2015

A Safety Net for Social Networking: Development of a Predictive Tool for Domestic Terrorism

Matt Campbell III
University of South Alabama, mattcampbell@southalabama.edu

Jennifer Langhinrichsen-Rohling
University of South Alabama, jlr@southalabama.edu

Bob Sweeney
University of South Alabama, bsweeney@southalabama.edu

Jordan Shropshire
University of South Alabama, jshropshire@southalabama.edu

A SAFETY NET FOR SOCIAL NETWORKING: DEVELOPMENT OF A PREDICTIVE TOOL FOR DOMESTIC TERRORISM

Matt Campbell, Ph.D.
University of South Alabama
mattcampbell@southalabama.edu

Jennifer Langhinrichsen-Rohling, Ph.D.
University of South Alabama
jlr@southalabama.edu

Bob Sweeney, Ph.D.
University of South Alabama
bsweeney@southalabama.edu

Jordan Shropshire, Ph.D.
University of South Alabama
jshropshire@southalabama.edu

ABSTRACT

There is a growing trend of disturbed individuals expressing their unhappiness with elements of government or society in general by committing violent acts. In a high number of these cases, the perpetrators gave clues through social networks as to what they planned to do beforehand. While technology can identify certain key words and phrases, it has not advanced to the point of being able to quickly discern between trivial and non-trivial threats. This paper describes ongoing research into developing a tool that can assist in the prediction of terroristic behavior using data from social networks combined with personal knowledge of the individual.

Keywords

Text mining, threat identification, social networking

INTRODUCTION

Recent news reports are full of examples of disturbed individuals who, for whatever reason, decided to express their inner turmoil by attacking others. A few examples include Ismaaiyl Brinsley who shot and killed two NYPD officers, Wenjian Liu and Rafael Ramos, who were sitting in their squad car; Jared Lee Loughner who killed six people while trying to assassinate Representative Gabrielle Giffords in Tucson, AZ; Clay Duke, who shot at board members and eventually killed himself during a Panama City, FL school board meeting; and Joseph Andrew Stack who crashed his small plane into an Austin, Texas building that housed the local FBI, CIA, and IRS offices. What all three of these examples have in common is that each of the perpetrators posted comments on the Internet that spoke of their plans prior to the act of violence. Often these posts were made on the social networking site Facebook. The Facebook social network has more than 500 million users with more than 50 percent of them accessing the network at least once every day (Gays, 2011).

While technology can help identify certain key words and phrases that can indicate potential threats to society, it has not advanced to the point of being able to quickly discern between trivial and non-trivial threats buried in large amounts of unstructured data. The goal of this research is to produce a tool that can assist in the prediction of terroristic behavior using data from social networks combined with personal knowledge of the individual. This research attempts to improve the Department of Homeland Security's (DHS) ability to understand and counter terrorism within the United States by:

- Increasing the awareness of social networking users to possible threats within their sphere of influence,
- Increasing the number of individuals who report appropriate suspicious activity by providing them with automated decision support tools, and
- Increasing the quality of reporting by allowing individuals to easily send all relevant information to law enforcement organizations with the click of a button.

This research will result in an application, named Safety Net, which would notify Facebook users of potentially threatening posts made by their Facebook contacts. Once notified, the user would be prompted to evaluate the post using their personal knowledge of the author (the poster) and guidelines developed by mental health and law enforcement experts to determine if authorities should be alerted. The application would then guide the user through the process of notifying law enforcement if

appropriate. This application could also be expanded in the future to cover other platforms such as MySpace, Twitter, and Blogger.

This proposal breaks new ground in the prediction of terrorist acts because of the integration between predictive computer models and the insight of friends and colleagues who are familiar with the individual making the suspected threat. Current methods focus on text classification computer models alone; however, it is difficult for a computer program to determine the intent of statements such as "I'm gonna do something about Congressman Smith" and "I'm gonna kill my co-workers." Personal knowledge about the person posting the comment is necessary to prevent a high number of frivolous reports that waste considerable law enforcement resources.

