INTELLIGENT ALERTING AND RESPONSE USING SOCIAL MEDIA AS A SOFTWARE SENSOR

Tore Andre Nilsen & Jayesh Wadhwani
IntraPoint, Inc.[1]

Keywords

Sensor, Sentiment, software, analysis, media, social

Abstract:

Given that "one out of every seven minutes spent online is on Facebook" and "340 million tweets are sent each and every day" it is no wonder that Social Media has become so pervasive in our lives that it is used not only as a preferred medium for public alerting during a crisis but also as a medium to gather and share intelligence.

When collective intelligence is shared by so many, the accuracy and quality of the intelligence becomes suspect. The "roving gangs" rumors that spread and created panic after Hurricane Katrina is a prime example.

In this paper we focus on using "Social Media as a software sensor" and how it can be used for filtering of data to provide early warning and intelligent alerting.

Our research uses sensory data from Twitter, Facebook, Blogs, Websites, Pinterest, and Videos. It listens across the "social web" and tries to analyze real-time social conversations in conjunction with "location based searches" and "sentiment analysis" to filter out defective data.

Our concept of smart alerting not only uses social media but also the power of mobile communications to alert a location based focused populace.

We further leverage our research with "alert maps" to allow emergency responders to better manage the alerting process.

Finally, our research also reports on our work with sensors, smart objects, smart networks and the "Internet of Things", and also the results of our prototype implementation of these technologies in the realm of early warning and emergency management.

This next generation research is funded and backed by the Norwegian Research Council under the "IMMER" project.

---

[1] IntraPoint, Inc. NASA Ames Research Park, MS 19-46T, Moffett Field, CA 94035
tore.andre.nilsen@intrapoint.com

Introduction

Examples and literature abounds on how Social media has and can be used in Crisis response and management. The pervasiveness of Social Media is mind-boggling. 100,000 tweets are sent every minute while Facebook has 552 Million daily active users (Bennett , 2013).

Though these numbers may be awe inspiring there is a risky downside. Accuracy of the data that is retrieved from Social Media channels is suspect. Emergency responders acting on this information may get involved in a scenario they did not expect while the disaster victims may end up in a direr situation than they bargained for.

It goes without saying that filtering and mining of data during an emergency would be a herculean task.

In our research we take a more pragmatic and different approach to using social media data for early warning detection, crisis response and recovery.

We use the concept of "sensors", "smart objects" and "The Internet of Things" which allow us to make crisis management more intelligent.

We further show how social media can be integrated with emergency response plans.

Theory and Method

Accuracy and Information Overload

The freedom of the Internet and hence of the Social Media allows anyone to post information without fear and audit. If one were to make an incorrect statement about an emergency situation it could have severe consequences.

After the marathon bombings when social media was abuzz with updates, photos and prayers it also contained fake requests for donations. During hurricane Sandy FEMA setup a special page on their site titled Hurricane Sandy: Rumor Control (FEMA, 2012).

While traditional media prides itself with accuracy of content, social media triumphs in its speed of delivery. Accurate news not available at the time of crisis serves the purpose of being informative rather than actionable.

In a time of crisis inaccurate news or information only fuels fear and doubt. This fact came to life during the thousand of tweets during the Mumbai terror attack that were so widely dispersed in quality that the Government of India had to request everyone to stop tweeting as it was causing more confusion that helping the situation.

To make matters worse responders and victims are faced with a deluge of information. At the peak of the Japanese Tsunami more than 5000 tweets went out every second.

With people embracing Social Media as their primary medium of information it is difficult to filter out defective data in a time of crisis. Doubt during a crisis causes anxiety and in some cases chaos. People need to be assured of the accuracy and authenticity of the information they are presented with.

Though the problem of information overload can be taken care of by using technology at hand the answer to the question of accuracy remains elusive.

One of our primary aims of our research was to find a balance between speed and accuracy so that the response in terms of alerting is intelligent rather than ad-hoc.

Perception and Accuracy

To solve the accuracy problem we took a human approach. We decided to investigate how humans perceive accuracy.

