Fusing Decision Support Systems into the Fabric of the Context
A. Respício and F. Burstein (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-073-4-323

Context-Aware Twitter Validator (CATVal): A System to Validate Credibility and Authenticity of Twitter Content for use in Decision Support Systems

Saguna[a,b,1] , Arkady ZASLAVSKY[c,b] and Cécile PARIS[c]
[a] *FIT, Caulfield Campus Monash University, Australia,*
[b] *LTU, Luleå, Sweden,*
[c] *CSIRO, ICT Centre, Australia*

Abstract. Decision support systems (DSS) are beginning to use content sourced from social networks such as Twitter to provide decision makers with information to make timely and critical decisions. Misleading information obtained from Twitter can lead to adverse outcomes as well as cause trust issues within DSSs. In this paper, we propose and investigate a context-aware Twitter validator (CATVal) system to validate credibility and authenticity of Twitter content at run-time for use in DSS. We build, store and update a credibility index for Twitter users and verify user's context information each time a user tweets. The proposed system can benefit a DSS by providing credible and dependable information while detecting misleading and false information sourced from Twitter and possible other social media.

Keywords. Decision support system, context-awareness, social networks

Introduction

Decision support systems (DSS) in recent times are advancing towards the use of collaborative and cooperative decision making within organizations [1] which is one form of collective decision making. The last decade has witnessed the migration from individual decision making to group decision making [2, 3] in many organizational decision making processes. The Internet allowed for the faster access to information and the use of web-based tools for decision making proved beneficial as they enabled a larger number of users to participate in the decision making process [2]. Now, with the advent of social networking and its widespread use by people, there is immense potential for its use in decision support systems. Information disseminated within social networks can be efficiently used to make effective decisions within organizations [3, 4]. Out of the many social networks available to users today, 'Twitter' stands out in terms of how it efficiently disseminates information about events, news, occurrences within and around a user's daily life. It has also become a place to voice views and opinions about various activities relating to those who 'tweet' them. Twitter also maintains a

[1] Corresponding Author (email: saguna.saguna@monash.edu).

section on trending topics which displays the top 10 terms mentioned by its users at any given time. Along with trending topics, there is immense push towards aggregating twitter content by different organizations to capture information of use to them and that can, thus, facilitate improved decision making [4]. The use of such information, which is aggregated via social network users, is susceptible to risks that are inherent to Twitter security.

In recent past, Twitter has succumbed to a number of security attacks and breaches [5-7]. Although, spamming twitter accounts are detected [5-7], based on tweet content, account properties and other usage criteria, there is still further need to identify the authenticity of twitter message sources. Also, authors in [5] have shown how Twitter users can fall into mainly three categories, humans, bots and cyborgs. This raises the question "Are all those using Twitter credible users?" followed by "How can we source authentic and credible information from Twitter?" Thus, there is a need to build a system that utilizes information only from those Twitter users who can be trusted to provide credible and authentic tweets about any news, events and happenings relating to them or occurring around them.

In order to achieve this goal, we propose a context-aware Twitter credibility measurement tool which provides information from credible and authentic Twitter sources to decision makers within decision support systems. Recently, a large number of researchers have attempted to understand the usage and communities of Twitter [5, 8-11]. In this paper, we extend on [5] which utilizes a number of criteria such as timing entropy, device makeup, URL ratio, Bayesian text and followers to friends ratio to distinguish between Twitter users as humans, bots and cyborgs but lacks the use of context information. In [5], the authors do not investigate further to distinguish users on the basis of their context information. Thus, we incorporate the user and his/her environmental context information such as time, location, type of device, activity, network related context and API based tweet for identity validation of Twitter user. Further, we use this information to build on a credibility index of Twitter users which then enables the process of using only credible and authentic information in decision support systems from real and credible Twitter users as information sources.

This paper is structured as follows: The use of social networks in DSS and motivating scenarios are presented in section 1. Section 2 presents Twitter and its role in DSS. Section 3 details the security issues in Twitter. Section 4 presents our proposed context-aware Twitter validator with its architecture and section 5 gives the conclusion.

