18th ICCRTS

"C2 in Underdeveloped, Degraded and Denied Operational Environments."

Twitter as a Source for Actionable Intelligence

Paper number: 081

Topic 3: Data, Information and Knowledge

Bruce Forrester
Defence R&D Canada – Valcartier
2459 Pie-XI North
Quebec, QC, G3J 1X5
Tel.: (418) 844-4000 #4943
Bruce.Forrester@drdc-rddc.gc.ca

Abstract

During the recent uprisings in the MENA countries (Middle East North Africa), international media were afforded very limited access [1], hence social media (SM) was often used by media agencies to provide information on internal events. We know that SM played a role in communications for both the local population [2] and international media. It seems logical to believe that such information could be used by intelligence agencies to help overcome such degraded information flow, however this source was not used by most NATO countries. We are still not sure of the intelligence uses of SM nor its validity as a source. SM is produced by hundreds of millions of people around the world daily on thousands of SM sites; just one site, Twitter, with over 500 million active users as of 2012, generates over 340 million tweets and 1.6 billion search queries per day (http://en.wikipedia.org/wiki/Twitter). There is an ever-growing industry that monitors and analyses social media (SM) content for business [3]. However, these tools are limited in terms of capability and utility for intelligence work [4]. Hence, we must look further afield to find the types of required analysis. Research of real time monitoring of streaming data has increased since the advent of Twitter [5]. This paper presents the views of OSINT specialists from nine different countries on how SM could be used to inform intelligence products. The paper focuses on how data from one such source - Twitter - has been used and analysed by academia and highlights the most interesting results that are applicable in an intelligence context. Can similar analysis be used to produce actionable intelligence? We will examine philosophical and methodological research issues concerning large data sets such as Twitter data and explore studies using Twitter data and highlight promising metrics and methods. Discussion will focus on the interesting traits that seem to have potential benefits for intelligence analysis.

I Introduction

> *Populations in non-democratic states will increasingly employ social media tools in pursuit of democratic agendas. However, these governments are and will continue to develop more nuanced, insidious and effective mechanisms for exploiting social media while maintaining already pervasive control over traditional media sources. For these reasons, this analysis recommends that the Intelligence Community increase its attention to developing tools to observe, measure and report on the complex and evolving use of social media both by citizens and governments in largely closed societies.[6]*

Such is the conclusion of an article produced for the Office of the Director of National Intelligence's 2010 Summer Hard Problem Program. A similar conclusion was also reached by an international group of scientists and OSINT practitioners looking into the intelligence uses of social media for NATO [7]. However with hundreds of millions of people around the world using social media daily, navigating this ocean of activity is complicated. For instance, just one site Twitter, with over 500 million active users as of 2012, generates over 340 million tweets and 1.6 billion search queries per day (http://en.wikipedia.org/wiki/Twitter). It is clear that we need to better understand how we can tap into this wealth of data and information.

During the Arab Spring there was a significant rise in the volume of tweets. The first conclusive report of the role of social media during the Arab Spring states "Over the course of a week before Mubarak's resignation, the total rate of tweets from Egypt —and around the world — about political change in that country ballooned from 2,300 a day to 230,000 a day. Interestingly, the relative contribution of people not living in the region diminished significantly over this period" [2]. During this same period, the concerned governments were frantically trying to shut down access to the sites and were arresting identified social media activists [2]. It was also the first times that major news agencies significantly increased their use of reports and videos that were produced by the local populations within these countries. Such amateur and mostly unconfirmed sources are commonplace in today's mass media, used for both speed of reporting but also where media access is limited or prohibited [1]. So it seems promising that social media in general, and Twitter more specifically, could be used to help understand populations and governments in countries of interest – a novel sensor for instability[8].

There is a degree of maturity in the non-military applications of analytical methods, used by industry that monitors and analyses SM content [3], which can be used as a basis for the development of specific algorithms for the intelligence community. However there are some distinct differences between the military and civilian target populations that will require research and modification of current algorithmic models. This paper will discuss research that has been conducted using twitter for such purposes as election prediction [9-11], finding influential users [12-14], determining how information flows within the network [5, 15-18], and earlier work on producing meaningful metrics [13, 19-21].

The rest of the paper is organised as follows. Section II will examine philosophical and methodological research issues concerning large data sets such as Twitter data. Section III will explore studies using Twitter data and highlight promising metrics and methods. Section IV will look at the data from scientists and OSINT practitioners on the potential uses of social media. Section V will discuss the interesting traits that seem to have potential benefits for intelligence analysis. And finally section VI will conclude and discuss future work.

II Research Issues surrounding Social Media data

Black et al. [22] state “There are recognized empirical and theoretical gaps in the application of social science theories to raw, electronic data like that retrievable from Twitter”. This section begins with a short discussion on some of the major philosophical and methodological issues surrounding large collections of social data. It is followed by a specific discussion on twitter related issues.

