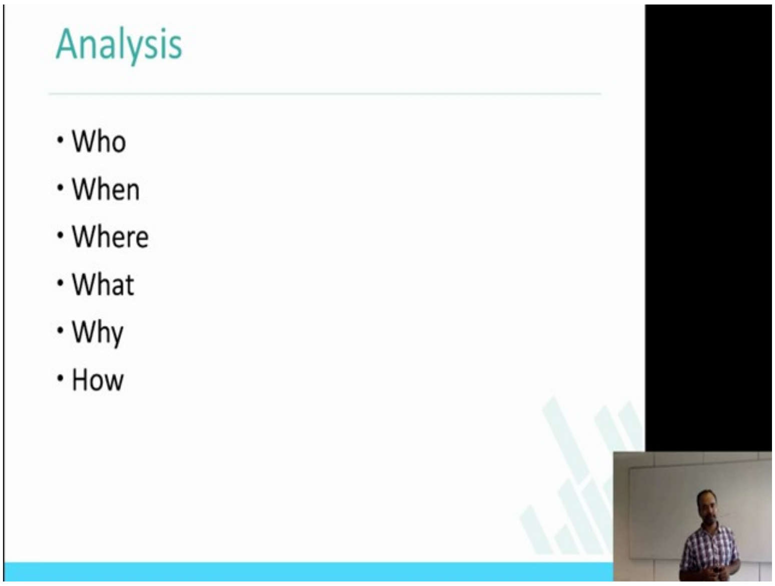
- One of the biggest fake content propagation datasets that have been studied by researchers

Of course the big question is that how will you get the ground truth because now, if we were to look at this tweets and say that which are fake, which are legitimate you need to know what the fake tweets are. So, the multiple ways that people have tried, multiple techniques that people try which is, we look at it some in the course in some probably I will just mention it as we look at the slides.

So, in this particular Hurricane Sandy analysis that was done in the way it was done was the Guardian, which is actually a media house they collected actually fake information, then manually annotated that these images are fake because they are a repository of lot of content that gets generated on social media, they were able to actually annotate and produce the dataset which is, which say that in Hurricane Sandy these are the fake post right. So, the reputable online resource to filter fake and real images, ‘Guardian collected and publicly distributed a list of fake and true images’ what did they distribute, tweets with fake images 10,350 tweets; users with fake images 10,215; tweets with real images and users with real images.

So, using the reputable guardian data, we actually looked at the data that was collected in this and then form how many tweets have these posts, how many images that was posted that were actually fake and how many unique users and things like that, that is what that was done.

(Refer Slide Time: 12:06)

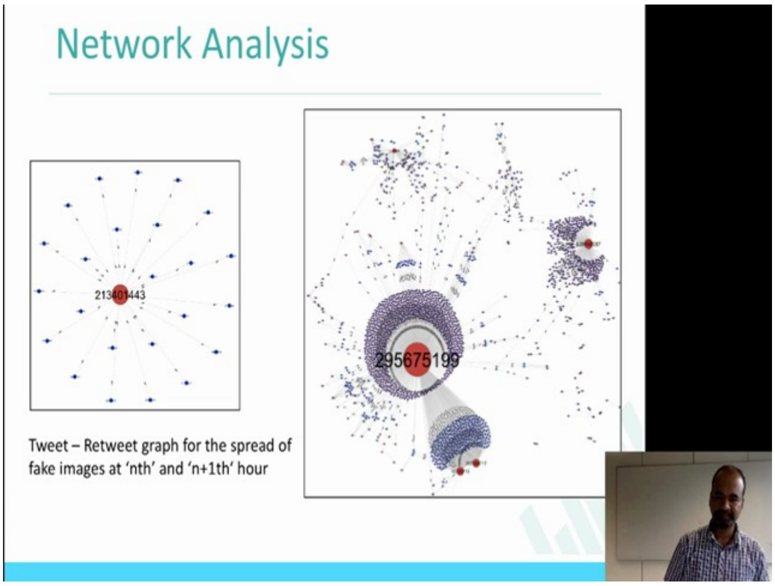


Analysis

- Who
- When
- Where
- What
- Why
- How

So, when you do these kinds of analysis when generally look at online social media analysis, it is always best to look at analysis like this; who, when, where, what, why and how. These kinds of analysis will actually help you to answer some interesting questions; who posted it? When did they post? Where did they post from? What did they post about? Why did they post and how did they post? So, why and how was slightly trickier here, it is hard to get, why did the person post a rumor? It is hard to tell unless the user, unless the person who posted and itself actually confesses, how do you, how did the course is probably is I mean probably possible to get which is to look at, what they why did they use how did they post, into the social media.

(Refer Slide Time: 12:56)



Network Analysis

Tweet – Retweet graph for the spread of fake images at 'nth' and 'n+1th' hour

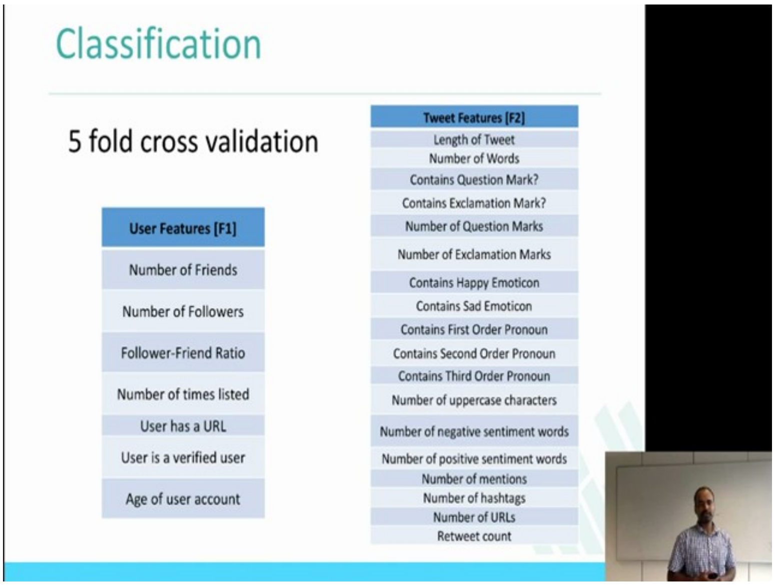
Now, let us look at more specifically the analysis on who, when, where and what and how and what. So, in this graph what I am showing you is a network analysis which is to show you, who is the person who posted the tweet and how the information is getting diffused. The one of the left is the user who posted content on this particular event and when the user actually post this content obviously, in Twitter there is going to be this retweets, favorites and mentions of the user. So, in this case the blue the red dot is the one of a particular user who posted this content and the blue dots are the ones where the users are actually retweeted this content. So, if we just look at this the users' content is actually spreading among the other users in the network.

So, the one in the left is giving you the nth hour, the one the right is giving you the n plus 1th hour what is the difference here. So, in the first one there is only one user where as content is getting spread, whereas in the one on the right if you see the post is actually diffused so heavily in the network, within one hour. But there is also other observations that you can actually have, if you look at the number in the left of the user id. So, this basically the user id is the one through which we can actually collect data from Twitter it is a unique for a every particular user. So, the number in the left and then one that numbers that are prominent in the right are actually going different.

What does this show? This shows that the content, that some body starts, let us take if PK starts content and his content gets diffused in the network, he may or may not be the one who is actually more popular at after given point in time. This basically shows that the information is, we can actually draw multiple inferences from this analysis which is, who is posting the content? How the information is getting diffused, for example, if you say 3 plus see, on the left last 3 digits are 443 which is the prominent user, whereas if you look at in the right it is 199 the user which is in the center of the network.

So, that is one important analysis that you can do inference, that you can draw from these kind of analysis, this is called network analysis. We later in the course will actually see some tools where you can actually draw these graphs with the data that you collect from Twitter or a facebook. So, in this graph you can also see in one of the user on the right hand top corner which also has more number retweets. There are some users which is bottom of the graph, on the right which is, who are also more the tweets are getting more retweeted.

(Refer Slide Time: 15:43)



Classification

5 fold cross validation

User Features [F1]
Number of Friends
Number of Followers
Follower-Friend Ratio
Number of times listed
User has a URL
User is a verified user
Age of user account

Tweet Features [F2]
Length of Tweet
Number of Words
Contains Question Mark?
Contains Exclamation Mark?
Number of Question Marks
Number of Exclamation Marks
Contains Happy Emoticon
Contains Sad Emoticon
Contains First Order Pronoun
Contains Second Order Pronoun
Contains Third Order Pronoun
Number of uppercase characters
Number of negative sentiment words
Number of positive sentiment words
Number of mentions
Number of hashtags
Number of URLs
Retweet count

So, now let us look at a different kind of analysis from the data that we collect from Twitter, one of the problems that we can actually solve this is to actually classify whether the post is post that is given to us, but that comes Twitter is actually fake or real. So, in that context, we will actually apply a technique called classification given that this is not a machine learning or an information retrival class. We will not go into detail about what the classification is? What are different techniques? I will just only look at techniques that we applied with the data that we collect from Twitter.

So, if you look at this the different kinds of features that we actually get from the post that we get from Twitter or user features and tweet features. There are actually three kinds of content that you can actually look at from Twitter, one is the user profile which is, who am I in my case the faculty at IIIT, Delhi got my PhD from Carnegie Mellon university and things like that. Second, the people who I am connected with, that is my network my network would be faculty that are around the world, students that are around the world who are doing cyber security, people who are, people who studied on social media and things like that is my network.

Third is actually the content that I post itself what I am talking about I am talking about my students I am talking about PSOSM course, I am talking about some random things on social media right. So, those of the three broad categories of content that you can actually draw from the social media data which is user profile, the content that somebody post and the third one being the network that somebody is connected to.

So, in this case we are actually looking at the user features which is more like feel, which is more like the profile tweet features. Tweet features are from the tweet itself; let me just go through few of them that I have listed in this slide, which is in terms of user features number of friends. Somebody has number of followers, follower-friend ratio is also one of the important things that we can actually use while making the decision on whether this user is actually legitimate or fake, for example, if somebody is very popular user, the number of followings that they would have is actually much lower that is the people that PK will follow is actually lower than the number of people who would actually follow PK.