The Safety Net application can work as a suite of efforts to increase the public's awareness and to report possible terroristic threats. It retains both the spirit and the goals of Homeland Security's "see something, say something" campaign that was launched in 2010. The investigators also believe that data generated by this application could be used by researchers studying groups and terrorist activity at the macro level.

METHODOLOGY

This research focuses directly on predicting and preventing domestic terrorism behavior. In the examples cited earlier there was opportunity for individuals who were acquainted with the behavior of the terror suspects. However, no action was taken. Five possible reasons for this inaction are:

1. People saw the posts but did not want to get involved.
2. People saw the posts but did not believe the person was serious.
3. People saw the posts but did not know who to contact with their concerns.
4. People did not see the posts until after the acts had occurred because they were not logged into the social network.
5. People did not see the posts because of an overload of information on the social network.

In order to reduce the occurrence of attacks in which the perpetrator posted warning signs to a social networking platform, it is necessary to take steps to address the reasons listed above. To do this, the following questions must be addressed:

1. What prompts users to report or not report suspicious online behavior?
2. How can reporting be increased or encouraged?
3. Can technology be used to accurately identify online posts that reflect intent to harm self or others without a high number of false positives or false negatives?
4. What characteristics of a post and its author can be used to identify the serious threats (e.g. time of posting, words used, other recent posts, recent relationship status change, etc.)?
5. How can the friends or family of these posters best be alerted of the individual's behavior so that steps can be taken to avoid these types of attacks?

This research will enable us to create an application that social networking users can utilize to provide a safety net that would identify potentially threatening posts made by their social networking contacts. The social networking user would then be prompted to evaluate the post using their knowledge of the poster's character along with guidelines developed by mental health experts to determine if authorities should be alerted. The application would then guide the user through the notification process.

The Safety Net application would be a user friendly Facebook application that could be installed with a few clicks. Once installed, the Safety Net would alert the user by text message or email if any of their friends make a post that the behavioral model indicates as a probable threat. Safety Net would then provide the user with a reporting decision support system to determine if the post should be reported to law enforcement and an easy way for the user to report the post to law enforcement if they agreed it constituted a threat.

An example of how the Safety Net application would work:
Jane Doe decides she would like to participate in the Safety Net program and installs the application to her Facebook account. Soon after, one of her Facebook contacts (friend list) named John Smith posts a comment that states "My biology teacher is such a jerk. Tomorrow after class I get my revenge." The Safety Net application would send Jane Doe an email that contained the suspect post. This email would also contain information from mental health and law enforcement experts that describe why this statement may indicate John Smith's intent to harm his professor and other students. The email could

also note information contained in John Smith's Facebook profile such as a recent relationship change (breakup) that may be partially influencing his behavior. Lastly, the email would provide a way for Jane Doe to submit this information online with the click of a button to the appropriate law enforcement agency and a contact phone number in case she would rather talk to someone in law enforcement instead.