Philosophical Issues

The basis for accepted intelligence analysis is data and information gathered from multiple sources. Social media can be thought of as potentially having an unlimited number of signals that could be analysed depending on what answers are being sought. In traditional experimental research, data collection is conducted in a predetermined manner and is chosen based on the experimental design using a set of hypothesis. Data is purposefully chosen. We can think of this as a “goal-driven” method. This method ruled when we lived in a data deprived world. We are now overloaded with data and are preoccupied by how to filter to leave only relevant data. The methods of text and data mining [23] have exploded, as have statistical and correlational analysis due to improved processing power and techniques and due to the increased availability of data sets online. These methods are excellent for finding trends and interesting patterns in huge data that would normally be hidden. However, we must remember that “running the numbers and finding the correlations will never be enough” [24]. This “data-driven” or inductive method will provide an initial model that will need to be developed into a theory and verified through more rigorous scientific methods. Detecting trends or patterns from massive data in order to analyze possible underpinning relationships is not enough. It has been recognized that “a record of interaction through technology does not necessarily act as a proxy for social interaction” [22].

Data cleansing presents another danger with large data sets. Researchers will tend to remove data that they believe is erroneous or determined to be outliers and because there is so much data available, the samples remain statistically significant. However, outliers often provide just the indication for which intelligence analysts are looking. “Jeff Jonas of the IBM Software Group believes that ‘bad data’ is good for you. You *want* to see that natural variability. You *want* to support dissent and disagreement in the numbers. There is no such thing as a single version of the truth. And as you assemble and correlate data, you have to let new observations change your mind about earlier assertions” [24]. Cleaning and filtering (clustering, grouping, etc.) data also presents the problem of possible breaking of patterns that are larger than the clustering.

Finally, there are also concerns with the aggregation of large amounts of social data. There are hundreds of social media sites and at least 28 different categories of conversation types or purposes [25]. The types of analysis services and tools looked at in [3] used data that had been aggregated from multiple sources without necessarily considering the purpose for which postings (tweets, updates, blog entries etc.) were made. It is highly probably that individual postings were made within a certain context for a certain audience. In their aggregation with other similar topics, which potentially are centred on different time periods or audiences, interpretation and contextual errors easily occur [26].

Methodological Issues

There are a multitude of issues surrounding twitter data and the following should not be taken as exhaustive.

The first issue most researchers in this area need to address is how to collect the data. As Twittter data is assumed to be public by default [19], we should be able to do what we want with it. In the early days of Twitter, access to all data was available and researchers were able to collect and store this data for analysis [27, 28]. Restrictions to the amount of data that can be collected at any one time have since been put in place. There are currently three main application programming interfaces (API) for collecting Twitter data: the REST, search, and streaming methods. Unfortunately, these APIs are not clearly defined and are susceptible to change at any time [22]. Hence, each of these API methods is likely to produce a different set of data and researchers to date have not been clear on why they chose a particular API. In the implementation of these APIs, Twitter has in effect significantly limited access to the “fire hose” of data. This complicates an analyst’s ability to discover the entire dataset. It also makes the issue of capturing data on a suitable target population more difficult.

Next, how does one go about defining the required target population? At the out start, we can say that only a percentage of the population uses the internet. Of that, a subset use social media and of that a smaller subset use twitter [2]. “Thus random sampling on social media is biased sampling” [11]. This may not be a problem if one is looking for the opinion of the young tech-savvy middle-class. For more balanced results, analysts will require the ability to determine the significance of the population sample answering the questions: how well does the group analysed match the overall population of interest? “It should be noted that social media does not reflect the demographics of the society” [11]. Target populations need to be stratified in order to be capable of determining what different groups are saying. Such details are important to intelligence work.

There is a vast multitude of different ways that researchers use to retrieve, clean, store, and analyse Twitter data [22]. While most research referenced in this paper provided detailed descriptions of how they determined their data set, no one method was the same. For instance, some researchers used a single tweet as the unit of analysis while others combined all tweets from an individual user. Each analysis required a different determination of the population sample. As such, one needs to be very careful in the application of the algorithms presented in the research. A consequence of these non-standardized methods is a difficulty in coherently make comparisons between results in any meaningful way.

Social media sites are in a constant beta mode of development which means that new features are frequently being introduced. Some features will pass the test of time while others will be abandoned based on user usage and feedback. This lack of steady-state conditions presents a challenge for researchers trying to compare or build upon past studies. In addition, the number of users is constantly changing. For example, a 2009 study [29] captured twitter data on the top four trending topics over a two day period and obtained 7215 total tweets. As of March 21, 2012 according to Wikipedia, there were over 340 million tweets a day, and as the New Year began on January 1, 2013 there were 33,388 tweets per second in the Japan standard time zone alone. This issue warns researchers to be very careful when making comparisons or when trying to duplicate methods that were valid on relatively small data sets but are now being applied to massive amounts of data.

The above philosophical and methodological issues make it hard to combine studies in this area in order to produce more comprehensive theories. However, we can certainly learn from previous studies being cognisant of the problems.

III What attributes have been studied in the past?