So, that ratio can be actually used to make a judgment on whether the network, whether the user is actually legitimate or not, number of times listed list is another feature in Twitter where let us take, if I want to create a list of all the students who are taking

PSOSM course I will create a list as PSOSM course on NPTEL and I will actually add all the Twitter users on to this list.

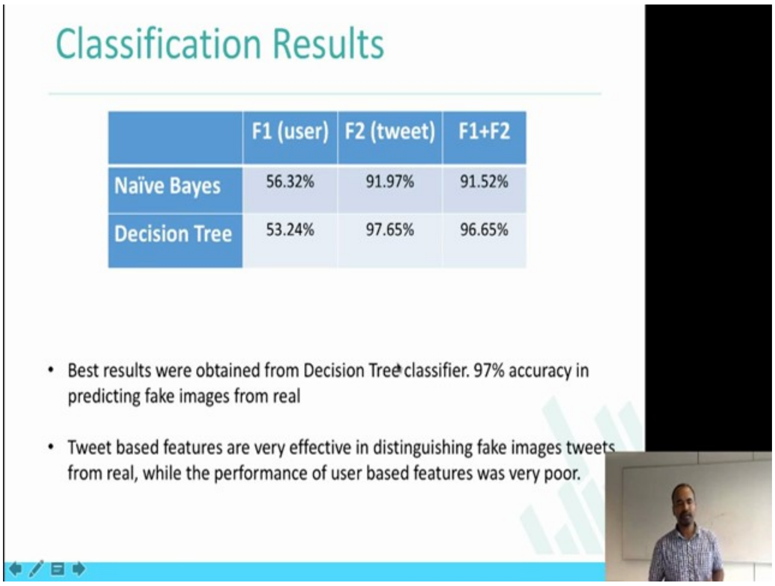
So, that is called a list and particularly this list has been used in different interesting ways if I were to find the experts around the world on particular topic list could be actually one of the good ways to find out which is if I were to create a list or somebody else creates a list on cyber security, and if they add PK onto it there is high probability that the person believes the PK is actually an expert and therefore, he is adding him or her into that list.

So, number of times listed is one feature that you can use, please go through the Twitter network, do play around with the list and other features that I am taking about, user has the url which is in my profile I will actually say that I am faculty at IIIT, Delhi and this URL called precog dot iiitd dot edu dot in, that URL is there how could the you can actually use that feature to predict user is a verified himself, verified user is another important feature that you could use because of their total number of users on Twitter there is only few a hundreds and thousands of users, who are actually verified, verified takes some process and you have to be you are to have a larger followers and things like that.

So, verified user can be a good feature to decide whether the user is legitimate or a malicious user. Age of the user account, and this has been a feature that people have used in traditional internet security methodologies, where they have actually used age of a web site, age of the domain registration to actually find out whether the domain is legitimate or not. It is same feature which is PK created an account 5 years back it is more legitimate and there is a PK account which means or there is Amitabh Bachchan account or Rajinikanth account which is created recently which may not be actually legitimate account. So, that is the intuition behind using age of a user account. So, next is tweet features let us just look at little bit about tweet features itself, in tweet features length of the tweet is a good information that you can actually use to find out whether the post is legitimate or not.

So, these are the features that you can use in general from many different analysis I am only using it for the problem of trust and credibility that we are talking about. Also length of the tweet number of words in the tweet, contains question mark, contains exclamation marks, number of question marks, number of exclamation marks, contains happy emoticon, contains sad emoticon, and things like that. So, this these are the different features that you can actually draw from tweets and the features that I told earlier which were actually user features. So, five fold cross validation is a technique which is used to make sure that the confidence on the classification accuracy that we are building is higher that is the reason why we use actually five fold cross validation and there are many other techniques, I am not going to into details of different other techniques that are available.

(Refer Slide Time: 21:22)



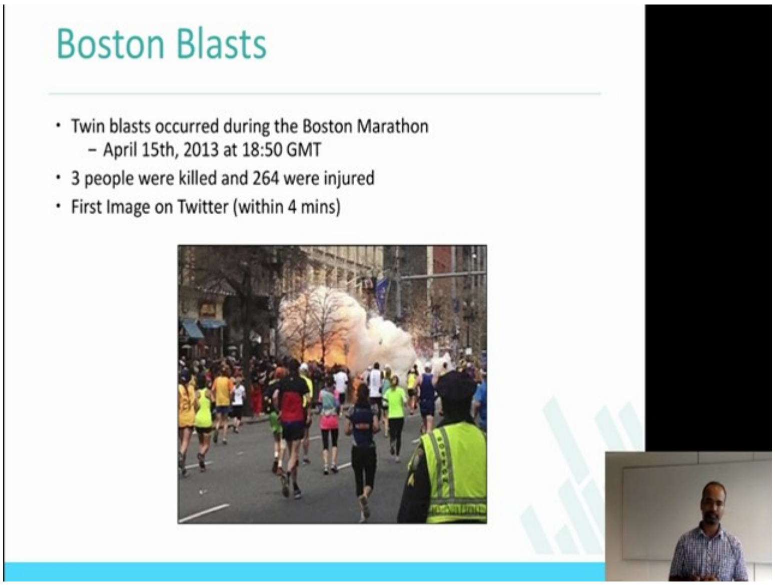
Classification Results

	F1 (user)	F2 (tweet)	F1+F2
Naïve Bayes	56.32%	91.97%	91.52%
Decision Tree	53.24%	97.65%	96.65%

- Best results were obtained from Decision Tree classifier. 97% accuracy in predicting fake images from real
- Tweet based features are very effective in distinguishing fake images tweets from real, while the performance of user based features was very poor.

Now, let us look at the results from the classification that we did. So, as I said F1 is a user feature F2 is a tweet features, and in classification techniques we can actually use F1 F2 separately and also create a features of F1 plus F2 the two techniques that were applied one is Naive bayes, which I actually uses bayes theorem to find out whether particular post is legitimate or fake and which features actually influence a lot in making the decision. That is also another technique which is a graph based technique which is decision tree, where all the outcomes and the probabilities are actually layed down on as in the form of a graph and it is a very popular machine learning technique which is applied to make decisions. And this particular case decision tree actually seem to a work better the people while using the tweet features, while the efficiency was about 97.65 percent, in predicting whether the post is fake or real. The tweet features I have choosed seems to a you have played well in both a naive bayes and decision tree where as use the features did not that play that much well in making the decision.

(Refer Slide Time: 22:32)

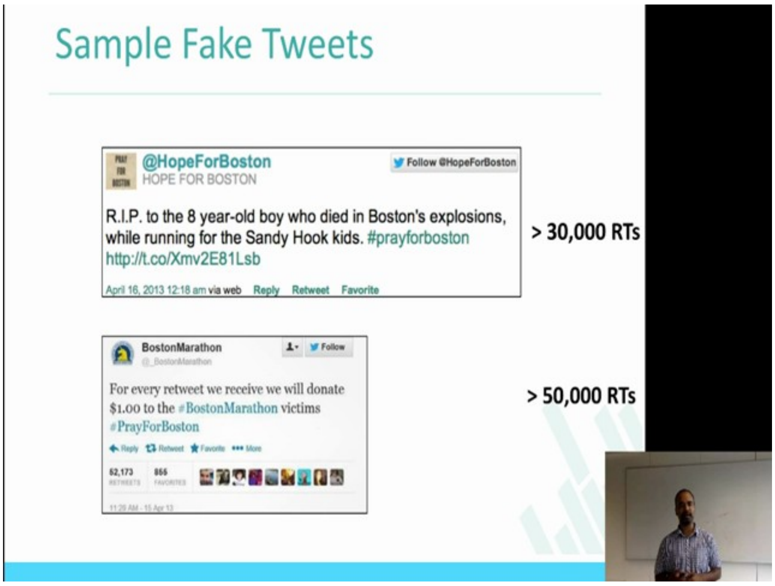


Boston Blasts

- Twin blasts occurred during the Boston Marathon
  - April 15th, 2013 at 18:50 GMT
- 3 people were killed and 264 were injured
- First Image on Twitter (within 4 mins)

Now, let us look at an event Boston blast again, we will use this technique could be looking at events through the events I will actually inject a lot of techniques that we will actually study in this class, and terminologies also that we will see. It is a twin blast that happened in 2013 in April and 3 people were killed and 264 were injured, first image that come on Twitter was within 4 minutes. It is basically a Boston marathon that was going on and the blast of the finish line was the event that happened.

(Refer Slide Time: 23:01)



Sample Fake Tweets

@HopeForBoston
HOPE FOR BOSTON

R.I.P. to the 8 year-old boy who died in Boston's explosions, while running for the Sandy Hook kids. #prayforboston http://t.co/Xmv2E81Lsb

April 16, 2013 12:18 am via web

> 30,000 RTs

BostonMarathon
@_BostonMarathon

For every retweet we receive we will donate $1.00 to the #BostonMarathon victims #PrayForBoston

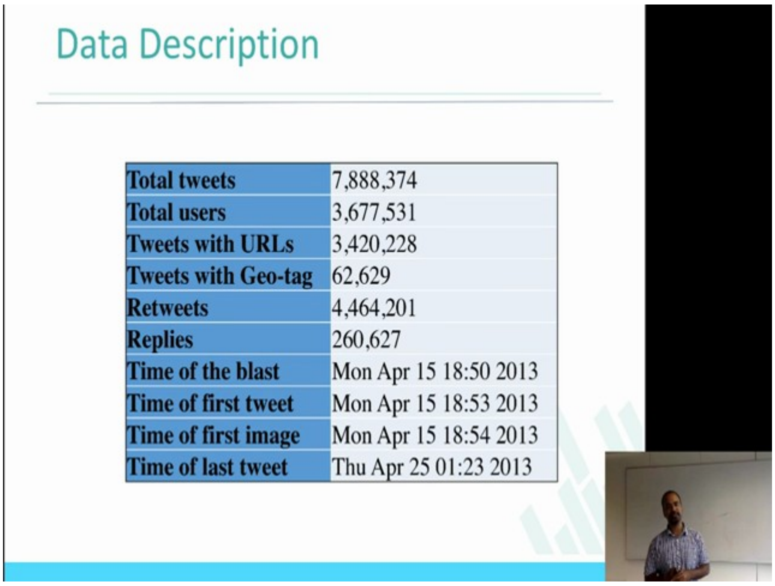
62,173 RETWEETS 855 FAVORITES

11:29 AM - 15 Apr 13

> 50,000 RTs

And there were actually multiple fake tweets here I am just showing you two popular fake tweets that were actually floating around, the first one I have showed this tweet in the past also 'R.I.P. to the 8 year old boy who died in Boston's explosions, while running for the Sandy Hook kids'. There was no kid who has participated in the marathon and then other post the at the bottom you see, for every retweet we will donate one dollar to the Boston marathon victims, and it is posted by an account called underscore Boston marathon something that you want to keep in mind which was not a legitimate account and this post was retweeted for about 50,000 times and these are the two popular tweets that were floating around during the event which were fake