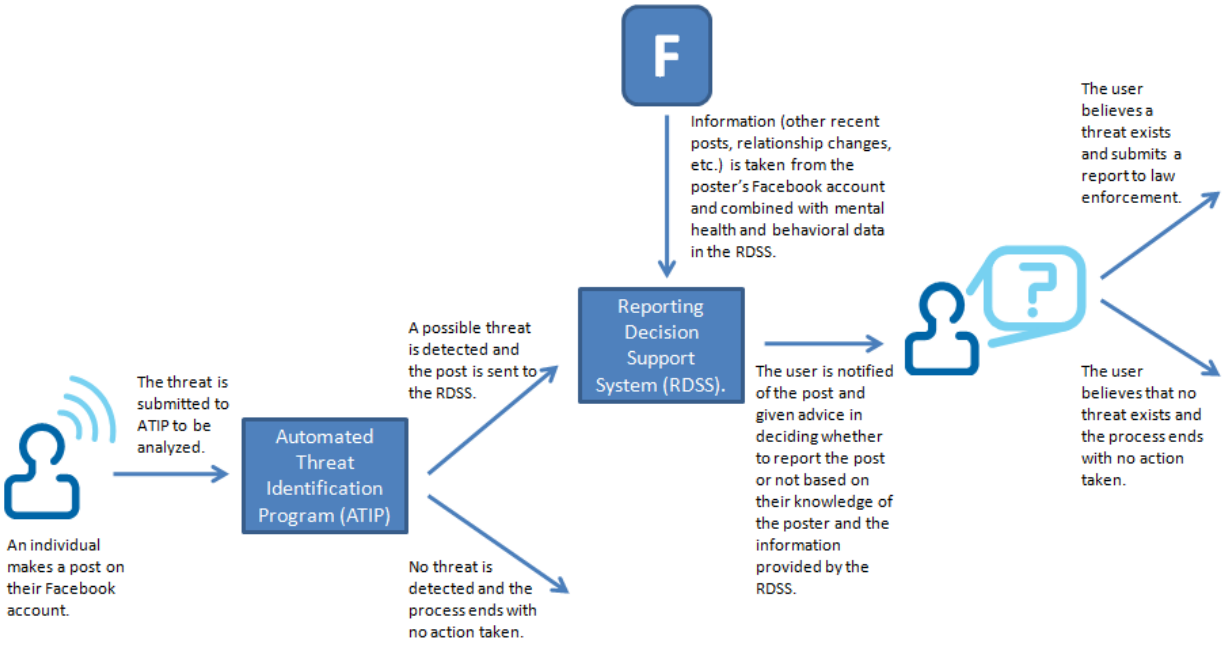

Figure 1: A Diagram of the Safety Net Application

In this example, the program does not violate John Smith's privacy because Jane Doe already has access to all of his profile information by virtue of being on his friend list. Jane Doe would then initiate a report to law enforcement, if appropriate, using her knowledge of John Smith's behavior.

The Safety Net application is made up of two distinct subsystems: the Automated Threat Identification Program (ATIP) which evaluates each post to determine if a possible threat exists and the Reporting Decision Support System (RDSS) which provides the user with expert advice about when threatening communications should be taken seriously and to whom each type of threat should be reported.

The Automated Threat Identification Program (ATIP)

The first level of refinement for evaluating suspicious posts will be performed using text classification technology to classify posts as terroristic or non-terroristic.

The investigators will identify potential message characteristics to be used by the filter by reviewing previous literature on deviant and terroristic behavior. The investigators will also obtain information on previous threats from law enforcement agencies as well as the media. Data elements used to build the model could include content of the post, time of posting, content of other recent posts, recent relationship status changes, and others.

The Reporting Decision Support System (RDSS)

Once a posting has been identified as potentially terroristic in nature, the RDSS would need to compose a message to the user stating that one of their contacts has posted a message that appears to be threatening and asking them to review it. Some postings are clearly terroristic in nature while others are harder to classify (i.e., Mandrusiak, et al., 2006). The investigators must provide the user with the correct supporting information so that they can make an accurate decision based on the post itself and their prior knowledge of the poster. Some work of this type has been conducted to detect suicidal individuals. Specifically, researchers have considered differences in the Internet posts of suicidal individuals as opposed to individuals

who are distressed but not suicidal and those who are non-distressed and non-suicidal (Barak & Miron, 2005). The intensity with which emotions are expressed within the post may also be an important factor; accordingly, recent computer applications have been developed to identify and quantify sentiment strength (Thelwall, Buckley, Paitoglou, & Di Cai, 2010). Studies of written threats made by bomb terrorists have also noted that religious references and social injustices are common themes in these communications, with terrorist acts often functioning as a way to right a grievance or seek revenge (Bomb Threat Management, 1996). Religion has also been shown to be a common theme among suicide terrorists as has the desire to seek vengeance (Townsend, 2007).