Studies of users' traits

At the heart of Twitter is a 140-character content space, which for many, may seem extremely limiting. However, Twitter provides a much faster means of communication than regular blogs. Shorter posts, and the corresponding time needed to read, encourages greater use and frequency of updates (every few days for a regular blog compared to several times per day for twitter) [8]. We have also seen the utility of this speed in earthquake and disaster relief efforts [11, 12]. In fact, much of the content of tweets is centered on the present and its value is ephemeral [13]. Kawk et al. [18] found that half of retweeting occurs with an hour and 75% within a day. A 2009 study reported that up to 40% of tweets could be classified as noise (spam)[15]. This causes a very low signal to noise ratio. However, 2009 was still early days for Twitter and current research is required to determine if this is still the case. Recently, social media sites have started to enable a much wider set of functionality, which has likely led to greater relevance. For Twitter, users can now easily add links to other sites and sources to enrich their tweets. As well, the ability to tweet from within other sites is now wide spread, as is the ability for mobile tweeting. This has led to affordances not envisioned by the creators. It has also complicated (or enriched) the types of analysis that can be conducted. So, can we really extract pertinent information from social media sources that leads to actionable intelligence?

It seems logical that if these social tools are being used to communicate we need to understand what types of communication are occurring and perhaps the motivations of users before we can understand the potential intelligence uses. An early study of twitter [8] found that people use this microblogging environment for three main reasons; information sharing, information seeking, and friendship-wise relationships. Cheong & Lee [29] found that, apart from individuals, Twitter was used by groups (non-profit or researchers), aggregators (publishing or collating info), marketing (to push a product or for spam) and for satire (humorous, satirical or parodying purposes). More recent studies have found that people with similar life outlooks and interests tend to “hang out” together [9], to talk about headline news and respond to fresh news[18], and that there is a high degree of homophily and following due to interest in similar topics [14]. This means that we should be looking at using social media to help us answer questions that are of a social nature and that change over time or are time sensitive. Social media has become an effective communication tool for organizers of events of a social nature. At the same time, due to the open nature, this communication becomes an accessible source.

Rao and Yarowsky [21] looked at the ability to detect latent user-properties within social media. While most user profiles ask users to describe certain personal attributes, provision of such information is optional. Rao and Yarowsky used the contents of tweets and posting behavior to help classify users by gender, age (above or below 30 years old), regional origin and political orientation (US only – Liberal, left or Democratic leanings) where no such information was explicitly provided. For instance they found that the presence of a sequence of exclamation marks is indicative of a female user and that women laugh with “LOL” while men use “LMFAO”. They built models (sociolinguistic-feature, Ngram-feature, and stacked) as binary classifiers using Support Vector Machines to determine which model worked best. They found that the staked model was best for gender and age, N-gram for region and the sociolinguistic model for political orientation. They determined accuracy between 72 and 83 percent.

Studies on influence

Google was the first to solve the issue of finding the most relevant sites on the internet through their PageRank algorithm that uses weighting of the hyperlinks between pages to determine their relative importance. As any Google user will attest, it does a pretty good job at finding relevant pages for one's search terms. However, the problem changes when one is trying to determine which user is the most influential, say, in a particular conversational thread. PageRank type analysis ignores the interests of the users and simply using the *indegree* (the number of people who follow a user) is not sufficiently granular to determine influence, and in fact reveals very little about the influence of the user [27]. Users can follow others and can be followed by others independently and without permission. Weng et al. [14] reports that there is a high reciprocity of user-follower relationships. So how do we go about identifying what to measure? The challenge of identifying the most influential users is discussed below.

Cha et al. [27] added the *retweet* (mean number of times other followers "forward" a user's tweet – implemented by `RT @username` or `via @username`) and *mentions* (mean number of times others mention a user's name – implemented by `@username` within the text of a tweet. If a tweet starts with `@username` it is only broadcast to that user as a private tweet not like a tweet to all users). They then compared all three measures of influence across three of the most popular topics in 2009. Topics can often be easily identified through the use of the #hashtag. Indegree influence can be thought of as the size of the user's audience. Retweet influence provides a metric on how well a user produces content that has pass along value. Mention influence is an indication of the ability of a user to engage others in a conversation.

Cha et al. spent a lot of effort in determining their target population which was typical of the majority of researchers in this area. Their final sample size was only two percent of their original population, but due to the huge numbers, it still contained 13,219 users. They investigated a user's influence across topics, the rise and fall over time, and how influence is maintained. They also looked at how an ordinary user's influence could rise. They found that top users required a concerted effort to gain and hold influence. Users like CNN consistently produced high value tweets that were often retweeted, while celebrities garnered mentions due to their name value. Ordinary users were able to gain influence by sticking to one topic (usually information about protests or controversial news) and posting interesting and insightful tweets, but their influence soon waned in relation once interest in the topic died out.

In another study, Weng et al. [14] also looked at the problem of finding influential users. They used an extension of PageRank called TwitterRank which adds topical similarity between users to the link analysis. In order to distinguish individual topics, Latent Dirichlet Allocation was used. LDA if a form of unsupervised machine learning and uses probability distribution to determine the value of overall vectors of word counts. As such, the totality of users' tweets had to be used in order to gain significance. Again, substantial data preparation was needed. They removed all words that they felt were not useful for topic forming including non-English words. They performed a correlational analysis on indegree plus the two ranking methods. They found that TwitterRank outperformed the other two in measuring the topic-sensitive influence of users, but they also have identified ways to improve this metric. They were also the first researchers to identify the behaviour of homophiliy; a phenomenon showing the homogeneous nature of peoples' social networks with regard to many sociodemographic, behavioral and intrapersonal characteristics [14]. So birds of a feather really do flock together in social media.

