CMI Center for the Management of Information

BORDERS™ A Department of Homeland Security Center of Excellence

THE UNIVERSITY OF ARIZONA

Defending Homeland transportation systems through application of automated linguistic analysis to transcripts of emergency calls for deception detection

Mary B. Burns, Kevin Moffitt, Jeffrey L. Jenkins, Jay F. Nunamaker, Jr. (PI)
Center for the Management of Information, University of Arizona, Phoenix, Arizona

911 calls, a rich source of verbal deception indicators, can be used as a proxy for bomb threat calls in exploratory linguistic cue analysis

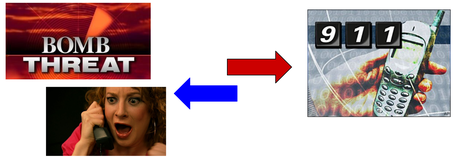

Figure 1. Emergency phone calls as data for linguistic cue analysis

A quick and decisive response to bomb threats aimed at America's complex transportation systems is crucial to prevent, mitigate, respond to, and recover from terrorist strikes. Rapidly and accurately classifying a bomb threat as 'real' vs. 'hoax' is imperative. Due to natural 'truth bias', humans, including those with special training, have disappointing results, slightly better than chance at 54%, in detecting deception (Aamodt & Custer, 2006; Bond & DePaulo, 2006). Thus, the purpose of this research is to determine if natural language processing (NLP) data mining techniques can augment human capabilities by detecting deception accurately in a real-world, relatively unrehearsed context. As a proxy for bomb threat calls (Figure 1), we will use publicly available emergency calls to explore the following research question: can automated linguistic analysis techniques accurately classify deceptive vs. truthful callers in transcripts of 911 homicide calls?

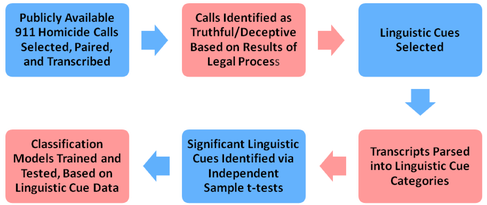

Figure 2. Steps in collecting, preparing, and analyzing 911 homicide calls

Methodology

From reputable Internet sites, we collected and transcribed fifty audio files of 911 calls that reported homicides, equally split between truthful and deceptive callers (Figure 2). To establish ground truth, we corroborated subsequent arraignment, prosecution, and/or admission of guilt via news articles of the crimes. Next, we removed the 911 operators' utterances. Informed by theories of deception detection, linguistics, and computer science, we applied linguistic cue analysis techniques to the transcripts to parse the data into linguistic cue categories, such as words expressing anxiety. Table 1 shows the hypotheses for significant cue categories with examples from transcripts. Next, we trained various machine-learning classification algorithms on the cues and tested their classification accuracy using a 10-fold cross-validation as a bootstrap technique to increase the validity of the results. Table 2 shows the results of cross-validation classification methods.

Results and future research

Our results suggest that truthful vs. deceptive calls can be discriminated via linguistic cue features such as more negative emotion and anxiety, reference to others in the singular or plural (non-immediacy), or suppression of verbal responses. The overall performance of the classification techniques ranged as high as 90% for the cross-validation set (Table 2). Our findings provide critical knowledge about how deceivers communicate during unrehearsed verbal exchanges. While this study focused on 911 homicide calls because of the availability of the data, future studies should investigate bomb threat calls. An integrated system for deception detection that can be used real-time against bomb threat calls could be added to an overall strategy for safeguarding complex transportation systems.

Hypotheses	Results	Example from transcripts
H1: Deceivers will use more non-immediacy terms (e.g. 3rd person plural pronouns) than truthtellers.	**Supported**	Yes **they** said, **they** said if **they** heard anything **they** were going to my house.
H2: Deceivers will control, limit, or suppress answers (e.g., use more terms of negation or assent) than truthtellers.	**Supported**	**No nothing**, he's gone.
H3: Truthtellers will display more feelings such as negative emotion of anxiety than deceivers.	**Supported**	I found out about an hour ago and I've been in a **panic** ever since.
H4: Deceivers, suffering from cognitive overload, will not be able to respond with clear details to operator (e.g., details about location) than truthtellers.	**Supported**	**Five seventeen** West Doty Street

Table 1. Hypotheses for significant linguistic cue categories, results, and significant variables with examples from transcripts

Classification methods	Logistic regression	Naïve Bayes	Random forest	Neural network	Discriminant analysis
Overall performance	82%	78%	70%	74%	90%
Truth performance	82%	76%	64%	64%	88%
Deception performance	84%	80%	76%	84%	92%

Classification results yielded predictive models with much higher accuracy than that of unaided humans.

Table 2. Results of classification algorithms (cross-validation set)

Literature cited

Aamodt, M., & Custer, H. (2006). Who Can Best Catch a Liar? *Forensic Examiner, 15*(1), 6-11.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality and Social Psychology Review, 10*(3), 214-234.

Acknowledgments

This project was funded through the BORDERS Center of Excellence by a grant from the Department of Homeland Security, Science and Technology Directorate, Office of University Programs.

For further information:

Please contact mburns@cmi.arizona.edu