1. Social Networks for Decision Support Systems

1.1. Motivating Scenario 1

Consider a scenario where a customer John interacts with Harry, an employee of the company, A in regards to some work. John is unhappy with his dealing with Harry and he utilizes the social network 'Twitter' to vent his anger and disappointment. John creates a number of new fake user profiles or hijacks/hacks into other existing profiles of Twitter users to tweet adversely about Harry. By targeting Harry in his tweets John spreads the word about his negative performance. This in-turn affects both the company and Harry's individual performance adversely. The manager notices these developments and sets up an inquiry into the matter. Unfortunately for Harry, the Manager is unable to determine the authenticity of the Twitter content based on which

the DSS created the negative reports and fires Harry. This is illustrated in figure 1. Thus, there is a need to validate the sources of negative posts on Twitter. To understand whether there are many real and genuine unhappy customers creating negative Twitter content for Harry and his company or just one unhappy customer creating a lot of negative publicity for Harry.

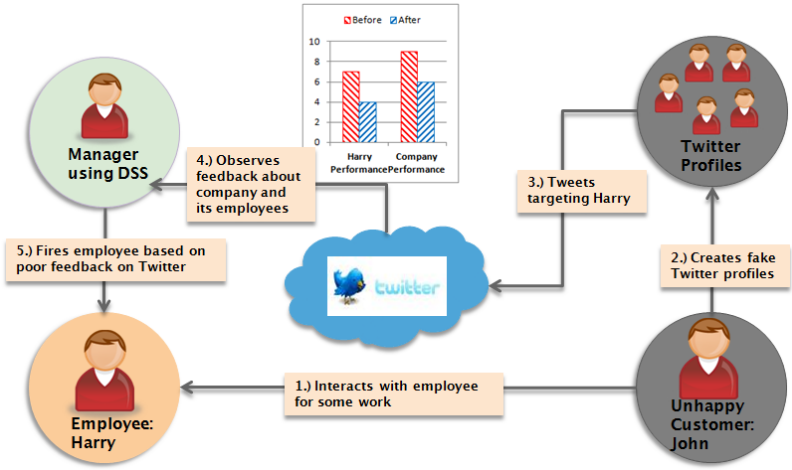

Figure 1. Motivating Scenario 1.

1.2. Motivating Scenario 2

Similarly, in scenario 2 shown in figure 2, a rogue, misleading entity/person is creating negative opinion and views about an organization's performance (corporate or government). Such Twitter content needs to be validated before being incorporated in a DSS for consideration by decision makers. It is important for organizations not to be

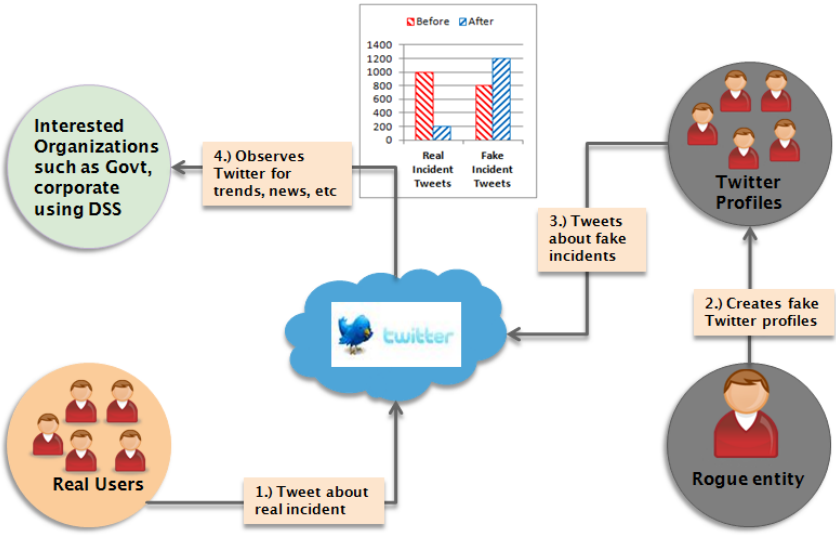

Figure 2. Motivating Scenario 2.

mislead by non validated Twitter content. Decision makers should only receive DSS content originating from real and verified users. Thus, a mechanism is required to validate the source of Twitter content before it is used in a DSS.