Suh et al. [13] looked at both content and contextual features of retweets (number of followers & followees, age of account, number of favorited tweets, and the number and frequency of tweets) in order to try to understand the factors that might affect the retweetability of a tweet. They used a 74 million set of tweets that constituted about 2- 3% of all tweets for the period of time when their data set was collected. Their method included a factor analysis using a principal components analysis followed by the production of a generalized linear model. For the content features, they found that inclusion of a URL has a significant impact on retweetability as did the domain of the URL. As well, the inclusion of a hashtag correlated with a retweet. For the context features, they found that number of followers and followees plus the age of the account affected retweets. Interestingly, they found no indication that past tweets (total number of tweets posted since account inception) factored into retweet rate [13].

Virality of tweets, through the measure of the retweet, was studied by Hansen et al. [30]. They set out to determine how sentiment (positive or negative) and content (news or not news) affected the virality of tweets. They used three corpora of which one was used to train a Naïve Bayes news classifier. Sentiment analysis used a list of 1,446 words with a valence between -5 and +5 (good separation compared to most sentiment analysis methods). The results showed that the news classifier had a high accuracy of 84% and that 23% of their "random" tweets were news. They concluded that negative sentiment surrounding news content promoted virality (the probability of retweet) but not for non-news content [30].

The ability to identify communication roles has been researched. In the early days of Twitter, a 2007 study [19], users on twitter basically would talk about their daily activities, and share or seek information. More recently Tinati et al. [15] used dynamic behavior method to classify users according to Edelman`s *topology of influence*. The five roles are: 1) Idea starter - An individual who starts a conversation meme, 2) Amplifier - An individual who collates multiple thoughts and shares ideas and opinions, 3) Curator - An individual who use a broader context to define ideas, 4) Commentator - An individual who detail and refine ideas, and 5) Viewer - An individual who takes passive interest in the conversation. They used the retweet, hashtag usage, and tweet timestamps in their analysis and made two important conclusions - users can be classified into these roles based on retweets and that comparison to real world roles of the users was logical.

Studies on Validity of content

Yang et al. [31] studied users retweeting behavior. If we agree that a retweet is an indicator of value as suggested [27], then analysis of such behavior could provide an interesting metric to help detect deception. Through a series of experiments, they looked at tweeting and retweeting activity as well as the importance of the content and the interest of that content to the user. This research has helped to identify the factors that influence the likelihood of a retweet (user, message, time). Also important is the use of a semi-supervised framework that allowed for a 29% prediction precision of retweet behavior [31].

Research to determine the reliability of tweeted information following a disaster was conducted by Mendoza et al. [16]. They studied propagation of information during the Chilean earthquake (Feb 27, 2012) comparing rumors and news. They used a retweet as an indication of importance or relevance of the original tweet. Through content analysis, they determined that a collaborative filtering affect differentiated news from rumors. Rumors were questioned by the community to a much larger extent than confirmed truths. This is a very interesting finding and one that could be utilized to help find deception.

Studies on Prediction

Asur & Huberman [32] studied the use of twitter to forecast box-office revenues for movies. They looked over a three month period at 24 movies released on Fridays with a wide circulation and collected 2.89 million tweets from 1.2 million users. Using attention, popularity and sentiment analysis to construct a linear regression model, their results outperformed the traditional Hollywood Stock Exchange predictions. They conclude that "this work shows how social media expresses a collective wisdom which, when properly tapped, can yield an extremely powerful and accurate indicator of future outcomes"[32]. More research is required to determine how this might transfer to other domains.

Yu & Kak [11] conducted a meta-study of realms that are currently being predicted using social media. Of course, prediction is limited to human related events. They covered marketing, movie box-office, information dissemination, elections and macroeconomics. They found that social media has some effect on all these areas but in general prediction using social media had relatively low accuracy due in large part to the prediction factors and methods used thus far in the research. Yu & Kak [11] do believe that improvements will be forthcoming as research matures in this area. Interestingly, they found a plus for using Twitter because of its short cycle length. Prediction in general is more accurate on contents with a short life cycle compared to that with long life cycle.

Gayo-Avello [9] also conducted a meta-analysis of 17 studies but concentrated on studies involving the prediction of elections. He concluded that it is just not possible to predict based on the methods and algorithms that were used by the studies. He found scores of methodological problems. First, the studies were not predicting, they were all post-hoc analysis. Second, there was no commonly used way to "count votes". Third he stated that the sentiment analysis used was applied as a black-box and with naiveté. Fourth, the studies ignored the presence of rumors, propaganda, and misleading information. Fifth, they neglected demographics and the representation of age, gender, social groups within Twitter. Finally, self-selection bias was simply ignored [9]. So, predicting elections at a granular level is hard using Twitter.

Choy et al. [10] looked at prediction and the 2012 US Presidential election. They collected over seven million tweets over a certain period relevant to this election and they used the AFINN sentiment analysis list and improved upon a previously used model to try and address issues raised by Gayo-Avello [9]. They determined their improved model to be far more accurate than the previous model and that the sentiment reflected in twitter acts as a good barometer of the electorate's opinion of the candidates [10].