(Refer Slide Time: 23:47)



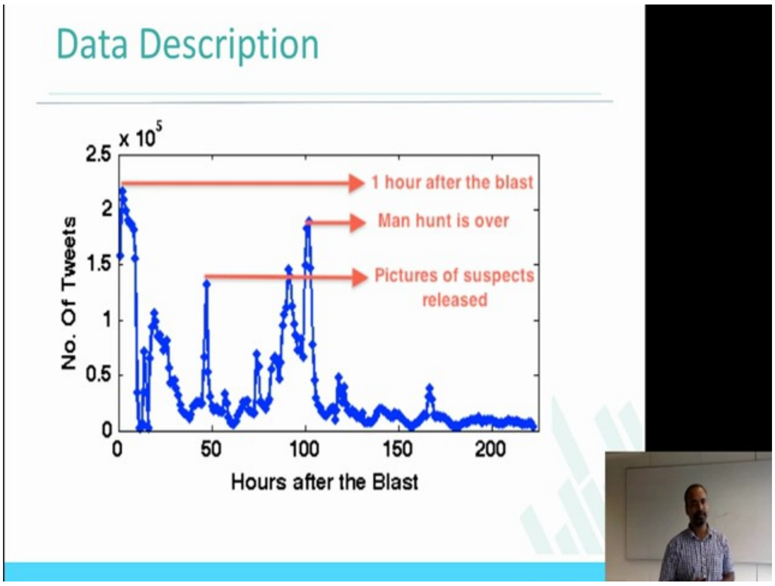
Data Description

Total tweets	7,888,374
Total users	3,677,531
Tweets with URLs	3,420,228
Tweets with Geo-tag	62,629
Retweets	4,464,201
Replies	260,627
Time of the blast	Mon Apr 15 18:50 2013
Time of first tweet	Mon Apr 15 18:53 2013
Time of first image	Mon Apr 15 18:54 2013
Time of last tweet	Thu Apr 25 01:23 2013

Data that was collected during this events was actually about 7.8 million tweets that were collected, 3.6 million users posted this tweet and if you look at the advantage of actually working in this space of online social media is actually this large numbers that we look at, tweets with URLs is about 3.4 million, 62,000 people are posts had geo tag and about 1 percent is what Twitter claims that the tweets that are posted on Twitter are geo tagged tweets, about 4 point 4 million replies 260,000 in the timeline of the blast. First tweet, first image, and last tweet, all of that is capture in this slide I will show it you because when you actually present and an analyze events. analyze a particular topic and I think studying this analyzing unit is only one way, but we are actually, you can actually adapt this to studying any topic.

For example in this case how do we collect the data we take hashtag Boston marathon which is actually trending and start collecting tweets, which has hashtag Boston marathon, look at other words that are in the post that has hashtag Boston marathon and use those key words to start collecting other tweets like query expansion concept and thereby we collect that post from Twitter. And this methodology can be used for collecting any data, data could be hashtag macbook pro, hashtag apple hashtag india and things like that. Not necessarily it has to be hashtag also, it could be any other words.

(Refer Slide Time: 25:19)

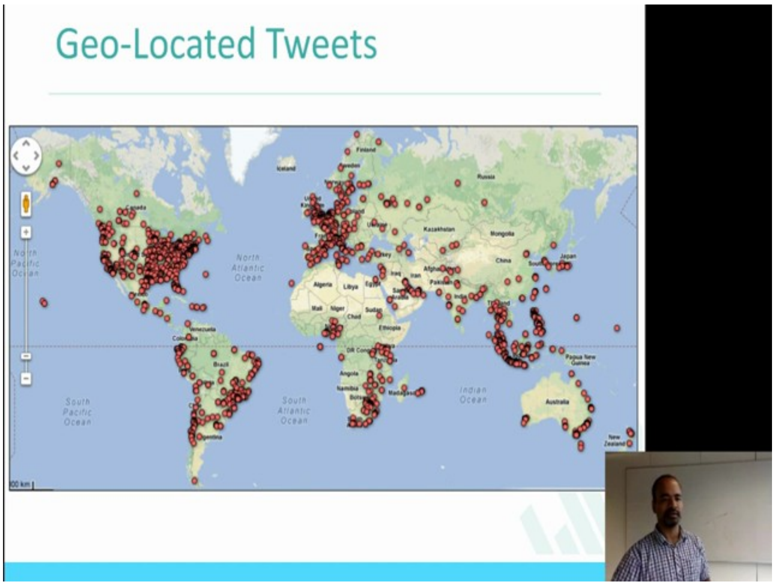


Data Description

Here I am actually showing you a graph which is on the x-axis is the hours after the blast in the y axis is number of tweets. So, the the crux of this slide is to show you that the spikes that happen on social media is actually very very connected or correlated to the events that happened in real world, for example, here are the first spike that happened is one hour after the blast then there was a spike in Twitter tweets which is pictures of suspects released and man hunt is over.

So, if you really look at it, that is the way that the content gets generated on social media is actually behaving and we've tried looking at these kind of blast for many, many events and it looks very similar, in terms of actually spiking correlated to the real world incident or an event that happens. And here is another slide which actually shows you the geo tagged coordinates of the tweets that were posted on Twitter.

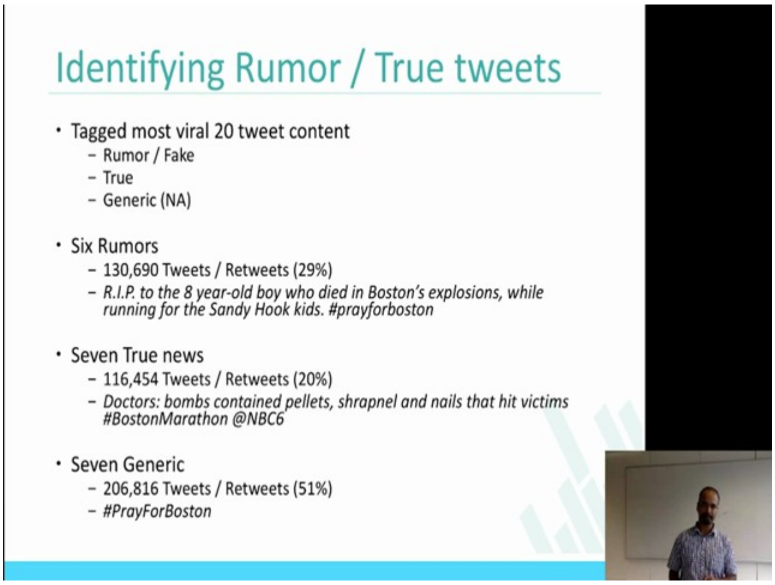
(Refer Slide Time: 26:16)



Geo-Located Tweets

Particularly for this Boston marathon blast, and every dot in this slide show you the tweets that have come from that particular location, understandably the posts have come mostly from the US.

(Refer Slide Time: 26:39)



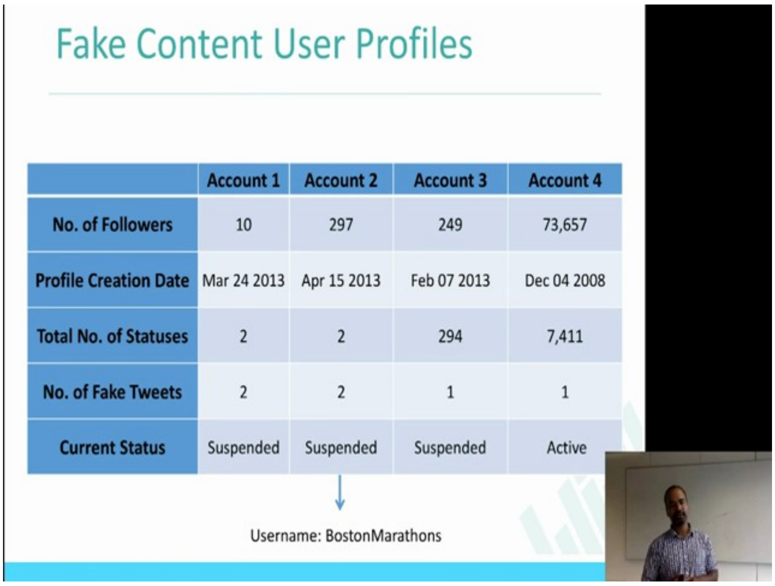
Identifying Rumor / True tweets

- Tagged most viral 20 tweet content
  - Rumor / Fake
  - True
  - Generic (NA)
- Six Rumors
  - 130,690 Tweets / Retweets (29%)
  - *R.I.P. to the 8 year-old boy who died in Boston's explosions, while running for the Sandy Hook kids. #prayforboston*
- Seven True news
  - 116,454 Tweets / Retweets (20%)
  - *Doctors: bombs contained pellets, shrapnel and nails that hit victims #BostonMarathon @NBC6*
- Seven Generic
  - 206,816 Tweets / Retweets (51%)
  - *#PrayForBoston*

So, now in this event on Boston marathon the technique that was applied for finding out whether a post is actually rumor, true or fake, first tagged most viral 20 tweet content which is whether it is rumor fake true or generic rumors, 6 rumors were actually collected from the posts that were talking about Boston marathon and seven true news was collected, which is doctored bomb contained pellets, shrapnel and nails that hit victims Boston marathon hashtag NBC6. So, those kinds of tweets were collected which is about true news that that was getting posted during the Boston marathon and six rumors were collected and seven generic posts that had pray for that the Boston. Pray for Boston also was actually trending during that time.