2. Twitter and its Role in Decision Support Systems

2.1. Twitter basics, type of content and third-party applications

Twitter is a micro-blogging platform where users are allowed to post tweets (Twitter status messages) up to a maximum length of 140 characters. Since its launch in October 2006, Twitter has witnessed immense growth globally with users from both developed and developing nations [4]. Initially, Twitter founders built the platform to help users share (with family, friends and other people who follow them) the answer to the question "What are you doing?" However, many studies have shown [4, 8, 12] that 'Tweets' (or Twitter status messages) are not only used for sharing the answer to this particular question but also for other purposes such as expressing views, opinions, support for a cause, marketing/publicity campaigns, news, simple communication, emergency or crisis situations, etc. In [13], a study conducted by Pear Analytics on 2000 tweets over a 2 week period classified tweets into six main categories as 'news', 'spam', 'self-promotion', 'pointless babble', 'conversational', 'pass-along value'. Similarly, different studies conducted in [8, 9], found that Twitter users tweeted for a number of reasons, such as a) keeping in touch, b) promoting certain kind of interesting information, c) collecting information of interest, d) asking for help and opinions and e) as an emotional stress reliever by sharing emotions and feelings.

Twitter is built using the publish-subscribe model. Subscribers can read and follow Twitter users. Tweets can be made via a SMS, website or a web-service application built using Twitter API. Twitter has approximately 383 million user accounts as of 1st January 2012 [14], with a prediction of 500 million accounts by March 2012 [15]. The difference between blogs and micro-blogging is the brevity of the content posted by users. Micro-blogging provides a faster means of communication whereas users used to create one blog in one or more days, they are now able to create shorter posts of 140 characters with lesser input of time and thought. This has led to the popularity and vast use of Twitter.

One way Twitter users create 'tweets' is by directly logging into the Twitter website using a web browser from a device such as mobile phone, laptop, PC or tablet. Users in certain countries can also post a tweet using SMS. Apart from these 2 ways, a large number of third party applications exist which use the Twitter API to post tweets. The tweets generated via such applications may be directly written by the users themselves or, at times, created by the application software. Applications include, for example, TweetDeck [16], Twitterific [17] and StatusShuffle [18]. Tweetdeck and Twitterific can be used from any user device and can help the user schedule their tweets. StatusShuffle creates tweets on behalf of users. Tweets can be funny one-liners or simple anecdotes or quotes from famous people. User tweets can also be automatically generated by sensors and detecting activities performed by them as shown in [19]. Other companies like SproutSocial [20], TweetAdder [21] and Twopcharts [15] are hired for marketing and/or monitoring content within social networks by corporate or individual users. These also help with social media monitoring, including monitoring the brand, industry and competition across social

media and the web. Further, these applications also help in engaging with customers over social media.

2.2. Aggregation and analysis of Twitter content

Recent times have witnessed an increased interest in the type of content Twitter users create. This interest is a result of the significant potential such content can have in providing different parties (such as corporate, media, government organizations, etc.) with knowledge, feedback, opinions and views about the products and services they offer to their customers. Such type of Twitter content aggregation is called 'Interest-based' aggregation. Another reason for aggregating user content on Twitter is to merely gain an understanding about Twitter usage and communities as well as about other characteristics relating to the type of content, Twitter activity and users.

2.2.1. Interest based aggregation and analysis for organizations

Twitter content from users is filtered and aggregated based on different interests of organizations. Such aggregation of content and its analysis provides crucial information required to make decisions. Capturing relevant information in a timely manner can facilitate the decision making process. Decision support systems are built with the capabilities to collect relevant information required for effective and timely decision making from social networks such as Twitter. Twitter also performs aggregation within the site by the use of 'Trending Topics'. It displays the ten top most topics that users tweet regarding at any given time. These are listed on the home page of the site and interested users can view the latest tweets about the topic. Users use a hash tag symbol (#name_of_topic) while tweeting about a particular topic or issue, and Twitter aggregates such keyword based tweets. Similarly, there are companies (e.g., SproutSocial [20]) that provide specialized services to aggregate and analyse Twitter content based on specific topics of interest[20]. Such social media based information is growing with time and is now considered crucial in decision making processes by organizations [3, 22, 23]. Organizations require a variety of information relating to the products or services they offer, the roles they play in society and their functioning and operations. Specific communities and groups of users on Twitter as well as those who form a company's consumer base can provide real time information in the form of opinions, views, feedback. This information can be crucial in articulating, for example, the popularity of a product and the type of customer base it holds in the market.