Wikipedia defines perception as "Perception (from the Latin perceptio, percipio) is the organization, identification, and interpretation of sensory information in order to represent and understand the environment"

Perception is not passive but is shaped by our experiences, environment, biases, references and many other factors. Perception is what allows us trust one source of news more than another. Unconsciously we apply weight to our references and sources of information.

As we evolve we learn to bring perception and reality close together. Living in an imperfect world, sadly, the two never overlap. Events such as disasters always causes a rift between the two.

During a disaster, perception plays a major role in our decision making because disasters and emergencies are not an every day thing for the average person. If it is your first fire your decisions are based on perception. If it is your fifth your decisions will be based on reality.

Now consider this simple scenario. Let's suppose that you are walking down the street and spot a local town newspaper that is usually given away for free lying on the sidewalk mentioning a terrorist attack on an oil rig.

One or many of the following goes through your mind:

1. If there was a terrorist attack in your recent memory then you are concerned
2. You keep walking and pass by a newsstand and see the same news in a major daily. You are now worried. You need details.
3. You get your cell phone out and check your favorite news source. There is no mention of this incident.
4. In the meantime you receive a text message from a friend who works in the same oilrig company telling you about the attack.

You wanted verification because the weight you gave the local town newspaper was low. When you saw the news in a major newspaper to which you gave more weight your concern grew. Even though your popular news source did not mention it you were certain when you received the text from your friend because you friend received your maximum weight.

One another human trait we found that played a very important part during a crisis is people's mood. Could we quantify that?

Jan Piotrowski of the Economist puts it aptly "An urge to know what others will be up to next is part of the human condition. Soothsayers, fortune-tellers, stockbrokers—and publications like this one—have been catering to that obsession since mankind first began making plans for the future. Their record has been mixed. The biggest hurdle is the apparent unpredictability of individual behavior. But if you knew the mood of all those involved, might a clearer picture emerge?"(Piotrowski, 2012).

Fortunately for us there was already conclusive research on this topic called "Sentiment Analysis". Researchers have used Natural language processing, computational linguistics and text analytics to extract the mood and the sentiment of a given scenario. Most of this research has been implemented as a web service. In our research we used the "sentiment analysis" API from Viral Heat.

We polled news sources, Facebook pages, Twitter hash tags, blogs and other sources. For every item we gave a "weight". Depending on the number of occurrences on each item we ended up with an accuracy score. This score was used a trigger for public alerting, alert levels and emergency procedures.

One problem remained though. How would we deal with a disparate range of sources such as Facebook, Twitter, blogs, micro blogs, feeds etc. Get the data, extract information from it and then make that data actionable.

<u>Software Sensors</u>

Wikipedia defines sensors as "A sensor (also called detector) is a converter that measures a physical quantity and converts it into a signal which can be read by an observer or by an (today mostly electronic) instrument"(Wikipedia)

An example is a "Thermistor" in which the electrical resistance varies in proportion to the temperature.

In this example of a "hardware sensor" the key is the interpretation of the change or measurement. For example in the thermistor, the resistance by itself holds no practical value till it is mapped to temperature. Also the quality of the mapping algorithm will dictate the accuracy.

If we now consider a snippet of software that reads from a data source such as a feed or a file then by itself the data is of no consequence. But if this piece of software were to read this data and trigger certain events based on some programmable criteria placed on this data we could safely say that this is a sensor… a "software sensor".

For example a snippet of software could in essence monitor user logins and convert the pattern of logins to a machine-readable pattern, which could show a probable hack attempt. We could also have a snippet of code output the rate of valid logins per hour and a human observer could make an interpretation from that.