So, essentially what we were trying to do is we were trying to study, look at the rumors that were posted, true news that were coming and some generic information, generic post that happened during in the Boston marathon event. In this kind of you see a generic sense of what are the different post that happened in an event like this and rumors about 29 percent were actually retweeted, true news about 20 percent were retweeted and 51 percent of the generic content was actually fit for retweeted.

(Refer Slide Time: 28:06)



Fake Content User Profiles

	Account 1	Account 2	Account 3	Account 4
No. of Followers	10	297	249	73,657
Profile Creation Date	Mar 24 2013	Apr 15 2013	Feb 07 2013	Dec 04 2008
Total No. of Statuses	2	2	294	7,411
No. of Fake Tweets	2	2	1	1
Current Status	Suspended	Suspended	Suspended	Active

Username: BostonMarathons

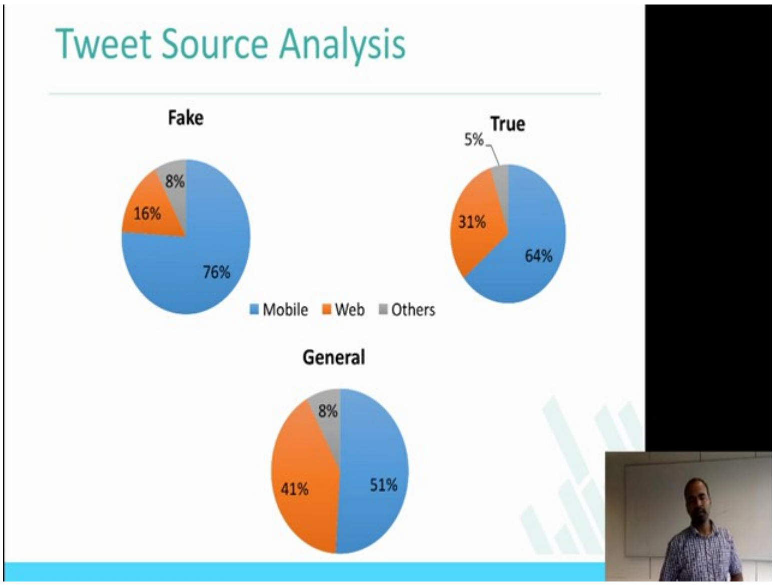
Here is another view of the data which is to show you the fake content user profiles. So, every time such event happens incidentally what happens is many of the fake user profiles get generated during the event, fake accounts gets generated to actually use the event to propagate malicious content.

So, in this example I am showing you account 1, account 2, account 3, account 4 different accounts were actually created, you will actually see that the accounts has number of followers which is account 4 being very high, account 1 being very low. So, this is and if you see the account 1 which is actually created on March 24, 2013 which was posting fake content, account 2 which was created on October 15, 2013 very close to the Boston marathon and account 3 February 2013 and account 4 2008. Total number of status this is the updates that they created and number of fake tweets that they posted is 2, 2, 1 and 1 respectively.

And if you see some of the accounts where get actually suspended and these suspensions happen because people report about this handle to Twitter in a multiple ways to actually keep down a particular account, while one is large number of people actually report a particular handle to Twitter and it gets suspended and there are through government processes you can actually apply for suspending an account.

Some if you see the last column, it is interesting that some of the user handles which are posting fake content on events like Boston marathon actually are active, even when we were actually collecting the data. This shows you that fake content propagated by fake user handles and these user handles created just after the event or just before the event, just after the event happens.

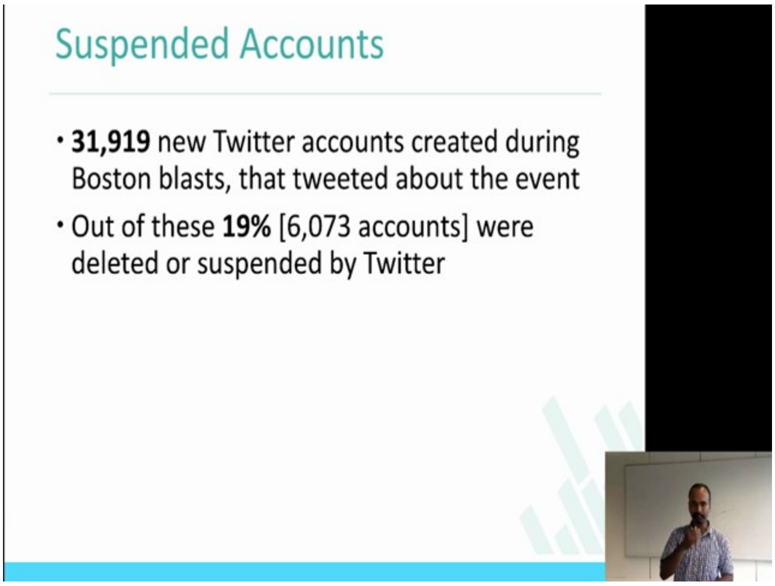
(Refer Slide Time: 30:04)



Tweet Source Analysis

Now, let us look at different view of the analysis which is tweet source analysis this gives you an insight about what devices do people use while posting the content, 76 percent of the post that was identified as fake was posted through mobile, whereas the 64 in true and 51 in general content that were posted. This insight about what device is being used, while posting this content can be very useful in making decisions, for example, if you wanted to targeted advertisement, what kind of devices are being used can be very useful in making the decision. So, the device this information is available in the JSON that you collect from Twitter for every tweet. So, you can use that to make this judgement.

(Refer Slide Time: 30:47)

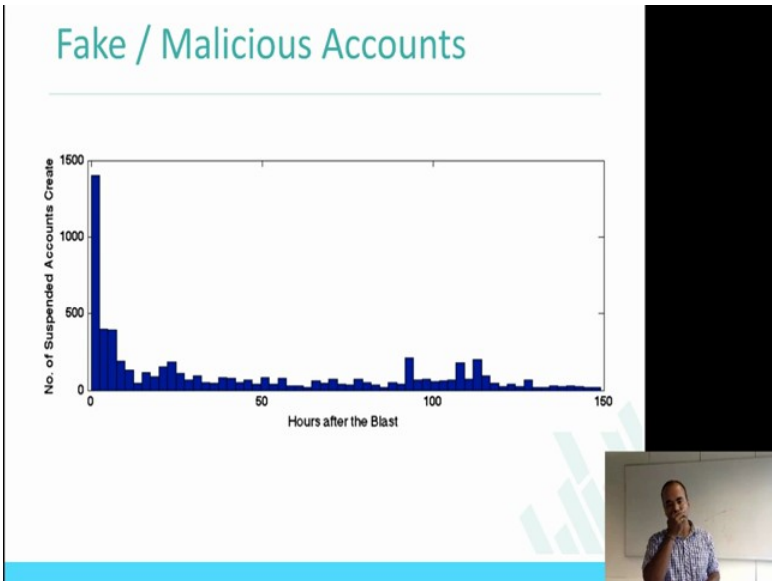


Suspended Accounts

- **31,919** new Twitter accounts created during Boston blasts, that tweeted about the event
- Out of these **19%** [6,073 accounts] were deleted or suspended by Twitter

So, if you look at the number of accounts that were created during this event, it was about 32,000 new Twitter accounts were created during this event, which were actually talking about this particular event. Out of this 19 percent were deleted or suspended by Twitter which again could have happen for multiple reasons and 19 percent of accounts that were created were actually suspended.

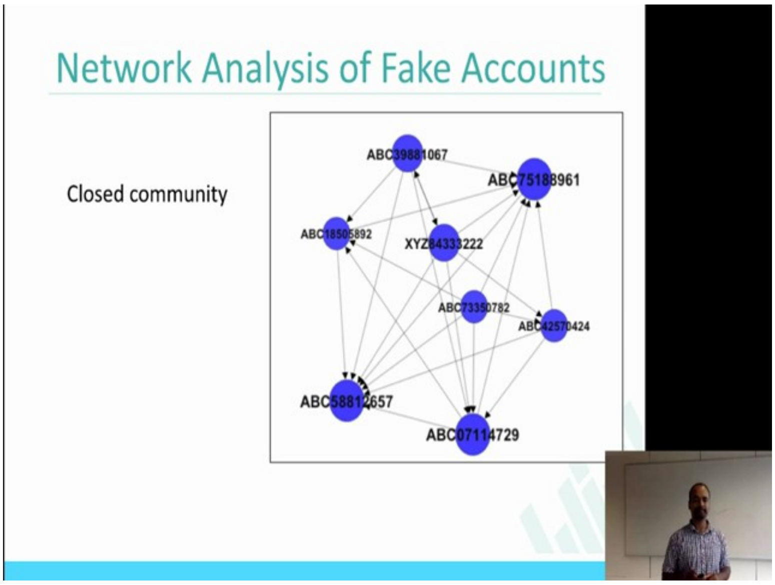
(Refer Slide Time: 31:12)



Fake / Malicious Accounts

So, this is graph to show you hours after the blast; x-axis being the hours after the blast and y-axis being number of suspended accounts that were created, which is to just to show you that the number of accounts that gets created immediately after the account is also high, in addition to that number of accounts getting deleted also high the fake or malicious accounts that were suspended, that were created and suspended were very high immediately after the event and after the event it kind of reduces little bit.

(Refer Slide Time: 31:45)



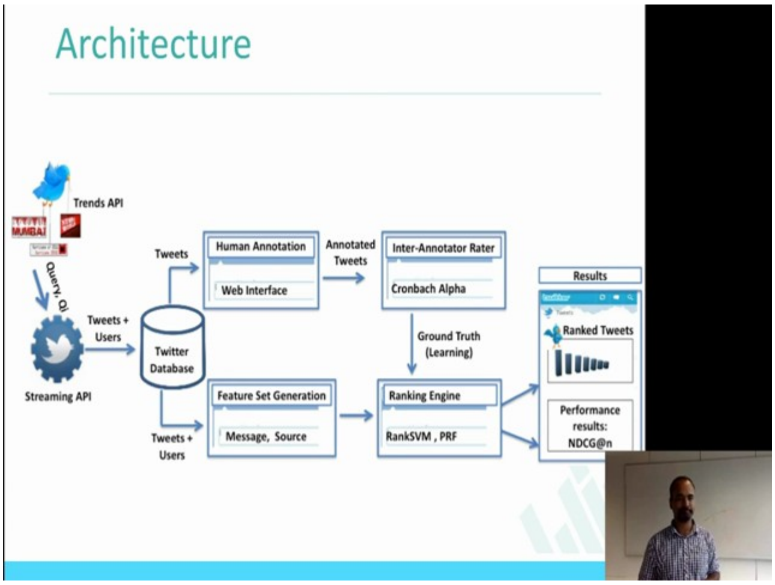
Network Analysis of Fake Accounts

Closed community

Let us look at the connection between the fake accounts itself. So, again just keep in mind the kinds of analysis that were doing is who, when, where and what, why and how. So, again that one insight into the analysis is this how were the people who were posting this content which is fake are connected. So, one insight is that they are actually pretty closed and this is not only in this domain you can actually seen this kind of analysis in many other domains also, for example, in classical security problem like phishing.