2.2.2. Aggregation and analysis for creating Twitter usage statistics

Generalized Twitter content aggregation is performed by a number of recently emerging companies and statistics about Twitter users, account types, type of content, location of content origin, purpose of content generation or user intention are some of the features studied. As mentioned previously in section 2.1 this content can be categorized into news, events, self-promotion/marketing, pointless babble, conversational, pass-along value, seeking/giving help or advice, broadcasting thoughts, emotional feelings and about other daily life activities. The location of tweets is also studied as, for example, in the case of an earthquake. Initially there are many tweets from the affected region about the event occurrence and experiences of those present in the location. This is followed by tweets occurring from different locations around the world, re-tweeting, querying/communicating or expressing grief and concern about the

event. Studies have shown how Twitter activity and trends in different locations have occurred during and after the occurrence of major events in the world [24]. Also, studies show how Twitter usage is affected by the number of followers (those people who follow a Twitter user) and followees (those people who the Twitter user follows) [5]. Twitter users with large number of followers tend to be more active and tweet with greater frequency, while users with large number of followees tend to be more of silent observers and seem to collect information passively. Studies also show how tweet activity varies across different times of the day as well as in different time zones [5]. The types of tweets vary in office hours, after work hours, weekends and holidays [5].

A number of other statistics reflect Twitter usage. We discuss these in the next section on security issues in Twitter as they mainly concern how Twitter can be exploited and misused.

3. Security Issues in Twitter

Since its inception, Twitter has faced a number of security breaches and threats. It is vulnerable to security threats from spammers, hackers and malicious content creators. We list these threats below and discuss the possible dangers to users as well as to those following Twitter content:

- Use of bots: Twitter is not used by human users alone. Spammers have created bots with a user account on Twitter. These bots follow a large number of users with the aim of getting those users to follow them back. Once unassuming users start following back their new bot followers, they are flooded with spam messages.
- Exploiting trending topics: These bots are also used to exploit trending topics by posting messages with the hash tag (#trending_topic) and links to malicious content and other spamming messages. Users following a trending topic easily succumb to such attacks.
- Worms on Twitter: In 2009, the Koobface worm infected a number of Twitter user accounts. The worm is installed on a PC/device when an inadvertent Twitter user clicks on a malicious link of a video posted by someone on Twitter. The worm then posts bogus tweets with links to malicious videos on behalf of the user every time they log on to Twitter. This further infects computers and devices (with malware) of those following the user, believe that they are clicking on links posted by a trusted user who is within their network of friends.
- Third-party applications and loss of user credentials: Recent studies have shown that a large number of Twitter users use API based third-party applications for posting tweets such as TweetDeck, Twitterffic, etc., many of which are Adobe Air based, thus easily susceptible to security attacks. These applications are vulnerable to attacks from hackers and lead to users losing their login credentials.
- URL shortening: Twitter uses a URL shortening service for users who wish to share web links on Twitter. This is an easy way to post malicious links. The shortened versions of the original web link direct a user to the original web

page. If misused, unassuming users fall prey to malicious content on Twitter by clicking as they are unable to determine whether it is malicious or not merely by viewing the link.