Bollen et al. [33] looked at how the Dow Jones Industrial Average (DJIA) correlate to daily twitter mood as measured by OpinionFinder and Google-Profile of Mood States (GPMS). They collected nearly 10 million tweets from 2.7 million users and only used tweets that were explicit in their statements of mood states. They found that public mood could indeed be determined using fairly simple techniques and that of the seven moods indicators (calm, alert, sure, vital, kind, happy, and that generated by OpinionFinder) the calmness as measured by GPMS was the best indicator of the DJIA. The mood correlated to the index but was shifted three to four day later. They highlighted some important areas of noise and bias that would affect more granular assessments of specific markets and will address these issues in future research.

Studies Concerning Language

Users of social media in general and Twitter in particular, due to its limited character space, have developed a special abbreviated language. Besides the obvious translation of foreign language problems, there is a real mash-up of language characteristic; the shortest, most easy way to get one's message across seems to be the maxim. Colbath & Srivastava [34] present good examples of how tweeters often mix of numbers, languages (English Arabic), jargon and even the use of how numbers and letters sound when mixed (such as "u r" for "you are" or "l8tr" for "later") . In addition, there are often cultural references added into this slurry of language styles. Many of the differences compared to journalistic text were described in [35]. As noted above in [21], researchers have used the fact that there are many idiosyncrasies in the use of language in tweets in order to find latent user traits. As a result analysts will need a significant level of language, cultural and current event awareness in order to understand and interpret individual tweets. Further, algorithms will require regular updates or will need to be built to constantly flag new language uses.

Misspelling and sentence structure can often be used as a latent indication of education level or interest [36]. However, it is not clear how this might apply to the very compressed nature of tweets and how it is affected by the size of mobile device keyboards or auto correction software.

Studies of Other Indicators

Estabrooke and Combs [8] developed a framework for characterizing the dimensions of volume, temporal change, and substance within social media. For volume, they state the need for a baseline and then monitoring for fluctuations in volume (posts, tweets, etc.). Generally, one could expect a fluctuation that correlates with membership and that any spikes in volume would be an indicator of an event of interest. For temporal changes, they establish patterns of life baselines for groups of interest. Finally, with substance, they point out some of the difficulties and mention that sentiment coding remains a very imprecise science [8].

Sentiment analysis was studied by Finn Nielsen [37] who compared the performance of several different word lists and scoring methods. There are several methods currently being used for assigning sentiment strength to words. Nielsen used +5 to -5, with most words receiving a 2 rating. Strong obscene words get either -4 or -5. His list (AFINN) contains 1468 words and a few phrases and tended to be biased towards negative words (65%) which is similar to OpinionFinder (4911 words with 64% negative). His results showed that AFINN and SentiStrength outperformed larger word lists such as ANEW, General Inquirer, and OpinionFinder. However, he recommends ANEW (Affective Norms for English Words) as it has been validated across several studies [37].

Geolocation is an important characteristic for intelligence work. However, it is not automatic in Twitter, where users must self-identify. Kawk et al. [18] state that “it is hard to parse for location due to its free form” and they considered time zone as an approximation for location. As more and more cellular phones come equipped with GPS, many people will likely opt into this feature. There are other ways to detect approximate location through various combinations of pattern of life analysis (active period trends), using contextual clues in user content, comparing time related content to timestamps, use of language, or if a user has several different social media accounts, information from another account might provide location [29]. During the Arab Spring, Howard et al. [2] were able to determine the percentage of tweets coming from inside Tunisia, from the neighboring countries and those from outside the region. This is an area that is vital to the intelligence domain especially in the application of kinetic force and one that is likely to receive much more attention in the future.

Deception is a problem that is present in most if not all intelligence domains. First and foremost, it is hard to find out who it is that is really who online. Luckily there are particular forms of deception in the online environment that one is able to detect using algorithms. Chen et al. [38] examined the "internet water army" or online paid posters in China. These hired help get paid for posting comments for some hidden purpose and are usually paid based on the number of posts. Chen et al. found that these paid posters have some special behavioural patterns that allow detection through statistical analysis. These patterns included percentage of replies, average interval time between posts, the number of days the user remains active, etc. They also found that user IDs were often shared and one could detect the use of the same ID in different geographical locations within a very short time period or that there were large numbers of IDs created in a short time. Lastly, Chen et al. [38] found that to save time, paid posters often copied posts and just slightly changed them hence leading to detection through semantic analysis. This is another research area that will require greater focus.

Examples of Tools

To finish this section, a few tools were reviewed. For a more comprehensive review of general tools and services used for social media see [3]. Black et al. [22] conducted a comprehensive review of twitter research concluding that there are significant methodological problems. Hence, they have defined a method and architecture for capturing, social transforming and analyzing the Twittersphere called Twitter Zombie. They use the Twitter search API with a cron scheduler to gather tweets that are then stored in a MySQL database. Tweets are then socially transformed into a representation of interactions in the form of node pairs. These pairs are in turn easily visualized and statistically analysed. Their research was completed with the goal of providing a common method and tool that can be used by researchers in the area of social media exploitation [22]. This may prove an interesting starting point for eventual standardization.

Byun et al[28] also provide guidance on an architecture for collecting and analysing tweets. They go into some detail on the design and functionality as well as how to circumvent the data restrictions imposed by Twitter APIs. Their tool was successfully used to gather data about the 2012 Super Bowl commercials.