The number of groups, number of accounts, number of sets of people who do this actually by it is small and they are all very well connected, similar kind of inferences is derived from this particular analysis also where for the Boston marathon if you look at people, the node here is the user and the edge between the nodes are the action of retweet following and followers. So, therefore, there is a there is a closed community that is actually operating in terms of posting this fake content.

(Refer Slide Time: 32:51)



Architecture

Trends API

Query, Qi

Streaming API

Tweets + Users

Twitter Database

Tweets

Human Annotation

Web Interface

Annotated Tweets

Inter-Annotator Rater

Cronbach Alpha

Ground Truth (Learning)

Tweets + Users

Feature Set Generation

Message, Source

Ranking Engine

RankSVM , PRF

Results

Ranked Tweets

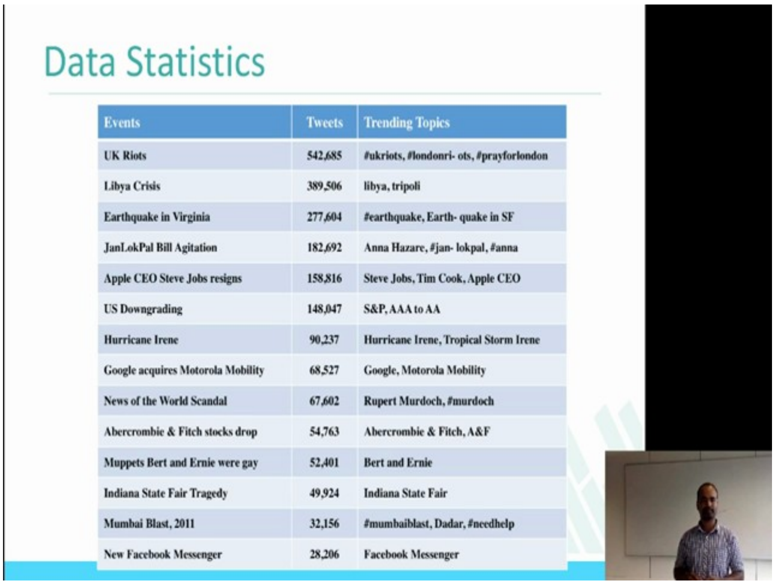
Performance results: NDCG@n

So, while doing this you could actually think about. So, earlier I think in week one I actually showed you some very high level slide about how machine learning and about how these kind of approaches of identifying fake and legitimate can be done, this is just to slightly zoomed in view of the same slide, which is to show you that the data is coming from Twitter through streaming API's. Assume you know, everybody knows now what is streaming API is, which is to collect data from the social networks, tweets are dumped to the database, human annotations which we saw earlier also in terms of annotating the post even now we saw about fake, generic and true our post those are all sometimes human annotated, sometimes you could actually use some simple techniques to do the annotation, one of the important thing that you also want to do in this annotations are done are inter annotator agreement which is if I say something that is legitimate and if you say something that is legitimate then probably more people saying post is legitimate, then the post must be legitimate, that is the kind of intuition that the Cronbach's Alpha, which is value that you may get for finding out inter annotator agreement and at the Cronbach's Alpha is generally about 0.7.

It is actually understood that the data has you can have more confidence in the inferences that you are drawing from this particular data. Cronbach's Alpha is the value that you will calculate while finding out inter annotator agreement and so as we discussed earlier also feature extraction, feature extraction is a technique by which we took F1, F2 and those things you will use that to find the model here.

I will just describe a little bit in the later slides about what model can be generated and you use that to find out whether this particular post is legitimate or not and then you can actually show that to the user also, that is the architecture that is presented here again is a very simple machine learning approach which is take the posts, use the post, do some feature extraction, use the feature extraction to create a model, use the model to actually predict whether the post is legitimate or not.

(Refer Slide Time: 35:14)



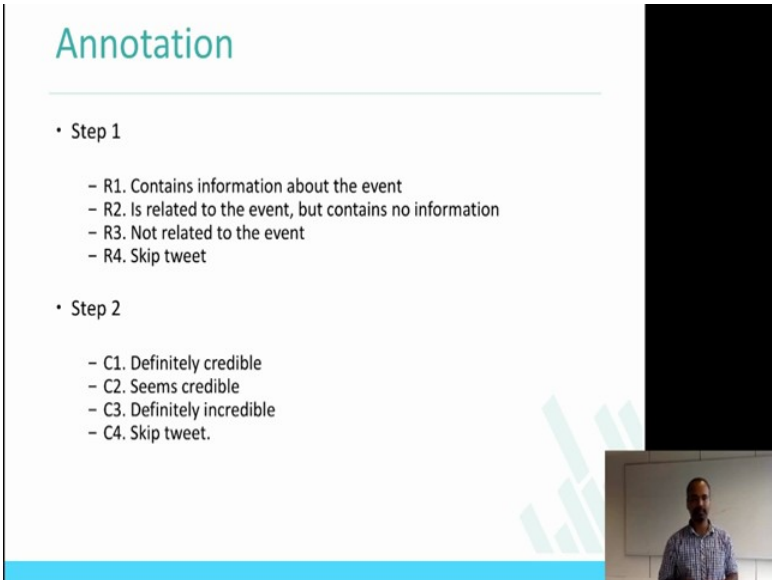
Data Statistics

Events	Tweets	Trending Topics
UK Riots	542,685	#ukriots, #londonri- ots, #prayforlondon
Libya Crisis	389,506	libya, tripoli
Earthquake in Virginia	277,604	#earthquake, Earth- quake in SF
JanLokPal Bill Agitation	182,692	Anna Hazare, #jan- lokpal, #anna
Apple CEO Steve Jobs resigns	158,816	Steve Jobs, Tim Cook, Apple CEO
US Downgrading	148,047	S&P, AAA to AA
Hurricane Irene	90,237	Hurricane Irene, Tropical Storm Irene
Google acquires Motorola Mobility	68,527	Google, Motorola Mobility
News of the World Scandal	67,602	Rupert Murdoch, #murdoch
Abercrombie & Fitch stocks drop	54,763	Abercrombie & Fitch, A&F
Muppets Bert and Ernie were gay	52,401	Bert and Ernie
Indiana State Fair Tragedy	49,924	Indiana State Fair
Mumbai Blast, 2011	32,156	#mumbaiblast, Dadar, #needhelp
New Facebook Messenger	28,206	Facebook Messenger

So, in using this architecture just taking, instead of just doing one or two events, multiple events data were collected and used to find out whether a particular technique, technology can be identified where this post is legitimate or not and here are the events UK riots, Libya crisis, earthquake in Virginia and US downgrading there are many, many events data were collected and in as I said before the column in the third column here which is trending topics.

These were the topics that were trending using which the data was collected; the column 2 shows you the number of tweets, again a large number of data was used in while doing this analysis.

(Refer Slide Time: 35:58)



Annotation

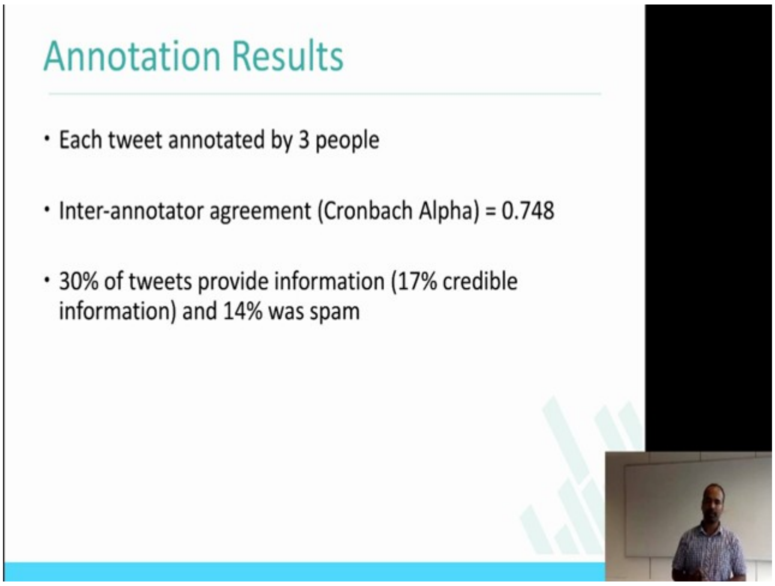
- Step 1
  - R1. Contains information about the event
  - R2. Is related to the event, but contains no information
  - R3. Not related to the event
  - R4. Skip tweet
- Step 2
  - C1. Definitely credible
  - C2. Seems credible
  - C3. Definitely incredible
  - C4. Skip tweet.

As discussed before, one of the methods used to find out whether particular post is legitimate or not, annotation was done there are multiple ways to do this annotation. Also you could get 3 of your friends to sit down together and I will tell you whether this every post is legitimate or fake, you can do through mechanical turk, mechanical turk is a crowd source mechanism by which you could actually show these posts in the platform of mechanical turk. The turkers also to look at it, turkers are basically people all around the world who are doing this task for a small money, step one in this case contains information about the tweet, post is shown to the user and in the user actually decides on one of these four characteristics which is contains information,is related to the event not only related to the events, skip.

If in the step one and says that contains the user decides that there is a information in this post, then the user is asked about definitely credible, seems credible, definitely incredible and skip tweet, again here, I am only going through the methodology which is post is taken. It is annotated you could annotate it for any particular topic that you would where you want to study, in this case it is credibility, but you could also think about it whether this post has phising URL or not, if this post is talking about a particularly event or not, this post is sensitive or not you could do many, many things in terms of annotations and in the topics that you are interested in studying from the post that is being collected.

