The Emerging Technology of Predictive Analytics: Implications for Homeland Security

Brian Lozada
Abstract

This case study examines the implications of using the emerging technology of predictive analytics in a homeland security context to provide increased intelligence for counterterrorism initiatives. The study addresses how this emerging technology is currently used, if such technology is viably transferable to a homeland security context, and if it will be able to provide increased intelligence in regard to the motivations behind terrorist attacks. In addition, the study evaluates whether this data will leverage the detection of potential terrorist profiles and potential terrorist targets to prevent future attacks. By utilizing such technology, homeland security stakeholders will gain a better understanding of how the enemy operates, or plans to operate, as well as identify possible threats before an attack occurs. Investing in intelligence gathering, analysis, and distribution capabilities is a major step in preparing for uncertainties, provided that the emerging technologies utilized for data collection are implemented effectively.
The Emerging Technology of Predictive Analytics: Implications for Homeland Security

In today’s ever-changing threat landscape, technology has aided terrorists in reaching larger audiences, as well as in recruiting and radicalizing via websites and social media. However, technology can also be used by the homeland security community as a tool in combating terrorism. By utilizing emerging technologies to collect intelligence, homeland security stakeholders are able to better understand how the enemy operates, or plans to operate, as well as identify possible threats before an attack occurs. This case study examines the implications of using the emerging technology of predictive analytics within a homeland security context to provide increased intelligence for counterterrorism initiatives.

Securing our nation is a complex, never-ending mission that requires collaboration, focus, and drive from the entire nation, including the federal government, state and local governments, the private sector, and the American people. This mission entails using the emerging technologies to gather intelligence in an effort to better prepare resources to either prevent an attack or respond to one; utilizing an information-sharing approach to distribute intelligence is also crucial to mitigating and preparing for threats.

Investing in intelligence gathering, analysis, and distribution capabilities is a major step in preparing for uncertainties, provided that the emerging technologies utilized for data collection are implemented effectively. The following questions provide a framework for the implications of the emerging technology of predictive analytics in regard to homeland security: How is predictive analytics used, or how can it be implemented, within the homeland security realm? Is such technology transferable to a homeland security context, and does it, or will it be able to, provide increased intelligence in regard to the motivations behind terrorist attacks?
Finally, would this data allow us to detect potential terrorist profiles and potential terrorist targets to prevent future attacks? First, an overview of what predictive analytics is and how it functions from a technical perspective is essential to understanding the technology’s implications for homeland security.

**An Overview of Predictive Analytics**

Predictive analytics in the form of the quantified-self movement has existed for decades in regard to individuals gathering, tracking, and analyzing information on their daily activities in an effort to make better health and behavioral choices for their future. In recent years, however, with advancements in online technologies, the concept of the quantified self has evolved far beyond the scope of the individual; for example, mobile devices, such as tablets and smartphones, include a great deal about people’s activities, including who they talk to, where they go, and what they search (Afeyan, 2014).

The ultimate goal of predictive analytics is to anticipate the outcome of future events by answering the following three questions: What happened? What is happening? What will happen? Predictive analytics is comprised of predictive modeling, which aims to address the who, when, and why questions regarding patterns of individuals’ current behaviors, and forecasting, which concerns their future behavioral patterns. Predictive analytics is used by corporations to foresee trends in customer behavior, product usage, and likelihood of purchases (Issson, 2012). For example, Amazon recently patented a new anticipatory shipping model, in which the company will utilize predictive analytics based on data collected from past purchases, wish lists, and online searches to pre-package an individual’s order before he or she even purchases it (Soper, 2014).

However, predictive analytics is also being used for unconventional means, such as predicting fraudulent insurance claims, military supply chain problems, and the spread of infections. During the 2009 H1N1 influenza outbreak, Google was able to predict the spread of the disease two weeks ahead of government reports based on search term activities and
consumer behavior. This information was essential in preparing local and state healthcare agencies to ensure that the proper medicinal resources were available to treat patients before they even became sick (Isson, 2012). In addition, research on the implications of predictive analytics is currently being conducted at Carnegie Mellon University in a study that analyzes smartphone data to predict the onset of depression by modeling changes in sleep behavior and social relationships over time (Afeyan, 2014).

The Technology of Predictive Analytics

The essential function of predictive analytics technology is to identify patterns among raw, historical data via complex event processing to forecast and assess potential risk; however, there are a variety of prediction analytic methods that are used. Predictive models can be classified as parametric or nonparametric; parametric models operate within a fixed number of parameters regardless of data growth and make particular assumptions about data based on parameter characteristics, whereas nonparametric models can adapt and grow with an increase in data size and do not make any assumptions of raw data (Orbanz, 2010).