Colbath and Srivastava [34] discuss the problem of language translation and the fact that most tweets include colloquialisms, dialect words, errors of syntax, etc. that greatly inhibit the use of formal language translators. They go on to describe a system, developed by BBN, called MAGPIE that allows for: 1) harvesting, translation and storage, 2) tracking of emerging topics and themes, 3) cross-platform identity and topic correlation, 4) language and dialect identification, 5) sentiment analysis, and 6) network visualization. This system is still underdevelopment but based on its described functionality will be worth further investigation.

An interesting project from the MITRE Cooperation is described by Costa& Boiney [39], and Maybury [40]. They recognize the need for social radar in order to “detect breakpoints that signal major sentiment shift likely to have effects on the behavior of populations or governments” akin to the how traditional radar collects data on the physical world. This vision acknowledges the need for understanding the links between sentiment, motivation and behaviours, in other words cultural models. They state ongoing research in the following areas: 1) sentiment analysis and topic discovery, 2) ideology identification in multiple languages, 3) emotion analysis of social media for instability monitoring, 4) automated breakpoints for change detection from online sources, 5) mapping influence via online posting, 6) cluster analysis, ranking, and exploration for online postings, and 7) course of action analysis using nation-state simulation models. This is a hefty research agenda that will surely help to advance to science of social media analytics.

IV What the experts say about potential intelligence uses

This section was derived from three international meetings involving scientists and OSINT practitioners from nine different NATO countries [7]. The basic question asked was: How could you envision using social media information and data from open sources? Table 1 summarises the answers.

Table 1 – Potential uses of Social Media Sources for Intelligence

Phenomenon	**Military/Intelligence impact**	**Intelligence Product**
1. Potential social uprisings: • What is the stability of current government in country? • What are the issues with the people? • Are things escalating? • What are the trends?	• Strategic and Operational • Contingency plans • Operational Plans • Peacekeeping	• Early warning and indicators • Trend watch • Response to standing RFI • Alerting service • Basic intelligence (baseline) • Threat assessment • Country studies

Phenomenon	Military/Intelligence impact	Intelligence Product
2. What is happening in remote areas where there are few other sources of information available?	• Current up-to-the-minute SA (situational awareness) of a particular area • Enables operational planning • Tactical threat assessment	• Response to targeted RFI • Alert service • Threat assessment • Information bulletin
3. Monitoring and pattern analysis looking for criminal / terrorist / insurgent activities.	• Cyber issues (taking down subversive sites) • Targeting • Understanding ECOA • Planning (collection, ops, tactical etc.) • Disrupting the insurgency cycle before the ACT stage	• Response to RFI • Threat assessment • Standing products
4. Targeting (non-kinetic) (i.e. profiling); identifying and getting information about particular person of interest, groups, organizations.	• Targeting • Understanding ECOA –Enemy Courses of Action • Planning (collection, ops, tactical etc.) • Understanding the ideology	• Response to RFI • Threat assessment • Profile • Structure of orgs

V Discussion

Social media by its online nature automatically comes with a rich set of metadata available for analysis. To date this researcher judges that the majority of studies have been based on mathematical or statistical analysis of this metadata and on simple manifest content. There have been significantly fewer studies of the latent characteristics and deeper meaning of content. This is probably due to the much higher level of difficulty and the requirement for a deep cultural and linguistic understanding of the population under study. Given that there is potentially the entire world's population to consider, deep content analysis will continue to be a research challenge for years to come.

Luckily machine learning and machine translation are increasing in maturity and sophistication compared to just 10 years ago. Computational power continues to increase based on Moore's law. Further, large internet companies like Google and Amazon have pioneered open source software, such as the Hadoop project and Mahout, which are helping to solve the big data challenges surrounding social media data. So the future looks bright.

Based on the studies reviewed in section III, the following can be inferred as relevant for intelligence:

1. There is a large base of users from which data and information can be acquired.
2. Users can be identified for the most part to a fair degree of granularity.
3. We can determine the specific roles played by users within a topic.
4. There are methods that can be used to determine the most influential users.
5. We can determine where an idea started, and how and by whom it propagated throughout the network.
6. Ordinary users can have a major influence on the spread of news and information. These individuals can easily be identified.
7. There is a wealth of information that can be gleamed from tweets through the included references to links and URLs.
8. The results of prediction using twitter are varied but promising and are likely to improve as methods improve.
9. A rough degree of sentiment can be calculated.
10. Language and culture are and will continue to be significant barriers to understanding content on a deep level.
11. A geolocation can be found or inferred in many cases.
12. There are several methods that could be used to help detect deception and rumors.

Given this, there appears to be a suitable base of research (albeit in a non-military context) that will allow for the development of methods for both monitoring and mining for trends (data-driven analysis) and for research at a finer level of granularity (goal-driven research – requests for information). An earlier study [41] identified and analyzed 21 likely SM categories and 150 commercial products and concluded that 6 categories and 50 products stood out for their potential use for insurgency activities. There is certainly a much greater scope to exploit other social media than just Twitter which is the focus of this paper.