- Devices Tweeting on Twitter: Twitter has recently seen devices like power meters and heart rate monitors posting messages on Twitter. These devices belonging to users are a means of interacting and gathering information relating to different aspects of users' lives, such as their health, electricity consumption at home, etc. Such devices can be easily hacked to post malicious content on Twitter, as they are not equipped to handle such security threats.
- Injecting malicious or undesirable tweet content over the network: Another possible means of posting twitter content is by 'over the network hacking', where a user's tweet content can be changed without his/her knowledge. 'Over the network hacking' involves intercepting the network traffic of a Twitter user and implanting or replacing an original tweet message with a fake/malicious content. This is more likely when using third-party applications on mobile devices or desktop machines as they do not have robust security mechanisms in place.
- Fake User Profiles: Twitter users are also vulnerable from rogue or fake user creating fake profiles taking over the identity of someone and then creating malicious or undesired twitter content on the real user's name. This can lead to serious implications for the real user in whose name the fake profile is created. Some of the known forms of such attacks are common amongst celebrity profiles, where Twitter has countered this by the introduction 'Verified accounts' for well known personalities or public figures. But in the case of normal users, such type of attack still poses a serious threat and can affect their daily life adversely. This can also affect adversely in a DSS, when a large number of seemingly real but fake profiles are created which produce Twitter content that is incorrect or can lead to incorrect information being fed into the DSS.

In this paper, we attempt to address a number of issues relating to posting of Twitter content without the user's knowledge by using a user's context information and building a context-based tweet authenticity check mechanism. Malicious or incorrect tweets can lead to a number of problems for the user as tweets are read by family, friends and other interested parties. If a hacker intends to cause harm to the user or is attacking an organization, for example posting a negative opinion about that organization from a hacked user account, it can lead to the loss of reputation for the user or the organization that collects information on Twitter from users it considers to be genuine.

In [5], twitter user accounts belonging to humans, bots and cyborgs are identified based on the following criteria:

- Timing entropy of tweets: the tweeting interval is used to measure a Twitter user's behaviour and helps in detecting automated users who tend to have a periodic or regular timing of tweets. Though, it is possible that advanced bots may be able to find a way to overcome such detection by employing

techniques which eliminate such regular/periodic tweeting by making it more random.

- Inclusion of spam in content: This is detected through machine learning techniques and the content of tweets is checked for spam. It is observed that the presence of spam is likely to mean that the post was generated by bots.
- Account related properties: Other Twitter user's account related properties are checked for detecting humans, bots and cyborgs, for example, whether the tweet is posted manually using a Web browser or via a mobile device or if it is generated automatically using Twitter application based on Twitter API via HTTP. The followers to friends ratio was another useful way of detecting bots, as bots previously had more friends than followers, athough in recent times bots have changed the way they operate by keeping the ratio close to one and not following their friends, i.e. if those friends do not follow them back. This makes it hard to use this criterion for detecting the newer bots.

In this paper, we utilize user and his/her environmental context information along with existing criteria to detect an authentic tweet originating from a credible user. We present our proposed architecture of a context-aware Twitter user credibility check system in the next section.

4. Context-aware Twitter Validator System (CATVal)

We propose a Context-aware Twitter Validator System (CATVal) for the validation of Twitter message source (user) to insure the correct and reliable delivery of information in a DSS. Such a system is crucial for decision makers in order to reliably make decisions based on authentic information sourced from Twitter. As mentioned in the previous section, Twitter content can be injected with false or malicious information through a number of ways such as compromising of user credentials or injection of content over the network. Further, fake users or profiles can be used to disseminate inaccurate content. Figure 3 shows the high-level CATVal system. The use of context

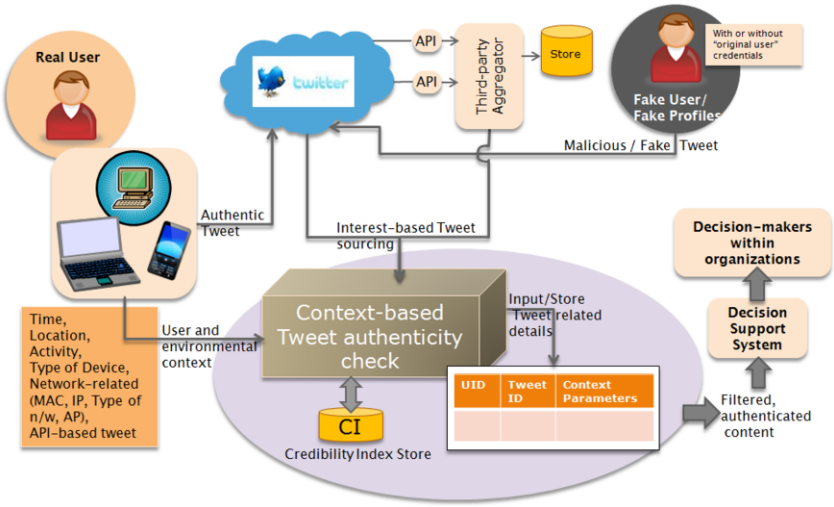

Figure 3. High-level CATVal System.