The RDSS should enable the user to make the same reporting decision as a qualified mental health expert would when presented with this post and knowledge of the person making it. In order to do this, the investigators must supply the appropriate information to the user without overloading them with too much or overly complicated information. Clinical decision making on an individual's level of dangerousness often relies on the characteristics of psychopathy (e.g., accounting for 60% of the decision making variance, (Elbogen, Huss, Tomkins, & Scalora, 2005) and there is ample evidence that assessing aspects of psychopathy is essential to predicting many types of violence (Langhinrichsen-Rohling, Huss, & Rohling, 2006). Therefore, key considerations of both mental health professional and actuarial decision making strategies about violence include determining whether the person has a history of violence, has engaged in juvenile or adult antisocial behavior, or routinely exhibits a lack of remorse, lack of empathy, and/or a failure to take personal responsibility for his or her actions. Poor behavioral controls or low frustration tolerance, impulsivity, medication noncompliance, substance abuse, poor anger control and experiencing high levels of hostility are also important predictors (Elbogen et al., 2005; Langhinrichsen-Rohling et al., 2006). The program must also take into account how well the user knows the person posting the message. If the poster is only an acquaintance, then more general guidelines will be used to make the reporting decision.

The appropriate reporting action may be as simple as ignoring the post altogether (e.g., if the user knows for a fact that the poster was making a joke) or as serious as submitting the threat to law enforcement. If a threat is deemed to be serious enough to report, the system will assist the user in identifying the correct agency to be contacted and including the necessary information in the report. The investigators feel that this application should not raise serious privacy concerns since it does not retain identifiable information on covered individuals. It simply acts to alert users to questionable posts made by people on their friends list and only uses information that is already available to the user.

DISCUSSION AND CONCLUSION

In order to gauge the success of this tool, it is necessary to identify appropriate methods of evaluation. We propose both qualitative and quantitative metrics for evaluation. There are described below:

A. Qualitative Evaluation

- During development of the reporting decision support system, the investigators will conduct test exercises with student volunteers evaluating past threats and associated biographical information supplied by mental health and law enforcement experts. The volunteers using the reporting decision support system should be able to reach the same conclusions as the mental health and law enforcement experts concerning whether or not to report the threat. The application will also be evaluated for ease of use.

B. Quantitative Evaluation

- The investigators will evaluate the number of valid and non-valid threats reported to law enforcement agencies during the pilot test.
- The investigators will compare the list of ATIP identified threats to the list of actual incidents reported to law enforcement during the pilot test to see if related a Facebook post or comment was made that were not identified by the application.

Although this research describes a proposed pilot test of the Safety Net software, it is our desire that this research will lead to a tool that can actually result in fewer violent attacks and a safer society. The expected outcomes of the project are discussed below.

A. Expected outcomes

One expected outcome from this project includes an increase in the number of terroristic threats reported to law enforcement agencies. These reports would contain a great deal of detail about the author of the threat since data could be taken directly from their Facebook profile.
A second expected outcome would be an increased awareness among the public of disturbed behavior in others and an increased willingness to report this behavior to the appropriate authorities.

B. Pathways to real world application
This research will be conducted in order to foster the development of the Safety Net application. The application will be created for use on the Facebook social network platform, but it could be modified to be used on other social network and blogging platforms such as MySpace, Twitter, and Blogger. The investigators believe this technology could also be used in the detection of suicide threats made online.

C. Potential end users
This application would have two groups of potential end users. The first would be the users who would install the application to their Facebook accounts and would agree to report posts that they believe constitute a threat. The other group of potential end users would be the law enforcement agencies that would receive the reports that users submit using the application.

REFERENCES

1. Agarwal, R. and Karahanna, E. (2000) Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage, *MIS Quarterly*, 24, 4, 665-694.
2. Ajzen, I. (1988) Attitudes, personality, and behavior, The Dorsey Press, Chicago.
3. Ajzen, I. (1991) The theory of planned behavior, *Organizational Behavior & Human Decision Processes*, 50, 2, 179-211.
4. Ghani, J. A., Supnick, R. and Rooney, P. (1991) The experience of flow in computer-mediated and in face-to-face groups, in Janice DeGross, Izak Benbasat, Gerardine DeSanctis and Cynthia Mathis Beath (Eds.) *Proceedings of the Twelfth International Conference on Information Systems*, December 16-18, New York, NY, USA, University of Minnesota, 229 - 237.
5. Tractinsky, N. (1997) Aesthetics and apparent usability: Empirically assessing cultural and methodological issues, in Steve Pemberton (Ed.) *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 97)*, March 22 – 27, Atlanta, GA, USA, ACM Press, 115-122.