The Bayesian model averaging method is a nonparametric model of predictive analytics that has proven to have better accuracy than other traditional models and is becoming increasingly popular in forecasting (Zhu, Kui, & Wang, 2013). During complex event processing, events are detected based on a continuous monitoring and reacting to various event sources, which include but are not limited to social media, global positioning systems (GPS), and radio frequency identification (RFID). First, primitive events are extracted from raw data and time-stamped before they are correlated with one another; then relationships between events are processed, and a response is generated. Complex events are a combination of primitive events that have similar, definable characteristics, such as the same semantic intent or structure (Zhu, Kui, & Wang, 2013).
For example, in a recent study conducted by Gerber (2014) that examined the effectiveness of Twitter-derived information in predicting crime, data was collected on past crimes that occurred during a specified timeframe. This raw data was then time-stamped and defined by the following set of characteristics: latitude and longitude coordinates, type of crime, and frequency of crime. During the same timeframe, tweets tagged with GPS coordinates near where the crimes occurred were also collected and categorized to reveal potential correlations.

In addition, researchers set definable characteristics in regard to semantic intent through the use of topic modeling by providing a list of key words to correspond to their crime prediction model. By analyzing trends within historical data, such as this, predictive analytic modeling can be useful in predicting future, similar events.

**Predictive Analytics and Crime Prevention**

Software tools with predictive analytics have already been adopted by an increasing number of law enforcement agencies across the United States and globally. In Memphis, for example, serious crimes have decreased by 30 percent, and violent crimes have decreased by 15 percent as a result of implementing predictive analytics in 2006 (France-Press, 2012). Predictive policing tools, such as these, aim to predict crimes before they happen based on computer-generated algorithms of patterns found within collected data. Law enforcement agencies are using these predictive analytic tools to help predict, prevent, and respond to crimes much like retailers are using predictive analytics to identify trends in consumer purchasing (France-Presse, 2012).

According to Colleen McCue (2012), a behavioral scientist who works closely with the Department of Homeland Security and local law enforcement agencies, “Studying criminal behavior is not that different from examining other types of behavior like shopping” because people are “creatures of habit” (as cited in France-Presse, 2012). The tendency of people to stay
within their comfort zones makes it easier to create profiles based on their behavioral patterns. The emerging technology of predictive analytics is evolving rapidly and assists in connecting the dots to provide law enforcement agencies better insight as to when and where crimes are most likely to occur.

One element of the predictive analytics technology is the use of geospatial data, which identifies the areas of likely crime based on trending criminal data or scheduled events. The use of geospatial data allows resources to be better allocated to those geographies that have the highest likelihood of crime; therefore, geospatial analysis allows law enforcement agencies to quickly develop a response plan to mitigate a threat as soon as it is detected (Peet, 2012). An example provided by Peet (2012) describes a scenario where a real-time Suspicious Activity Report identifies an unmarked box truck with no visible driver pulled over on the shoulder of the Brooklyn Bridge. Through the use of geospatial technology, predictive analytics can assist with determining the risks to the surrounding geographic location should an explosion occur, as well as identifying the schools and hospitals in the immediate vicinity and compare this information with the strategic needs for the situation.

Another example of using geospatial data and predictive analytics to detect crime has been recently studied by researchers at the University of Virginia. A study on “geo-tagged tweets” revealed that Twitter could be an instrumental tool in predicting certain types of crimes, such as stalking, thefts, and assaults, if the correct analysis is applied. According to Matthew Gerber, researcher in the University’s Predictive Technology Lab, “Future crimes often occur in the vicinity of past crimes, making hot-spot maps a valuable crime prediction tool.” In addition, people tend to tweet about their routine activities, which bring them into environments where crime is more likely to occur. Twitter data is also easy to gather because tweets are publicly available and often tagged with location information (Lever, 2014).
Recent studies in social-media-based predictive analytics, such as the survey conducted by Kalampokis et al. (2014), have been exploring the use of social media, specifically Twitter, to predict various events, including but not limited to pandemics, election results, financial markets, and natural disasters (as cited in Gerber, 2014). A study conducted by Gerber (2014) analyzed tweets posted by residents in Chicago to determine their value in predicting local criminal activity. Results show that crime prediction performance improved with the use of Twitter data for 19 of the 25 crime types included in the study; improved crime prediction, as a result of social media, could have the potential to improve resource allocation and decrease expense, by assigning police officers to the vicinities where crime is more likely to occur (Gerber, 2014).