information is essential for checking the authenticity of the tweet. The real user and the fake/rogue user are distinguished using context information sourced from the user, his/her devices and the sensors present in his/her environment. A credibility index is built for each user, and this information is stored in the credibility index store. Next, we present the architecture of our CATVal system.

4.1. CATVal System Architecture

Our proposed CATVal System architecture consists of the following 4 layers as shown in figure 4:

Context Collection Layer: This layer collects and processes all the user and his/her environmental context information such as time, location, activity, type of device, network related context (mac address, IP address, type of network, access point used) and API or manual web based tweet. The collected pool of context information is then passed on to the Tweet Source Validation Layer.

Twitter Interest-based Aggregated Content Retrieval Layer: This layer retrieves aggregated Twitter content from Twitter users based on the 'Interest' of the organization and its requirement of such content in the decision making process. This aggregated content can also be sourced via a third-party aggregator to an organization. A simple crawl on Twitter can retrieve required content from users tweeting about a product or service directly relating to the organization demanding such content via the DSS.

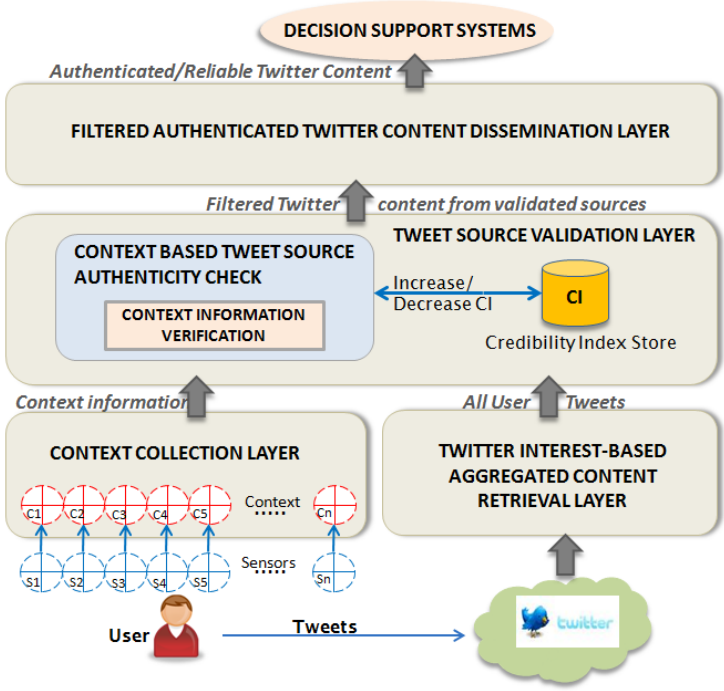

Figure 4. Layered CATVal System Architecture.

Tweet Source Validation Layer: This layer receives the context information of each Twitter user for which a validation is required or when a fake user attempts to tweet on

behalf of the real user. After checking for the credentials, i.e., the context information obtained from the user and previous record of user tweets, a credibility index is maintained in the credibility index store. The credibility index can be increased or decreased for each user depending on the validation criteria or credentials of the Twitter user being verified. After the credentials are verified and the Twitter source is validated, the particular Twitter user in question is cleared. The content originating from this user can then be passed on to the next layer which disseminates the required information.

Filtered Authenticated Content Dissemination Layer: This layer receives Twitter content from the authenticated and validated Twitter sources and contains the interface to pass on this required information to the DSS for use by the decision makers. This layer provides the required interface between the CATVal System and the DSS to forward the authenticated and reliable Twitter content to enable decision makers to make reliable and correct decisions within organizations.