So, from step one you take the data and then you ask the users to classify, it as definitely credible, seems credible, definitely credible and if there is nothing the user cannot make a decision, skip the tweet.

(Refer Slide Time: 37:41)

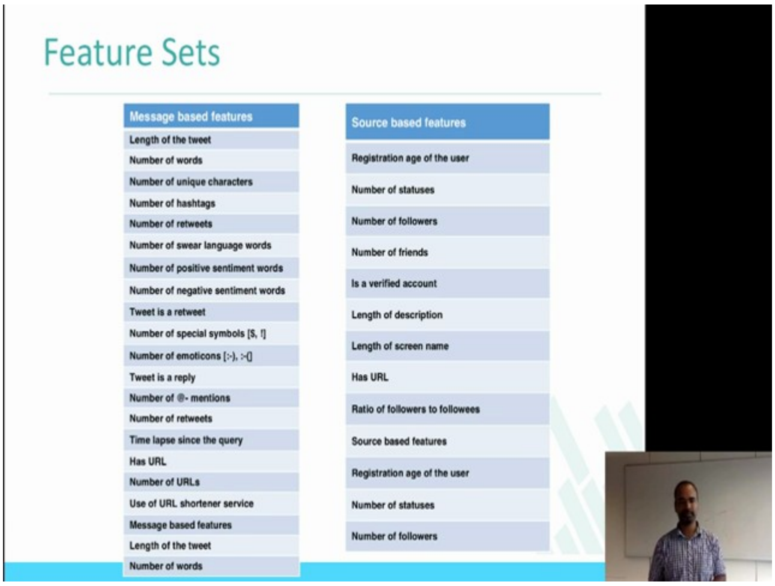


Annotation Results

- Each tweet annotated by 3 people
- Inter-annotator agreement (Cronbach Alpha) = 0.748
- 30% of tweets provide information (17% credible information) and 14% was spam

And that is why I said about Cronbach's Alpha which is something I will emphasize here, each tweet should be annotated by at least three people because that will give you more confidence in the data and when you do this it is called inter annotator agreement or Cronbach's Alpha. If you calculate that and if it is more than about 0.7 it is generally accepted that the data has more value or confidence in it, 30 percent of the tweets provide information which is in the step one users agree that 30 percent of the tweets are shown to them add information, only 17 percent have credible information and 14 percent was spam.

(Refer Slide Time: 38:21)



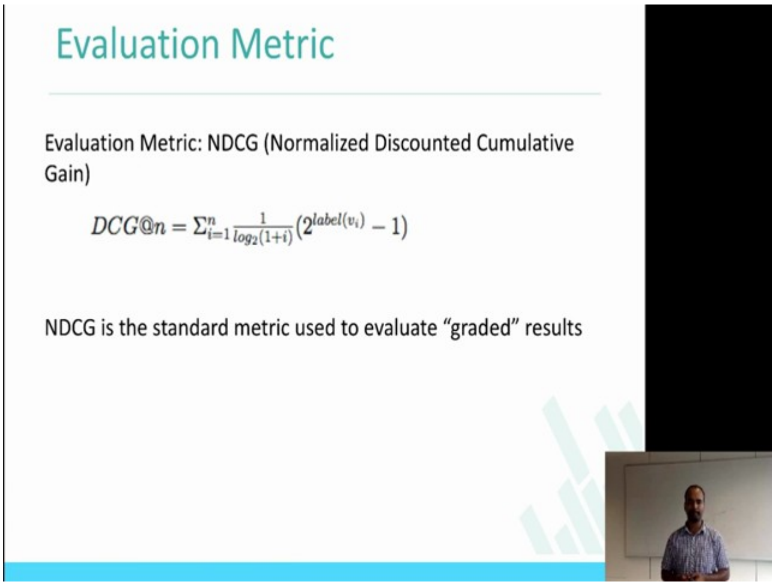
Feature Sets

Message based features
Length of the tweet
Number of words
Number of unique characters
Number of hashtags
Number of retweets
Number of swear language words
Number of positive sentiment words
Number of negative sentiment words
Tweet is a retweet
Number of special symbols [$, !]
Number of emoticons [:-), :-(]
Tweet is a reply
Number of @- mentions
Number of retweets
Time lapse since the query
Has URL
Number of URLs
Use of URL shortener service
Message based features
Length of the tweet
Number of words

Source based features
Registration age of the user
Number of statuses
Number of followers
Number of friends
Is a verified account
Length of description
Length of screen name
Has URL
Ratio of followers to followees
Source based features
Registration age of the user
Number of statuses
Number of followers

So, featue sets, we now discussed just now some time back about different features F1 and F2 in that slide here message based features and source base features which is again, if you look at it I told you about features from the posts which is message based features. So, features from the profile which is a source based features.

(Refer Slide Time: 38:46)



Evaluation Metric

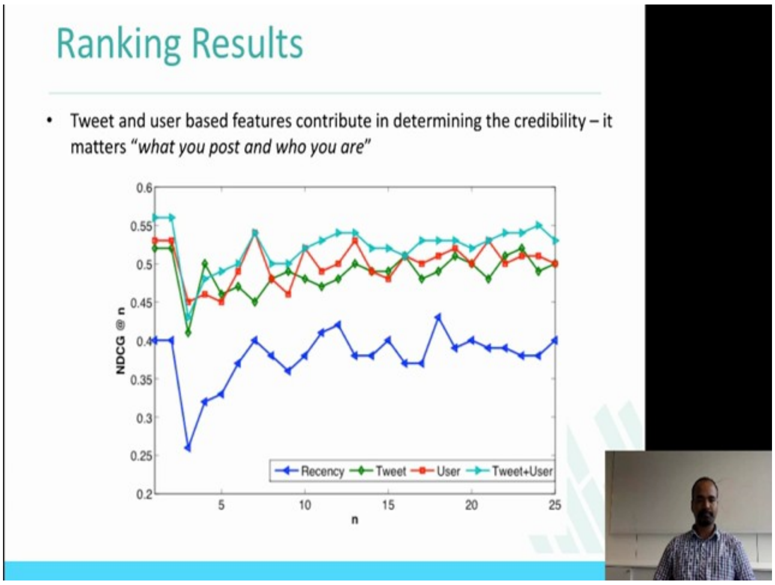
Evaluation Metric: NDCG (Normalized Discounted Cumulative Gain)

$$DCG@n = \Sigma_{i=1}^{n} \frac{1}{log_2(1+i)} (2^{label(v_i)} - 1)$$

NDCG is the standard metric used to evaluate “graded” results

Using these features we used a metric called the NDCG, which is normalized discounted cumulative gain it is nothing, but way by which you can actually mention the efficiency of the search, NDCG is being commonly used in finding, how good a search engine is performing in this case, we are using it to find out how what is the quality of the classification that we make whether it is legitimate or fake in this metric called NDCG.

(Refer Slide Time: 39:20)



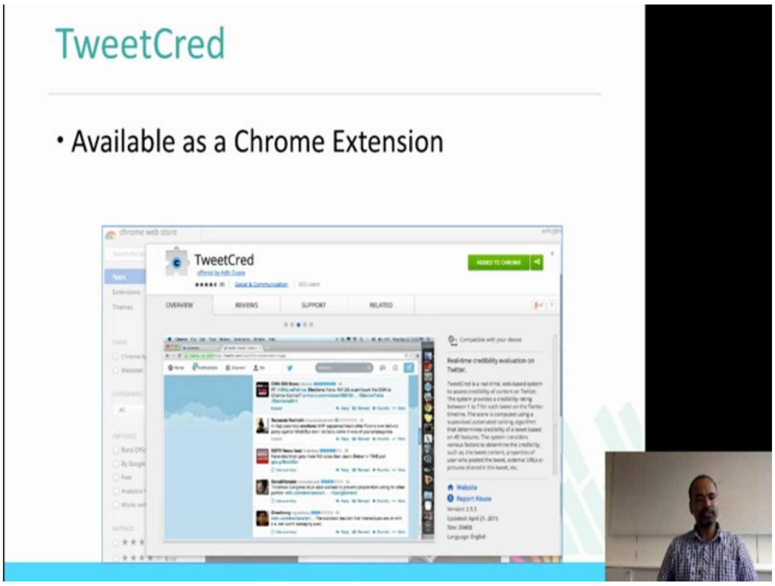
Ranking Results

- Tweet and user based features contribute in determining the credibility – it matters *"what you post and who you are"*

Also here is a graph for looking at the content from the tweets that we collected looking at the post it that we collected, recency, tweet, user, Twitter plus user, these are the features that we used. If you remember to find out whether a post is legitimate or not we are here we are drawing graph of n, which is on the x-axis and NDCG value, which is on the y-axis, you can clearly see that the tweet plus user which is at the top of the graph doing well in terms of the NDCG values. This basically helps to understand that what you post and who you are are a good features to make judgment on whether the content is legitimate or not, that is the kind of inferences that you should be chasing while you are analyzing the content from social media which helps in some actionable information also.

For example, here what you post and who you are helps to find out whether the post is credible or not which helps in making lot of decisions.

(Refer Slide Time: 40:34)

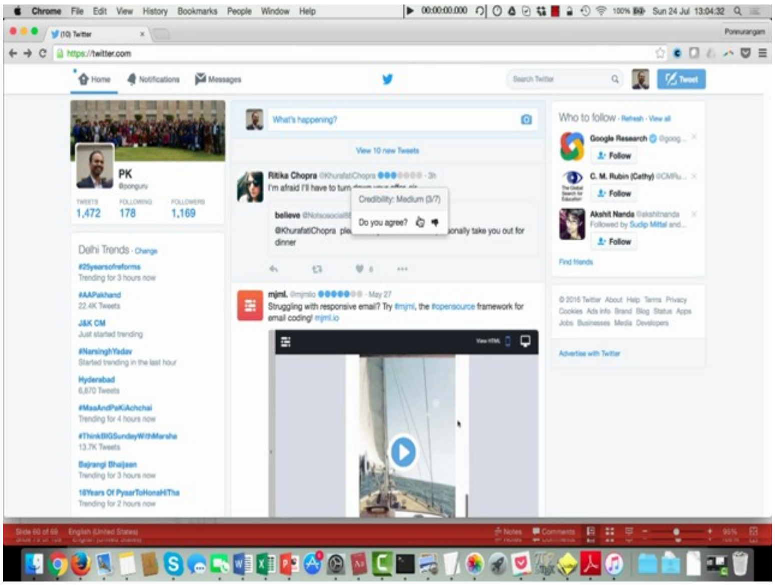


TweetCred

- Available as a Chrome Extension

Using this understanding of what feature works and what features do not work and what feature actually influences in finding out whether the post is credible or not, the TweetCred, a chrome browser extension was built and this extension it helps you to identify whether this particular post is credible or not. I'll just show you a light demo of the TweetCred extension and then I will walk you through, what this available in the chrome extension.

