# Using Author Topic to Detect Insider Threats from Email Traffic ⋆

James S. Okolica[1], Gilbert L. Peterson[2], and Robert F. Mills[3]

Air Force Institute of Technology, AFIT/ENG, Bldg 641 RM 220, 2950 Hobson Way, Wright Patterson AFB, OH 45433-7765, USA,
james.okolica, gilbert.peterson, robert.mills @afit.edu

**Abstract.** Despite a technology bias which focuses on external electronic threats, insiders pose the greatest threat to commercial and government organizations. One means of preventing insider theft is by stopping potential insiders from becoming actual thieves. In most cases, individuals do not begin work at an organization with the intent of doing harm. Instead, over time something changes resulting in their becoming an insider threat. By detecting warning signs it is possible to discover potential insiders before they become actual insiders. Using the Author Topic [1] clustering algorithm, we discern employees interests from their daily emails. These interests provide a means to create two social networks that are used to locate potential insiders by finding individuals who either (1) feel alienated from the organization (a key warning sign of a possible disgruntled worker) or (2) have a hidden interest in a sensitive( e.g. proprietary or classified) topic. In both cases, this is revealed when someone demonstrates an interest in a topic but does not share that interest with anyone in the organization.

The dataset used for this research is the Enron email corpus. Unlike most organizations, Enron has a known whistleblower, Sherron Watkins, who was considered an insider threat by her boss, Andy Fastow, who was engaged in the illegal business practices [2]. The first step of the research resolves the Enron email into a collection of stemmed words and frequency counts (i.e. the number of times each word and each individual occurs in each email). These frequency counts are then fed into Author Topic producing four probability distributions: the probability of a word given a topic ($p(w|z)$), the probability of an individual given a topic ($p(u|z)$), the probability of a topic ($p(z)$) and the probability of a topic given a document ($p(z|d)$). The second step creates two social networks for each topic. The first, the implicit interest network, is constructed by linking individuals who have shown an interest in the topic. An individual has an interest in a topic if the conditional probability for an individual ($p(u|z)$) is 1.64 standard deviations above average conditional probability for that topic. The second, the explicit email network, is constructed by linking individuals who have passed an email related to that

⋆ The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government

topic. An email is considered to be related to a topic if the conditional probability for that email ($p(d|z)$) is 1.64 standard deviations above average conditional probability for that topic. Individuals who have links in the implicit interest network but not the explicit email network and classified as having a clandestine interest in that topic.

In looking for potential insiders using the Enron dataset, the first step is selecting the social networks related to a sensitive topic. For this investigation, this topic concerns the off-book partnerships called the Raptors. Four topics emerge with a non-zero conditional probability for the word "raptor". The next step is checking which individuals have clandestine interests in these topics. These individuals have a link in the implicit interest network but none in the explicit email network. This reduces the list of potential insiders from over 34,000 to 71. This process is then repeated using the topics that concern socializing in order to determine which employees may be feeling alienated. Since there is no clear word that defines socializing, several words are used including *dinner, drink, fun, tonight, love, weekend, family* and *game*. Two topics emerge with a non-zero probability for all of these words. When the individuals with clandestine interests in socializing are compared with individuals with a clandestine interest in the off-book partnerships, only three individuals emerge with a clandestine interest in both. Sherron Watkins is one of the three individuals.

Author Topic emerges from this research as an effective tool at revealing potential insiders by datamining email. The topics generated by Author Topic are easily identifiable both based on the most probable words as well as the most probable individuals. In addition, Author Topic effectively reveals Sherron Watkins as a potential insider by flagging both her interest in the Raptor topic as well as her failure to communicate that interest via email to any of her colleagues. However, it is one thing to show that the signs existed that an individual was an insider after the fact. What is needed is to show that Author Topic would have revealed her before the fact. The analysis shows this as well since Sherron Watkins is one of only 3 individuals (out of a possible 34,000) with both a clandestine interest in the Raptor topic and a clandestine interest in socializing. If there was cause for concern that someone might leak information about the Raptors, this analysis would have generated, prior to the leak, a short list of people to pay closer attention to.

## References

1. Rosen-Zvi, Michal and Thomas Griffiths and Mark Steyvers and Padhraic Smyth. "The Author-Topic Model for Authors and Documents". *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*. 487-494, 2004.
2. McLean, Bethany and Peter Elkind. *The Smartest Guys in the Room*. Penguin Group (USA), New York, NY, 2003.