As an example, here are some sample questions concerning a potential social uprising that could likely be answered at least in part by social media:

- What is happening in country X?
- What is the population of country X talking about online?
- What are the hot topics?
- How are the grassroots discussions different from the mainstream media or government discussion?
- What are the issues that other countries are discussing with respect to the country X?
- What is the sentiment of the discussions?
- Is sentiment changing? - getting more aggressive or passive?
- What opposition exits?
- What organizations are acting within country X?
- Is there any religious polarization?
- Who are the thought leaders that are emerging in discussions?
- What is the rest of the world saying about country X?
- Where is the discourse community? – Who are they?

VI Conclusion and Future Work

As a starting point, this research took as a goal to help determine the exploitability of a single social media source.   The framework was the phenomena identified in Table 1. This preliminary work by scientists and OSINT practitioners identified relevant areas that could be used to understand potential social uprisings (what is the stability of the current government in a country; what are the issues with the people; what things are escalating or trending), could allow for a window into remote and isolated areas (where few other sources are available), and could be used in identifying and gathering information about particular persons of interest, groups or organizations. The studies looked at in this paper reflect these areas of possibility.

The majority of the studies found and read for this paper uses mathematical and statistical models. In fact, overwhelming the studies were conducted by researchers in the computer or information sciences.   As a result, many studies concentrated on the metadata associated with tweets and did not touch the content in any culturally significant way. Indeed at the heart of social media are people who are culturally bounded and embedded.   This fact shouts to the need for social-cultural models in order to delve deep into the understanding of social media. Ultimately a multi-disciplined team will be required for comprehensive exploitation of social media.

There is clearly existing civilian research that describes methods and techniques of interest and value and that can be adapted to intelligence work related to social media exploitation. The research as a field is still very immature and as it advances we will likely see a move towards standardization of methodologies. Several related fields, such as artificial intelligence, sentiment analysis and deep content analysis still require extensive research and the development of new methodologies and tools. As we further our understanding of the potential intelligence uses, in combination with a greater appreciation for the characteristics of social media, there exists great potential for researchers and intelligence analysts to leverage this understanding to advantage.

For intelligence work, especially related to areas where there is limited access and external communications, social media promises to be potentially rich environment for exploitation. Clearly, based on the studies identified in this paper, there are interesting tools and methods that act as a starting point. Certainly social media will help in providing input for situational awareness and will add a new understanding of current events based on the population for country reporting. Future research will need to look in much greater detail at the analysis of content and the cultural interpretation that will be required. There were very interesting techniques explored for the detection of deception and further research in this area will prove fruitful. Finally for prediction, or more likely estimation, there are several promising areas that need further exploration.

References

[1] UN. *Freedom of the Press: in the Middle East, widely curtailed and often violated*. 2012 [cited 2013 11 Jan]; Available from: http://www.ohchr.org/EN/NewsEvents/Pages/FreedomofthePressintheMiddleEast.aspx.

[2] Howard, P.N., A. Duffy, D. Freelon, M. Hussain, W. Mari, and M. Mazaid, *Opening Closed Regimes What Was the Role of Social Media During the Arab Spring?*, N.S. Foundation, Editor 2011, The Project on Information Technology and Political Islam: Washington.

[3] Labrèque, A., *Study of Social Networking Exploitation Tools*, B. Forrester, Editor 2011, Defence Research and Development Canada: Quebec City.

[4] Bruyn-Martin, L., E.-A. Filardo, and Y. DeWit, *Study of Intelligence Analytical and Collaborative Technologies and Tools: Final Report* D.R.a.D. Canada, Editor 2012, DRDC Valcartier: Quebec City.

[5] Shakarian, P. and D. Paulo. *Large Social Networks can be Targeted for Viral Marketing with Small Seed Sets*. in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM-2012)*. 2012.

[6] Helen, B. and P. Benjamin, *Stop looking for the Next Twitter Revolution*, D.o.N. Intelligence, Editor 2010.

[7] ET.BY, *Technical Activity Proposal - Intelligence Exploitation of Social Media*, 2012, NATO RTO.

[8] Estabrooke, I. and D.J.Y. Combs, *Social Media Defining the Problem: A Research Perspective*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, R. NATO, Editor 2012: Tallinn, Estonia.

[9] Gayo-Avello, D., *A Balanced Survey on Election Prediction using Twitter Data.* arXiv, 2012.

[10] Choy, M., M. Cheong, M.N. Laik, and K.P. Shung, *US Presidential Election 2012 Prediction using Census Corrected Twitter Model*, 2012.

[11] Yu, S. and S. Kak, *A Survey of Prediction Using Social Media*, 2012, Oklahoma State University: Stillwater, Oklahoma.

[12] Leavitt, A., E. Burchard, D. Fisher, and S. Gilbert, *The Influentials: New Approaches for Analyzing Influence on Twitter*, in *Web Ecology Project*2009.

[13] Bongwon, S., H. Lichan, P. Peter, and H.C. Ed, *Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network*, in *Proceedings of the 2010 IEEE Second International Conference on Social Computing %@ 978-0-7695-4211-9*2010, IEEE Computer Society. p. 177-184.