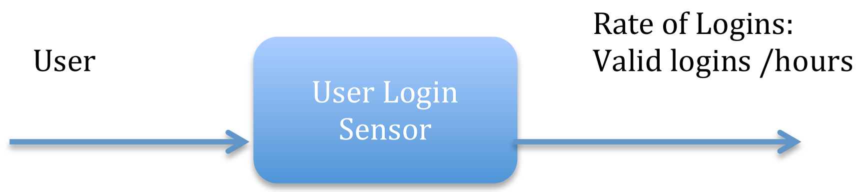

Figure 1: User login as a software sensor

Taking this analogy a step further we could have a piece of code that would read a twitter feed based on certain criteria, e.g. #terrorattack

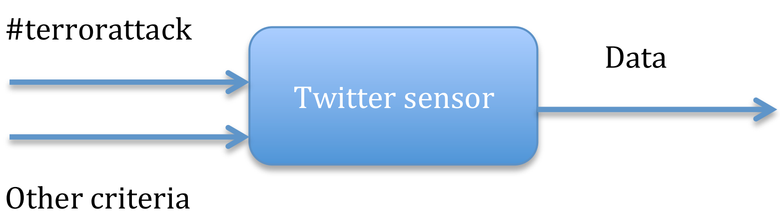

Figure 2: Example of a Twitter sensor

Another sensor is the "Sentiment Analyzer" sensor:

Figure 3: Sentiment Analyzer sensor

Modeling sources and analyzers as sensors allowed us to think of them in the realm of Internet of Things. To that end we decided to make them smart objects. In other words they would have their own identity and functionality.

Like any other smart objects "Smart objects using Social Media sensors" would talk to each other or to a central authority over the Internet. They would use a protocol that would encourage decoupling. This would allow us to add or remove smart objects on the fly. Our thought process can be visualized as follows:

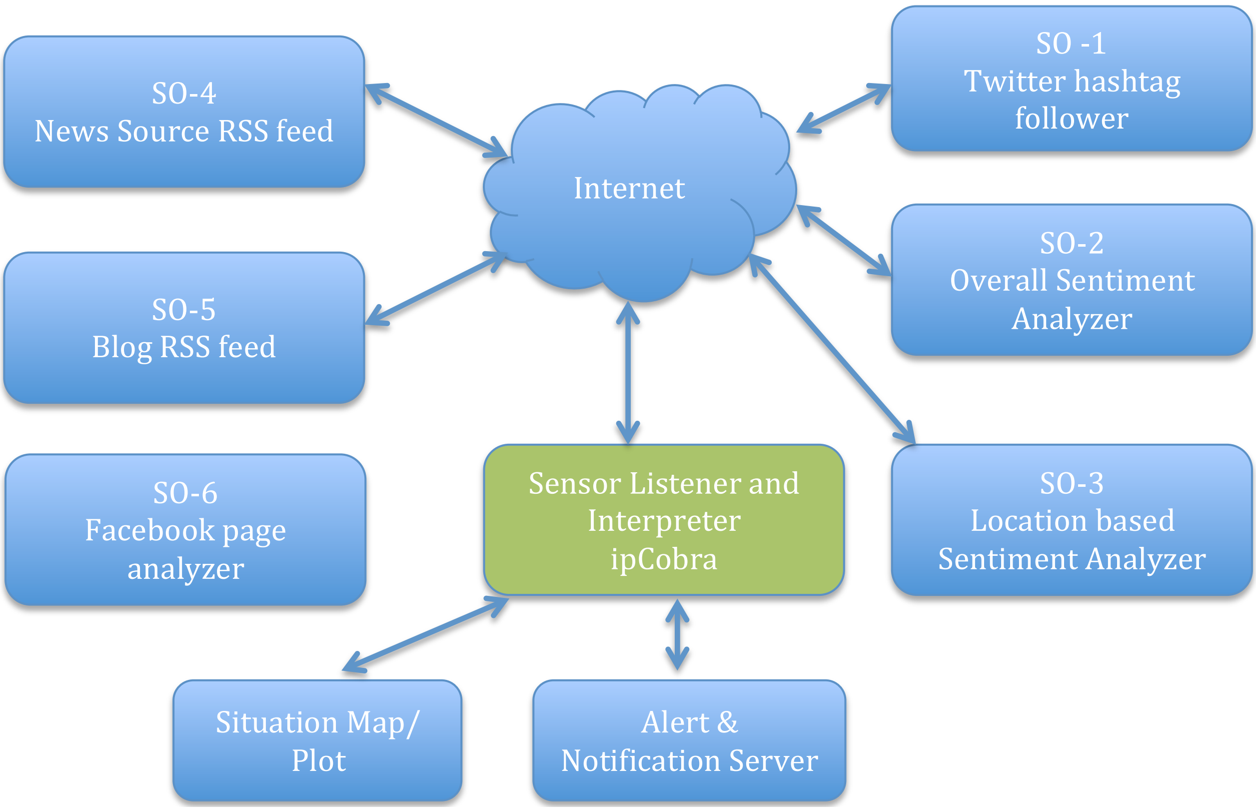

Figure 4: Social media as smart objects

In this case all the smart objects are sent their instructions by the Sensor Listener and the results are sent back for interpretation and execution of alerts. We also added in a Situation plot to provide a more location based experience.

This essentially solved our problem of managing disparate number of sources of information.

Results

Our final implementation contained the following:

- Sensor listener code named 'ipCobra' implemented as a real time event driven kernel
- SO – 1, Twitter hash tag follower using Twitter API
- SO – 2, 3 Viral Heat's sentiment API[5]
- SO – 4 News Source RSS feed using standard RSS architecture
- SO - 5 Blog analysis using standard text analytics.
- SO – 6 Facebook page analyzer using Facebook API and text analytics.

We created a light protocol that rides on top of the standard AMQP 0.91 protocol. As this is a publish/subscribe protocol coupling of objects is insignificant. The Internet was a private Wi-Fi network and all communication used standard http/https protocols.

We used "terrorist attack" as our subject of research. The following weights and rules were applied:

SO – 1: Used #HSEM, Homeland Security Emergency Management. Weight 10. Number of tweets about "terrorist attack" in one hour. Maximum 10.
SO – 2, 3: Used keyword "terrorist attack", mood = negative, probability < 1.0, weight 7
SO- 4: RSS feed from BBC – Top stories – Weight 8, 2 news reports or one news report with update every hour. Should contain the words "terrorist attack" and the location
SO – 5: New York Times blog. Weight 5. Should contain the words "terrorist attack" and the location with multiple updates in an hour. Maximum updates required 5.
SO – 6: FBI's Facebook page. Weight 9. Any status updates involving terrorist attack.

Maximum/ideal limits. Total: Weight * max occurrences

Smart Object	Weight	Occurrence	Occurrence normalized to 10	Total	Comment
SO - 1	10	10	1	100	Number of tweets/hour
SO – 2	7	1	10	70	Probability
SO – 3	7	1	10	70	Probability
SO – 4	8	2	5	80	BBC news
SO – 5	5	5	2	50	NYT blog
SO - 6	9	1	10	90	Facebook page

Total: 460

So, for a particular incident scenario our perception of accuracy is listed above. Closer the total is to 460 more accurate will be the severity of the incident.

We must warn the reader that this number is not a generalized number but specific to this incident. It was a hurricane we might use different weights for the same sources in the table above.

Creation of this table would be part of creating emergency response plans for given situations.

Reading 1 – News media not active. Social media moderately active

Smart Object	Weight	Occurrence	Normalization factor	Total	Comment

SO - 1	10	8	1	80	Number of tweets/hour
SO – 2	7	0.5	10	35	Probability
SO – 3	7	0.3	10	21	Probability – Location based
SO – 4	8	0	5	0	BBC news
SO – 5	5	0	2	0	NYT blog
SO - 6	9	1	10	90	Facebook page

Total: 226

Reading 2 – News media not active. Social media somewhat active

Smart Object	Weight	Occurrence	Normalization factor	Total	Comment
SO - 1	10	1	1	10	Number of tweets/hour
SO – 2	7	0.1	10	7	Probability
SO – 3	7	0.1	10	7	Probability – Location based
SO – 4	8	0	5	0	BBC news
SO – 5	5	0	2	0	NYT blog
SO - 6	9	1	10	90	Facebook page