The use of predictive analytics in these types of scenarios can be directly transferred to a homeland security situation should the threats be perceived as terrorist-related in nature. However, in order for any investigative approach to be successful, the key is gathering as much data as possible; for predictive analytics to be successful in the realm of law enforcement and homeland security initiatives, data gathering is of the utmost importance, as the more data collected, the more opportunity to identify patterns and utilize those patterns to make informed decisions that can help prevent, prepare for, and respond to viable threats.

**Homeland Security Implications for Predictive Analytics**

The emerging technology referenced in this case study will help the homeland security community in providing targeted information and developed intelligence on threats that are actively being discussed on social networks and online media sources. The use of predictive analytics can assist in providing homeland security stakeholders with information to better prevent, prepare for, and recover from an all-hazards event. While this technology was originally developed for private sector use, a partnership with the homeland security community can help develop a process and procedure to use predictive analytics to better safeguard the nation against terrorist-related threats.
The emerging technology of predictive analytics is effectively transferable to a homeland security context, provided there are proper partnerships and integration into the current homeland security framework. Predictive technology will provide stakeholders within the homeland security community the information needed to make appropriate decisions regarding terrorists who have been previously profiled based on their online behavioral patterns; this will allow the homeland security community the ability to actively respond in preventing or preparing for future attacks, as the intelligence discovered may indicate motivations behind terrorist-related threats. In addition, appropriate oversight must also be established for the use of this technology so that it is not misused or misappropriated in support of the homeland security’s mission to utilize emerging technologies of this nature.

From a homeland security perspective, data gathering through predictive analytic modeling can track an individual’s online search history, which can help establish trends and correlations, as well as potential areas of heightened interest to a known radical or terrorist. Once this information is collected and time-stamped, homeland security agencies can actively monitor these individuals’ whereabouts, through the analysis of GPS data, and behaviors, through the analysis of semantic patterning, to potentially mitigate a planned attack. In addition, this technology can be used to analyze and apply internet traffic patterns to determine potential reconnaissance that terrorists are plotting, as well as possible methodologies that terrorists are researching, in order to launch successful attacks against our nation.
Emerging Technologies and Privacy Concerns

As emerging technologies have developed in recent years in response to the need for advanced counterterrorism initiatives, the dual responsibility of gathering information to support such initiatives while protecting the civil liberties and legal rights of individuals as a result of increased information access continues to remain a priority. Yet despite privacy concerns, the trend of extensive data collection to track individuals’ behavior seems to continue to be on the rise.

In 2012, U.S. Attorney General Holder signed off on revised rules regarding government data sharing which would allow for the National Counterterrorism Center within the Department of Homeland Security the ability to analyze existing government records collected about citizens for suspicious patterns of behavior concerning domestic terrorist-related threats (Howard, 2013). According to Howard (2013), this action can be justified for counter-terrorism initiatives, as there are advanced persistent threats; the goal of such data collection and analysis is to find and detect terrorism plots before they occur.

However, these rules seem to override the mission of the Federal Privacy Act of 1974 that protected individuals’ civil liberties (Howard, 2013). Therefore, emerging technologies, such as predictive analytics and active intelligence engines, may result in public controversy should the public believe that infringement on personal liberties exceeds the perceived security benefit. In order to continue to maintain the delicate balance between developing new technologies while, at the same time, ensuring the protection of civil liberties, it is necessary that the nation adheres to homeland security policies that were instated to protect such liberties. In addition, it is necessary that all technologies are safeguarded against the potential threat of vulnerability amongst intelligence sources and methods utilized for information sharing to further increase the public’s confidence in counterterrorism measures.
Conclusion

The guiding principles of homeland security are founded on the ability for all levels of government, as well as the private and public sectors, to develop and improve their capacity to gather, process, analyze, and share critical information. In order for this foundation to remain strong, it is imperative that the nation stays up-to-date on new advances in emerging technologies in order to efficiently prepare, prevent, respond to, and recover from an all-hazards event and maintain interoperability.

All stakeholders within the homeland security community should be encouraged to embrace new technologies to store, transmit, and process information, as well as to utilize all of the tools and resources available to them to better improve future counterterrorism initiatives. Emerging technologies, such as predictive analytics, are just the beginning. Technology is the driver that will foster intelligent decision-making in support of safeguarding our nation, and it will only continue to evolve in the face of the ever-changing threat landscape.
References