[14] Jianshu, W., L. Ee-Peng, J. Jing, and H. Qi, *TwitterRank: finding topic-sensitive influential twitterers*, in *Proceedings of the third ACM international conference on Web search and data mining %@ 978-1-60558-889-6*2010, ACM: New York, New York, USA. p. 261-270.

[15] Tinati, R., L. Carr, W. Hall, and J. Bentwood. *Identifying Communicator Roles in Twitter*. in *WWW2012 - MSND'12 Workshop*. 2012. Lyon, France.

[16] Medoza, M., B. Poblete, and C. Castillo, *Twitter Under Crisis: Can we trust what we RT?*, in *1st Workshop on Social Media Analytics (SOMA'10)*2012: Washington, DC.

[17] Fink, C., J. Kopecky, and N. Bos, *Evaluating Social Media as a Source of Public Opinion in the Developing World*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, N. RTO, Editor 2012, RTO NATO: Tallinn, Estonia.

[18] Haewoon, K., L. Changhyun, P. Hosung, and M. Sue, *What is Twitter, a social network or a news media?*, in *Proceedings of the 19th international conference on World wide web %@ 978-1-60558-799-8*2010, ACM: Raleigh, North Carolina, USA. p. 591-600.

[19] Java, A., 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*2007, ACM: San Jose, California. p. 56-65.

[20] Asur, S. and B.A. Huberman, *Predicting the Future with Social Media*, 2009, Social Computing Lab HP Labs: Palo Alto.

[21] Rao, D. and D. Yarowsky, *Detecting Latent User Properties in Social Media*, 2009.

[22] Black, A., C. Mascaro, M. Gallagher, and S. Goggins, *Twitter Zombie: Architecture for Capturing, Socially Transforming and Analyzing the Twittersphere*, in *GROUP`12*2012: Sanibel Island, Florida.

[23] Miner, G., D. Delen, J. Elder, A. Fast, and R. Nisbet, *The Seven Practice Area of Text Analytics*, in *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*2012, Elservier Inc.

[24] Bollier, D., *The Promise and Peril of Big Data*, C.M. Firestone and P.K. Kelly, Editors. 2010, The Aspen Institute: Washington. p. 1-55.

[25] Solis, B. and JESS3, *The conversation prism representation*, 2010.

[26] Spark, D., *Real-Time Search and Discovery of the Social Web*, S.M. Solutions, Editor 2009.

[27] Cha, M., H. Haddadi, F. Benevenuto, and K. Gummadi. *Measuring User Influence in Twitter: The Million Follower Fallacy*. in *ICWSM '10: Proceedings of international AAAI Conference on Weblogs and Social*. 2010.

[28] Byun, C., Y. Kim, H. Lee, and K.K. Kim. *Automated Twitter Data Collecting Tool and Case Study with Rule-Based Analysis*. in *iiWAS2012*. 2012. Bali, Indonesia.

[29] Cheong, M. and V. Lee, *Integrating Web-based Intelligence Retrieval and Decision-making from the Twitter Trends Knowledge Base*, in *SWSM'09*2009: Hong Kong.

[30] Hansen, L.K., A. Arvidsson, F.A. Nielsen, E. Colleoni, and M. Etter, *Good Friends, Bad News Affect and Virality in Twitter*, 2010, Danish Strategic Research Council.

[31] Yang, Z., J. Guo, K. Cai, J. Tang, J. Li, L. Zhang, and Z. Su, *Understanding retweeting behaviors in social networks*, in *Proceedings of the 19th ACM international conference on Information and knowledge management*2010, ACM: Toronto, ON, Canada. p. 1633-1636.

[32] Asur, S., & Huberman, B., . *Predicting the Future With Social Media*. 2009 [cited 2010 6 October]; 8]. Available from: http://www.hpl.hp.com/research/scl/papers/socialmedia/socialmedia.pdf.

[33] Bollen, J., H. Mao, and X.-J. Zen, *Twitter mood predicts the stock market.* arXiv, 2012.

[34] Colbath, S. and A. Srivastava, *MAGPIE: A System for Triaging and Translating Social Media*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, R. NATO, Editor 2012, RTO NATO.

[35] Forrester, B., *Social Media Exploitation Tools: Understanding Where and How to Look*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, N. RTO, Editor 2012, RTO NATO: Tallinn, Estonia.

[36] Ellison, N., R. Heino, and J. Gibbs, *Managing impressions online: Self-presentation processes in the online dating environment.* Journal of Computer-Mediated Communication, 2006. **11**(2).

[37] Nielsen, F.A., *A new ANEW: Evaluation of a word list for sentiment analysis in microblogs.* arXiv, 2011.

[38] Chen, C., K. Wu, V. Srinivasan, and X. Zhang *Battling the Internet Water Army: Detection of Hidden Paid Posters*. eprint arXiv:1111.4297, 2011.

[39] Costa, B. and J. Boiney, *Social Radar*, in *HFM-201 Specialist Meeting on Social Media: Risks and Opportunities in Military Applications*, R. NATO, Editor 2012, RTO NATO: Tallinn, Estonia.

[40] Maybury, M., *Social Radar for Smart Power*, 2010, The MITRE Corporation: Bedford, MD.

[41] Labrèque, A., *Study of social networking technologies Social networking analysis in a counter-insurgency context*, 2011, Defence R&D Canada – Valcartier: Quebec City.