(Refer Slide Time: 41:02)



Ritika Chopra @KhurafatiChopra · 3h
I'm afraid I'll have to turn down your offer, sir

believe @Notsosocial
@KhurafatiChopra ple... onally take you out for dinner

Credibility: Medium (3/7)
Do you agree?

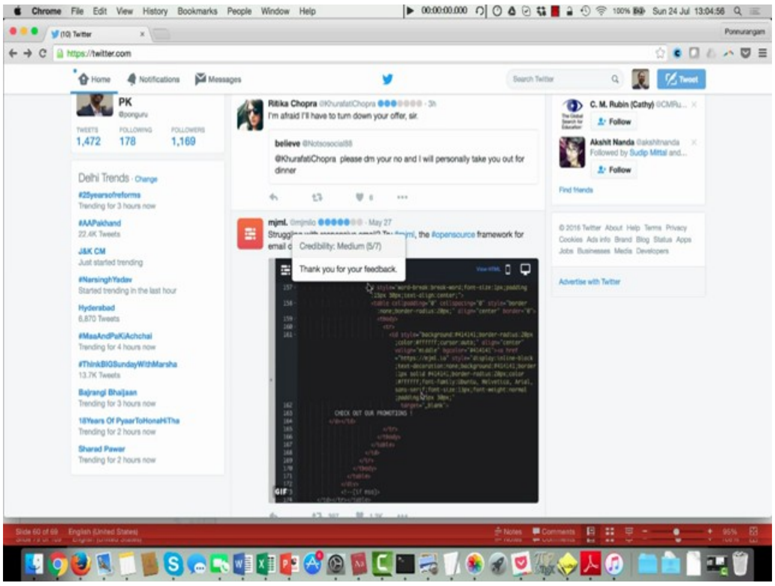
mjml. @mjmlio · May 27
Struggling with responsive email? Try #mjml, the #opensource framework for email coding! mjml.io

What I am showing you here is the Tweetcred browser plugin, chrome browser plugin which actually helps you in making a decision whether a particular tweet is credible or not. It just gives you; it basically uses all the features that we discussed until now, where it is being bundled with this chrome extension.

So, look at this tweet, if you go to Twitter dot com in your timeline, this information about whether a post is how credible is it will not be there this is coming from TweetCred. If you look at this it actually gives you a value of 3 on 7, it is calculating the value of credibility on a scale of 1 to 7 and in this case it is showing that this is my timeline in this case it is showing that this post is 3 on 7, this post is 5 on 7 and values like that. So, it is going to work for all the post that are in a timeline, it is going to work for what in your search its going to work for post in dm and things like that.

So, let us look at the values that it is presenting also. So, if you see here it is actually showing you a value of 3 on 7 and then when I hover it. It actually gives me information called credibility medium 3 on 7, do you agree? So, this is the way if you remember, the machine learning model that people using the feature you take that module and whenever we get feedback like this, we can actually go and update that model to make it more efficient.

(Refer Slide Time: 42:55)



PK @ponguru

TWEETS	FOLLOWING	FOLLOWERS
1,472	178	1,169

Delhi Trends · Change

#25yearsofreforms
Trending for 3 hours now

#AAPakhand
22.4K Tweets

J&K CM
Just started trending

#NarsinghYadav
Started trending in the last hour

Hyderabad
6,870 Tweets

#MaaAndPaKiAchchai
Trending for 4 hours now

#ThinkBIGSundayWithMarsha
13.7K Tweets

Bajrangi Bhaijaan
Trending for 3 hours now

18Years Of PyaarToHonaHiTha
Trending for 2 hours now

Sharad Pawar
Trending for 2 hours now

Ritika Chopra @KhurafatiChopra · 3h
I'm afraid I'll have to turn down your offer, sir.

believe @Notsosocial88
@KhurafatiChopra please dm your no and I will personally take you out for dinner

mjml. @mjmlio · May 27
Struggl... #mjml, the #opensource framework for email c...

Credibility: Medium (5/7)

Thank you for your feedback.

```
CHECK OUT OUR PROMOTIONS !
```

C. M. Rubin (Cathy) @CMRu...
Follow

Akshit Nanda @akshitnanda
Followed by Sudip Mittal and...
Follow

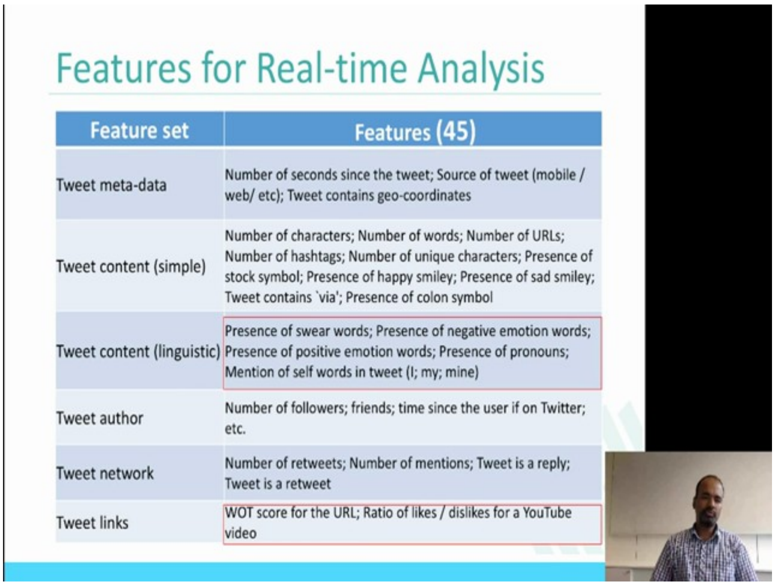
Find friends

© 2016 Twitter About Help Terms Privacy Cookies Ads info Brand Blog Status Apps Jobs Businesses Media Developers

Advertise with Twitter

So, in this case you could actually say that no, I actually agree with the value of 3 on 7, it gives you message saying thank you for the feedback. In this case, let us take it if I were to say that the value I do not agree with value then it actually asking what you are agreeing with. So, I say no this is actually more credible it should be actually 7, when it says thank you for feedback. So, what it basically does is, it is capturing these details from you and it is going to make use of it when we end up updating the model that was built at the back end for the TweetCred. This information can be used in making the judgment. So, that is the chrome browser extension of TweetCred, which basically takes the features in real time and makes the judgments and presents it to the user with the values of 1 to 7.

(Refer Slide Time: 43:27)

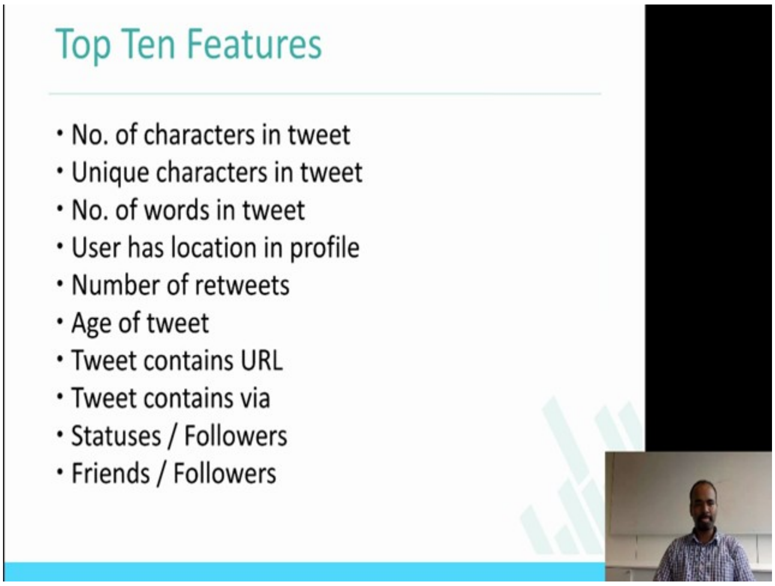


Features for Real-time Analysis

Feature set	Features (45)
Tweet meta-data	Number of seconds since the tweet; Source of tweet (mobile / web/ etc); Tweet contains geo-coordinates
Tweet content (simple)	Number of characters; Number of words; Number of URLs; Number of hashtags; Number of unique characters; Presence of stock symbol; Presence of happy smiley; Presence of sad smiley; Tweet contains 'via'; Presence of colon symbol
Tweet content (linguistic)	Presence of swear words; Presence of negative emotion words; Presence of positive emotion words; Presence of pronouns; Mention of self words in tweet (I; my; mine)
Tweet author	Number of followers; friends; time since the user if on Twitter; etc.
Tweet network	Number of retweets; Number of mentions; Tweet is a reply; Tweet is a retweet
Tweet links	WOT score for the URL; Ratio of likes / dislikes for a YouTube video

So, you may remember the features that we discussed in this lecture, but unfortunately all features cannot be actually used while doing it in real time, for example, finding out all the followers that you have and using some scores on the followers is actually hard. So, here are the 45 features that were actually used to while doing the real time analysis itself, specially I wanted to highlight on the presence of swear words, presence of negative emotion words, presence of positive emotion words, web of trust score, which is WOT score for the URL and ratio of likes and dislikes from the YouTube video which has links to the YouTube. So, these are the features that we did not discuss before. So, I kind of thought we'll highlight them when I'm presenting the slide. So, these features were used in building tweetcred demo that I showed you.

