Finding the needle in the social media haystack

Sam Fenwick looks at the current progress in the area of social media analysis and monitoring, drawing on discussions from PSCE Conference Oxford and other sources

A number of R&D projects are working to bring the power of social media to control room operators

Social media has become both opportunity and headache for PPDR organisations. It can provide so much data that mining it for actionable intelligence is a formidable challenge and there are many questions surrounding the best way to use it operationally and from an ethical standpoint. Returning to my initial analogy, it's a lot easier to find a needle in a haystack if you have a powerful magnet. To this end, considerable research has already been performed with the goal of bringing relevant information to control-room operators from social media during an emergency. Indeed, some systems have already been developed or are close to commercialisation.

Raúl Santos de la Cámara, project co-ordinator at HI Iberia, gave a presentation on SAMi2 – Semantics Analysis Monitor for the Illegal Use of the Internet – at Public Safety Communications Europe (PSCE) Conference Oxford. SAMi2 is a European project, which ran from January 2014 to December 2015, to create an automated tool for filtering social media so that users can quickly see relevant information for law enforcement. It uses natural language processing techniques to understand the content of social media, such as tweets, and its developers have taken issues surrounding data protection into account.

It has also been developed so that gathered information can be stored in a manner that allows it to be used as evidence. Currently, SAMi2 only works with Twitter, and multimedia analysis has yet to be implemented, but it has extensible architecture for the addition of new processing capabilities, such as integration with other types of social media. The developers are also working on sentiment analysis. SAMi2's architecture is based on open-source technologies. The system has two processing steps: a shallow step designed to identify tweets of interest and then a deeper one that analyses those tweets to better understand their content. It features a graphical user interface for social network analysis and allows various search profiles to be run continuously.

Madrid City Council's police force is using SAMi2 under realistic conditions and has provided feedback along the whole project cycle. Santos said this had led to some surprises and taught the developers a lot about users' expectations. He also mentioned the use of face-recognition algorithms to track people across different social networks as one of the potential areas for further development, and said they want to connect to the full Twitter application program interface (API) but this is expensive to do.

During a Q&A session, I asked if it was designed to detect instances in which criminals or terrorists substitute words for more innocuous ones. Santos said this is not the case, adding that this would be down to the user – in that, if they knew potential suspects were substituting words, they could use SAMi2 to search for these substitutes. A delegate asked Santos how the system would work with non-English languages. He said that in principle it is possible to change it to work for the majority of European languages – it is just a case of replacing the language tool, which hinges on the availability of tools for that language. While answering a question about false positives, he explained that is difficult to tackle as it is possible to overtrain algorithms to the point where they are no longer effective.

Jason Nurse and Ioannis Agrafiotis from Oxford University's Department of Computer Science reflected on the lessons learnt from the TEASE (Trust-Enabling Augmented-reality Support for information Environments) project, which sought to develop a system to measure the trustworthiness of information on the internet based on parameters such as how recent it is, its location, the source's competence and reputation, and corroboration. Importantly, the user can set the weightings of these parameters. Agrafiotis explained they had looked at various visualisation techniques, and found that users preferred a traffic-light system, which can be combined with tag clouds to show what proportion of users tweeting a particular tag meet the set criteria to be considered trust-worthy, while allowing users to quickly evaluate a high volume of tweets. One of the next steps, according to Nurse, is the use of algorithms to create "world views" – clusters of social media content that are internally consistent – using natural language processing. He also said TEASE is a proof of concept framework and that the next steps would be the development of a fully automated system and its user-based evaluation. The latter would involve two-stage experimentation: its use by responders in a controlled crisis situation and its use in a real-world context, feeding data to first responders. Both TEASE and SAMi2 are currently at the Tier IV/V point of product development.

Moving away from the event, one useful resource for those interested in this topic is *The state of the art 2015 – a literature review of social media intelligence capabilities for counter-terrorism* by Jamie Bartlett and Louis Reynolds.

In the authors' view, "forcing extremists to keep reposting content and creating fresh accounts adds energy to the network, provoking new ways of avoiding moderation and spreading their message. A preferable response is a small number of strategic mass take-down efforts, which would make the network harder to reconstruct…"

They also make the point that recent studies have challenged the idea of the 'self-correcting crowd'. One study examined three rumours and, according to the review, found that "while the ratio of misinformation to correction varied, tweets spreading misinformation outweighed those issuing corrections by between 5:1 and 44:1, and the number of corrections did not always increase in line with an increased circulation of the falsehood they addressed". This highlights the need for systems such as TEASE that can assess the reliability of social media posts.

It would be remiss of me not to mention SUPER (Social sensors for security assessments and proactive emergencies management), a project that started in June 2014 with a budget of €4.25 million, of which €3.12 million is from the EU's Seventh Framework Programme for research, technological development and demonstration. It is exploring "a holistic integrated framework for understanding citizens' reactions against emergencies in social media, while at the same time empowering security forces and civil protection agencies to fully leverage social media in their operations".

Some of its topics include event detection, a rumour identification and spreading framework to enable the detection of malicious content and rumours, along with community analysis allowing the identification of prominent social media communities present during emergencies, their bias and motivations.