Total: 114

Reading 3 – News media fully active. Social media not active

Smart Object	Weight	Occurrence	Normalization factor	Total	Comment
SO - 1	10	0	1	0	Number of tweets/hour
SO – 2	7	0	10	0	Probability
SO – 3	7	0	10	0	Probability – Location based
SO – 4	8	2	5	80	BBC news
SO – 5	5	5	2	50	NYT blog
SO - 6	9	0	10	0	Facebook page

Total: 130

Reading 4 - Only Social media – Only Twitter active

Smart Object	Weight	Occurrence	Normalization factor	Total	Comment
SO - 1	10	10	1	100	Number of tweets/hour
SO – 2	7	0.3	10	21	Probability
SO – 3	7	0.3	10	21	Probability – Location based
SO – 4	8	0	5	0	BBC news
SO – 5	5	0	2	0	NYT blog
SO - 6	9	0	10	0	Facebook page

Total: 142

Tabulating the results:

Reading	Score	Comment
Ideal/Max	460	Ideal or maximum
Reading 1	226	News media not active. Social media moderately active
Reading 2	114	News media not active. Social media somewhat active
Reading 3	130	News media fully active. Social media not active
Reading 4	142	Only Social media and only Twitter active

Discussion

It is clear from the above readings that we could essentially use the scores to trigger alerts by assigning the scores to alert levels like so:

Score < 150 – Green ALL OK
Score between 150 and 300 – Amber EARLY WARNING
Score above 300 – Red EMERGENCY

Though these are hypothetical alert levels, one could finely tune the scores to match the required alert levels. As emergency plans and notification/alert policies vary from company to company we have not made any attempt to generalize the scores to any specific alert level.

The beauty of using such a system is that one can add/remove smart objects or duplicate them and set the scores accordingly. Moreover the sensor listener also allows for training so that alert levels can be adjusted iteratively over a period of time.

We have also had success implementing objects on the fly depending on what the score is. For example if the score is less than 150 and we wanted confirmation then we would start up another smart object with a news source sensor.

In addition to this we have successfully implemented decision logic depending on the emergency plans. For example to send a specific alert if SO-1 > 50 and SO-6 = 90.

The number of ways this can be used is left to the reader's imagination.

References

(Bennett, 2013). 100 Amazing Social Media Statistics, Facts and Figures http://www.mediabistro.com/alltwitter/100-social-media-stats_b33696

(FEMA, 2012). Hurricane Sandy: Rumor Control http://www.fema.gov/hurricane-sandy-rumor-control

(Piotrowski, 2012). Getting in the mood. http://www.economist.com/node/21537049

(Wikipedia) Sensor http://en.wikipedia.org/wiki/Sensor

Acknowledgements

We are indebted to the support provided by the Norwegian Research Council and our industry partners. We would also like to thank Viral Heat for giving us access to their Sentiment Analysis tool.

Biographies

Tore Andre Nilsen has been working with Internet technology since the early 90's. Through studies at NTNU and his work at SINTEF, he has obtained extensive experience in Internet system development and strategic management applications. He has spent the last 10 years in combined management and technology positions working with Knowledge Management, Content Management, Project Management, Learning Management, Enterprise Architecture and Crisis Management.

Today, Nilsen is the founder and owner of multiple software businesses in Europe and the United States. He is also involved in multiple product development projects, research projects and consulting projects. He is the founder of Wise Online Services Inc., General Partner of IntraPoint AS and the Chairman of TAN holding AS. He currently holds the position of C.T.O. at IntraPoint.

Jayesh Wadhwani earned his Master of Science degree in Electrical Engineering from the New Jersey Institute of Technology. With an I.T. career spanning over 30 years he has held many engineering and leadership positions in various industries. Currently he works as a Senior Research Engineer at IntraPoint's Advanced Technology Labs.

His work encompasses software and hardware development particularly developing intelligent Crisis Management Systems in the realm of Internet of Things.