(Refer Slide Time: 44:24)

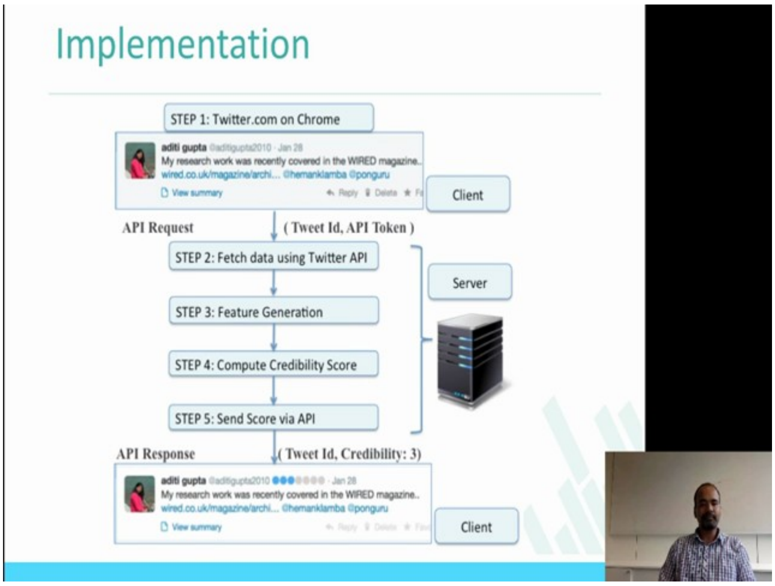


Top Ten Features

- No. of characters in tweet
- Unique characters in tweet
- No. of words in tweet
- User has location in profile
- Number of retweets
- Age of tweet
- Tweet contains URL
- Tweet contains via
- Statuses / Followers
- Friends / Followers

Of course, the common question is what are the top 10 features that actually makes the decision or which influences in identifying whether post is a credible or not more efficiently. It is number of characters in the tweet, unique characters in the tweet, number of words in the tweet, user has location in the profile, number of retweets, age of tweet, tweet contains URL, tweet contains via which is through, how the post was done, status and followers, friends and followers, those are the top 10 features of from Twitter, which can be actually used to make a judgment on whether a post is legitimate or not. Please keep in mind this is only for Twitter, the features that you look for Facebook, the features that you may look for Instagram, in other social network may be very different.

(Refer Slide Time: 45:16)



Implementation

STEP 1: Twitter.com on Chrome

aditi gupta @aditigupta2010 · Jan 28
My research work was recently covered in the WIRED magazine..
wired.co.uk/magazine/archi... @hemanklamba @ponguru
View summary

Client

API Request (Tweet Id, API Token)

STEP 2: Fetch data using Twitter API

STEP 3: Feature Generation

STEP 4: Compute Credibility Score

STEP 5: Send Score via API

Server

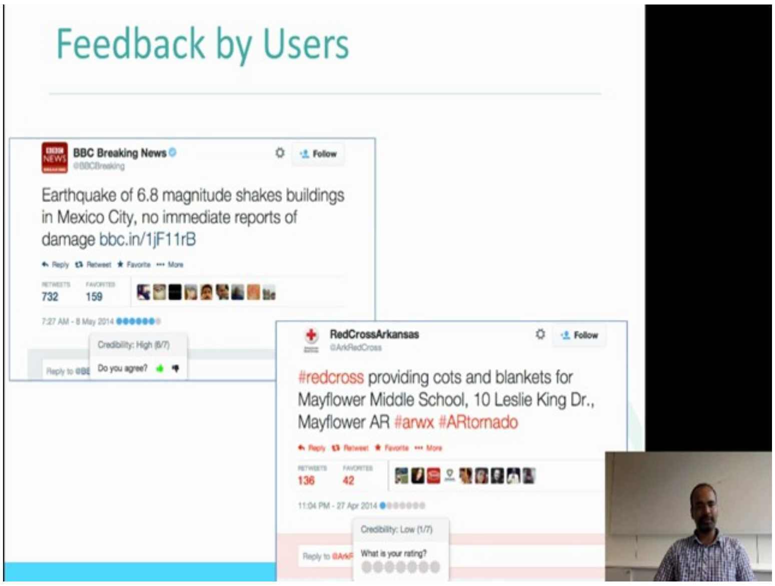
API Response (Tweet Id, Credibility: 3)

aditi gupta @aditigupta2010 ●●●○○○○ · Jan 28
My research work was recently covered in the WIRED magazine..
wired.co.uk/magazine/archi... @hemanklamba @ponguru
View summary

Client

Here is just a slide to show you how the implementation for the tweetcred was done which is chrome browser extension. There is a post that is on your timeline, it takes the post fetches data using Twitter API, which is the architecture that I showed you earlier, where feature extractions were done. Then, the model which is built, tweet is taken through API feature extracted the credibility score is computed with the techniques that we discussed until now and the values assigned back to the API, and then tweet ID and the credibility value comes back, it is presented in your timeline saying this value is actually 3 on 7, the demo I showed you. It is simple chrome extension that was about to show these values.

(Refer Slide Time: 46:05)



Feedback by Users

BBC Breaking News
@BBCBreaking

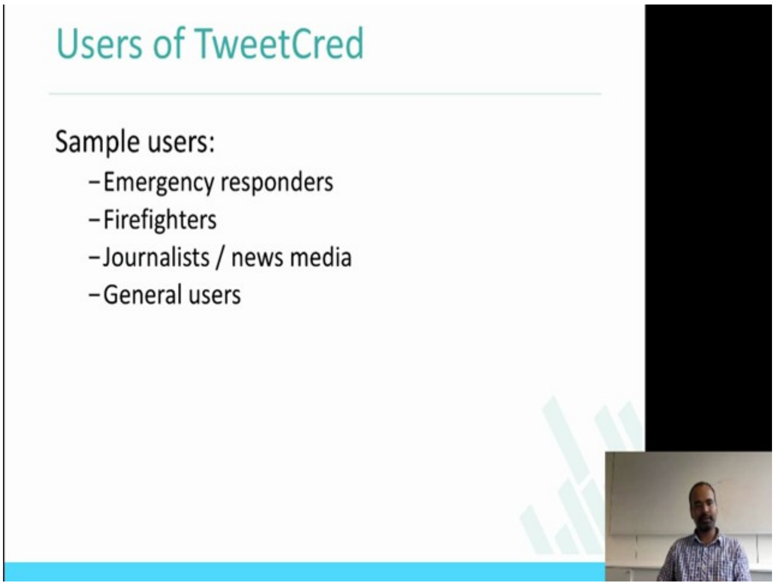
Earthquake of 6.8 magnitude shakes buildings in Mexico City, no immediate reports of damage bbc.in/1jF11rB

RedCrossArkansas
@ArkRedCross

#redcross providing cots and blankets for Mayflower Middle School, 10 Leslie King Dr., Mayflower AR #arwx #ARtornado

So, users can also give feedback to the system and that is showed in your demo TweetCred actually ask user to say agree or disagree with the values that is presented. If you agree that is okay, if you don’t agree please provide the information, please provide a value that it should be what you think it should be, that is what is presented in the top left which is when you agree , bottom right is actually saying if you disagree.

(Refer Slide Time: 46:35)



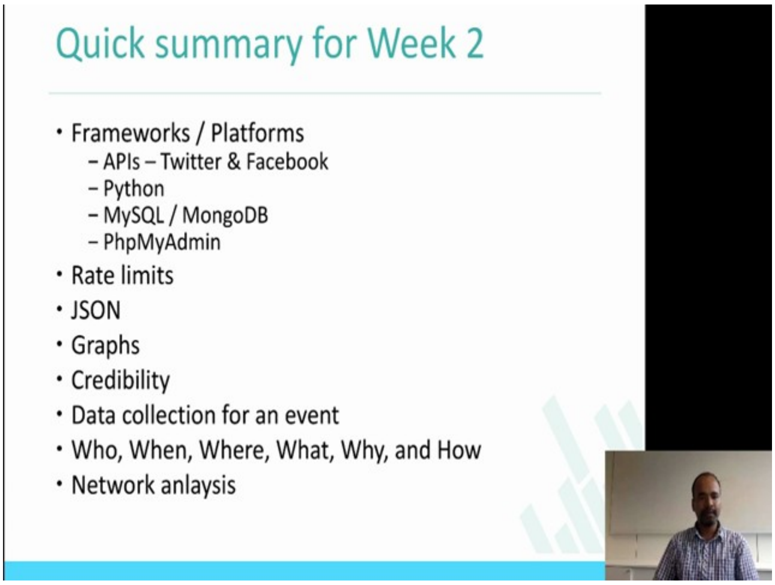
Users of TweetCred

Sample users:

- Emergency responders
- Firefighters
- Journalists / news media
- General users

Different types of users that can be foreseen using TweetCred type of tools. You can at least remember TweetCred is only one example that I am presenting here, there are many other tools that one to think of while analyzing social media content and information presented to the user. In this case emergency responders, fire fighters, journalists and news media and general users also have started using tools like this.

(Refer Slide Time: 47:01)



Quick summary for Week 2

- Frameworks / Platforms
  - APIs – Twitter & Facebook
  - Python
  - MySQL / MongoDB
  - PhpMyAdmin
- Rate limits
- JSON
- Graphs
- Credibility
- Data collection for an event
- Who, When, Where, What, Why, and How
- Network anlaysis

Let us do a quick summary of week 2, when we started we actually looked at API's, programming interfaces and we you also have a tutorial for Facebook in this week. Then we looked at very, very briefly what Python programming language, MySQL, Mongo DB, PhpMyAdmin and when you collect the data you are going to actually get the rate limits. Please remember that there is always going to be rate limits when you are collecting the data from these social media services and we talked a little bit about the format in which the social media service is store the data, which is JSON, when you collect the data, you are going to get JSONs which you have to analyze through your scripts and Facebook stores all the data in terms of a graph. We looked at that briefly then we started digging deeper into trust in credibility as a focus area. We looked at these concepts of trust and credibility through events being Boston marathon was one of the events, Hurricane Sandy is another event that we looked at we looked at these events. Through these events that I was trying to tell you how data is being analyzed, what kind of techniques are being applied on this data.

We looked at classification is one of the major technique that is used while designing whether a post is credible or not and during this analysis, I also told you about who, when, where and what, why and how are the basic questions that you can actually analyze using in the social media content and specifically we have also looked at some social network analysis techniques inputs. So, that is the week 2, hopefully you will go through the content and if you have any questions please go and ask in the forum, we'll be happy to actually answer there.