But what of the commercial solutions that are already out there? Currently being used in the US is Motorola Solutions' CommandCentral Social platform, which is a partnership with DigitalStakeout. DigitalStakeout handles the social media data-mining and relationships with social media companies such as Twitter and Facebook, while Motorola Solutions handles integration with the rest of its CommandCentral suite and presents the analysed and filtered data in a unified way, together with other sources of information, including computer-aided dispatch systems, ANPR and CCTV cameras.

According to Stephen Beach, smart public safety solutions specialist at Motorola Solutions, this allows someone sitting at their desk to see social media alarms, click on them and, if the activity comes with location information, see if there are nearby CCTV cameras or anything related to the alarm in terms of computer-aided dispatch activity, in addition to learning the poster's identify and the content of their tweet.

"One of the things that makes Digital Stakeout unique is that they don't just access the social media outlets we all know of – there's between 15 and 20 social media engines that it can access," Beach says. There's a distinction between social media monitoring, which looks at the content of posts, and social media analysis, which in Beach's view focuses on their metadata. "There's social media monitoring, which a lot of people do, and then there's social media analytics, which not many people do – and we're one of the companies that does provide that analytics piece."

He notes that currently only a small proportion of social media comes with location data. "A lot of this information, while it could be displayed on mapping engines to show where people are, only works if the social media they are using is GPS-enabled or they are using a device that is; 98 per cent of all social media is not GPS-enabled."

Beach gives an example of the potential benefit of such systems in a real-world context: the Twelfth, which commemorates the Battle of the Blaine between the Protestants and the Catholics on 12 July in Northern Ireland. "We knew that typically this event can get somewhat violent – people are drinking, there's fights – so we [decided to] run a simple search during that event, which occurs around Belfast. So in that area we told our software to identify any individual who [tweeted] about either 'attack' or 'police', and we found results.

"One of these individuals happened to be throwing rocks at the police and, as a result, someone took a photo of them and posted it on Instagram. The caption was something like: 'I saw this man throwing rocks at the police today.' The man is in disguise in the photo but you can see a tattoo on his forearm. If he has been incarcerated before, there's probably a record of that tattoo; simply by identifying it, the police could certainly narrow down the list of suspects to a few [people]."

Making it easier to do the right thing

Returning to the presentations given at PSCE Conference Oxford, Susan Anson, research analyst at Trilateral Research & Consulting, discussed one of its recent research projects, a Comparative Review of Social Media Analysis Tools (SMAT), which was funded by the Global Disaster Preparedness Center and American Red Cross. It assessed 94 social media analysis tools, narrowing them down to 31, based on their suitability for disaster risk reduction (DRR). These were then appraised in more detail, focusing on their features, usability, approximate cost and a number of other factors. The final report provides a catalogue of these tools along with the details that a DRR organisation would need to determine which are most appropriate for its needs.

Some of the barriers to the use of SMAT by such organisations, identified by the report's survey, were: financial limitations (100 per cent), lack of clarity regarding what tools can add to their work (50 per cent), lack of skills (33 per cent), and language barrier (17 per cent). The report also includes guidance for organisations looking to use SMAT, including four in-depth use cases.

Monika Büscher, director of the Centre for Mobilities Research and associate director at the Institute for Social Futures at Lancaster University, pointed out that while information sharing across PPDR agencies can lead to a more co-ordinated response, organisations have reasons not to share their data, such as compliance with data-protection legislation. However, Büscher also highlighted that the failure to share data can have a very real operational impact, giving the example of a victim of the London bombings who was repeatedly asked for her details. She noted the potential for greater sharing of data to allow better tailoring of emergency services to the needs of vulnerable individuals, giving the example of sending an ambulance out already knowing that the patient is diabetic.

Some of the seemingly most effective ways in which the PPDR community has used social media revolve around shaping the public's response to events, thereby reducing the risk that tweets and Facebook posts might hinder a response. Büscher highlighted the way Paris police used social media to ask the public not to tweet officers' locations during the attacks carried out by the so-called Islamic State.

One of the issues Büscher raised during the session was that a statistically normal proportion of false positives could lead to a large number of them, with potentially significant consequences, as seen with the death of Jean Charles de Menezes back in 2005.

She also discussed the 'SecInCoRe (Secure Dynamic Cloud for Information, Communication and Resource Interoperability based on Pan-European Disaster Inventory) project. It will design a secure, dynamic cloud-based knowledge base and communication system for first responders and police authorities, it includes a searchable taxonomy that allows users to rapidly pull up relevant information including guidelines on ethical, legal and social issues, with the idea of creating a "living resource". Büscher said part of the task involves translating the guidelines into suggested actions.

Hopefully this has given you a feel for what's out there and in development. It's clear that control-room operators will have more high-quality data at their fingertips in the years to come, but as any Spiderman fan will know, "with great power comes great responsibility".

Further reading:

- www.demos.co.uk/project/state-of-the-art-2015
- trilateralresearch.com/wp-content/uploads/2015/08/GDPC_SMAT_Short-Report-for-GDPC_Final.pdf
- super-fp7.eu
- www.secincore.eu