5. Test-bed and Prototype Implementation

Context Collection in CATVal: We collect user and his/her environmental context using a number of sensors such as accelerometer (body motion), indoor and outdoor positioning systems (location), temperature sensor, RFID tags and reader (object interaction) and other sophisticated software based context collection threads (user's activity on devices such as laptop and mobile phone). We place a Mulle v3 sensor (which has on it an accelerometer and a temperature sensor) on the user's waist to gather information about user body motion. For body motion detection using accelerometer we use the waist of the user as the most appropriate position. If a Mulle is not available the accelerometer sensor on a mobile phone such as Android phone can be used. The Mulle sends data at 15Hz which is collected on the user's laptop or mobile phone where our system utilizes Java based Weka API to process the acceleration data. We use J48 decision tree which is an extension of C4.5 to infer three body motion related low-level activities such as sitting, standing, and walking. For object interaction, an RFID reader is attached to the user's wrist which detects RFID tagged objects the RFID readings are inferred as activities using varying time windows for different activities. The user's activity on the laptop and the mobile phone are inferred using a software-based sensor which records user's activities in terms of using different applications (adobe reader, word document, power point presentation, eclipse, etc) on the laptop. The tool also helps in URL logging to log URLs visited by the user which gives the user's browsing activity such as email, online news, Google search, Google scholar search, library website, airline booking site, Flickr photos, etc. Location information is collected using three methods, GPS (outdoor), Wi-fi positioning (Ekahau WLAN positioning is deployed on university campus office of the user), and RFID based location (at home). We use the inbuilt AGPS on an android development phone (ADP1) to gather other context such as location, speed and temperature. The RFID smart home helps in locating the user at home using multiple fixed tags. Further information on our test-bed as well as collection and detection of activity information can be found in [19, 25]. In the future, we are building a system which mainly sources context information from a user's mobile phone based on its inbuilt sensors.

Credibility Index Computation: For each tweet the user makes, a set of context information is associated to it is saved in our 'Context Store' as <UID, TweetID,

Context Parameter<location, activity, environmental context (temperature, light on/off) network context (mac, ip address), type of device, api-based tweet>> shown in figure 3. This information is used for identity validation. Further, we build a Credibility Index (CI) and increase or decrease CI depending on criteria such as timing entropy(Te), follower-to-friend ratio (FrR), bayesian text analysis (BT) of tweet content based on a number of features along with the context (C) associated with each tweet. The Credibility Index function CI = f(Te, FrR, BT, C) checks for the values of the parameters Te, FrR, BT and C to determine the value of CI which lies between 1 and 0. For each user, the credibility index is updated and passed on to the DSS as required for different interest-based aggregated content on Twitter.

A Twitter application is created to which a user needs to subscribe and give permission in order to enable the collection of his/her context information. For those users who do not subscribe to the Twitter application a basic CI is computed without context information. They are encouraged to subscribe to our Twitter application in the future to enable enhanced computation of the CI. In the future such context information can be stored with trusted third-party storage facilities which can alleviate any user concerns about trust and privacy issues. The CATVal system is built using Java and Twitter4j API.

6. Conclusion

The growing use of social networks, specifically Twitter has led organizations to use content available through social networks in the decision making process. Misleading content on Twitter can lead to negative outcomes if used in DSS by decision makers. A DSS needs to provide authenticated, dependable and credible content to decision makers along with addressing trust issues arising from any misleading or even malicious content. In this paper, we propose, develop and discuss a context-aware Twitter validator system (CATVal) to validate credibility and authenticity of Twitter content at run-time for use in decision support system. We show how context information relating to the user and his/her surroundings can be used to build a credibility index for a user based on his/her Twitter status updates. Decisions made using false information can prove costly to organizations. CATVal helps to provide authenticated content from credible Twitter users thus having significant impact on providing quality information to DSS. Further, CATVal provides decision makers information which is trustworthy by verifying at run-time the credibility index of Twitter users. The CATVal system is able to update the credibility index based on context information of a user each time he/she makes an update. Future work will include extensive evaluations by drawing lessons from the gained experiences.

References

[1] M. Jankovic, "Prise des décisions collaboratives dans le processus de conception de nouveaux produits. Application à l'automobile.," Paris: Ecole Centrale Paris, 2006.

[2] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, present, and future of decision support technology," *Decision Support Systems,* vol. 33, pp. 111-126, 2002.

[3] S. B. Shum, L. Cannavacciuolo, A. De Liddo, L. Iandoli, and I. Quinto, "Using Social Network Analysis to Support Collective Decision-Making Process," *International Journal of Decision Support System Technology,* vol. 3, pp. 15-31, 2011.

[4] M. Cheong and V. Lee, "Integrating web-based intelligence retrieval and decision-making from the twitter trends knowledge base," in *Proceedings of the 2nd ACM Workshop on Social Web Search and Mining* Hong Kong, China: ACM, 2009.

[5] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who is tweeting on Twitter: human, bot, or cyborg?," in *Proceedings of the 26th Annual Computer Security Applications Conference* Austin, Texas: ACM, 2010.

[6] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proceedings of the 26th Annual Computer Security Applications Conference* Austin, Texas: ACM, 2010, pp. 1--9.

[7] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting Spammers on Twitter," in *Seventh Annual Conference on Collaboration, Electronic messaging, AntiAbuse and Spam* Redmond, USA, 2010.

[8] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: understanding microblogging usage and communities," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis* San Jose, California: ACM, 2007, pp. 56--65.

[9] D. Zhao and M. B. Rosson, "How and why people Twitter: the role that micro-blogging plays in informal communication at work," in *Proceedings of the ACM International Conference on Supporting Group Work* Sanibel Island, Florida, USA: ACM, 2009, pp. 243--252.

[10] B. Krishnamurthy, P. Gill, and M. Arlitt, "A few chirps about twitter," in *Proceedings of the 1st Workshop on Online Social Networks* Seattle, WA, USA: ACM, 2008, pp. 19--24.

[11] C. Paris, P. Thomas, and S. Wan, "Differences in Language and Style Between Two Social Media Communities," in *Proceedings of the International Conference on Weblogs and Social Media (to appear)*, Dublin, 2012.

[12] E. Mischaud, "Twitter: Expressions of the Whole Self: An investigation into user appropriation of a web-based communications platform," London School of Economics and Political Science, 2007.

[13] PearAnalytics, "Twitter Study," August 2009.

[14] Semiocast, "Brazil becomes 2nd country on Twitter, Japan 3rd, Netherlands most active country," 2012, (Access Date: 01 Feb 2012).

[15] Twopcharts, "http://twopcharts.com/twitter500million.php," 2012, (Access Date: 01 Feb 2012).

[16] "TweetDeck, http://www.tweetdeck.com/," (Access Date: 01 Feb 2012).

[17] "Twitteriffic, http://twitterrific.com/," (Access Date: 01 Feb 2012).

[18] "StatusShuffle, apps.facebook.com/status-shuffle/," (Access Date: 01 Feb 2012).

[19] Saguna, A. Zaslavsky, and D. Chakraborty, "CrysP: Multi-Faceted Activity-Infused Presence in Emerging Social Networks," *Smart Spaces and Next Generation Wired/Wireless Networking,* vol. 6294, pp. 50-61, 2010.

[20] "SproutSocial, http://sproutsocial.com/," (Access Date: 01 Feb 2012).

[21] "TweetAdder, http://www.tweetadder.com/," (Access Date: 01 Feb 2012).

[22] M. Cheong and V. C. Lee, "A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter," *Information Systems Frontiers,* vol. 13, pp. 45-59, 2011.

[23] Y. Lu and D. Yang, "Information exchange in virtual communities under extreme disaster conditions," *Decision Support Systems,* vol. 50, pp. 529-538, 2011.

[24] A. L. Hughes and L. Palen, "Twitter adoption and use in mass convergence and emergency events," *International Journal of Emergency Management,* vol. 6, pp. 248-260, 2009.

[25] Saguna, A. Zaslavsky, and D. Chakraborty, "Complex Activity Recognition Using Context Driven Activity Theory in Home Environments," *Smart Spaces and Next Generation Wired/Wireless Networking,* vol. 6869, pp. 38-50, 2011.