Cognitive hacking and intelligence and security informatics

Paul Thompson
Thayer School of Engineering and
Department of Computer Science
Dartmouth College
Hanover, New Hampshire 03755
Paul.Thompson@dartmouth.edu

ABSTRACT

This paper describes research on cognitive and semantic attacks on computer systems and their users. Several countermeasures against such attacks are described, including a description of a prototype News Verifier system. It is argued that because misinformation and deception play a much more significant role in intelligence and security informatics than in other informatics disciplines such as science, medicine, and the law, a new science of intelligence and security informatics must concern itself with semantic attacks and countermeasures.

Keywords: Cognitive hacking, semantic hacking, intelligence informatics, security informatics, deception detection

1. INTRODUCTION

Libicki first characterized attacks on computer systems in the context of information warfare as being physical, syntactic, and semantic, where software agents were mislead by misinformation deliberately fed by an adversary[1]. Recently Cybenko et al.[2] defined cognitive hacking as an attack on a computer system directed at the mind of the user of the system, which, in order to succeed, had to influence the user's perceptions and behavior. In addition to this work in computer security other related research has been undertaken on deception detection in the fields of psychology and communications[3] and in the fields of forensic linguistics and in literary and linguistic computing, in particular research on authorship attribution[4].

Semantic attacks and their countermeasures are expected to be an important area of research in a new science of intelligence and security informatics, as called for by the National Science Foundation and the National Institute of Justice[5]. This paper describes research begun on cognitive hacking countermeasures; generalizes the concept of cognitive hacking, placing it in Libicki's framework of semantic attack; and discusses the role of semantic attacks and countermeasures in the context of a new science of intelligence and security informatics.

2. BACKGROUND

In 1981, Landwehr provided a characterization of computer system security which has framed subsequent discussion of computer security[6]. His definition arose from a consideration of the requirements of military security. He postulated that:

> Information contained in an automated system must be protected from three kinds of threats: (1) the *unauthorized disclosure* of information, (2) the *unauthorized modification* of information, and (3) the *unauthorized withholding* of information (usually called *denial of service*)

Libicki described semantic attacks in the context of information warfare, where software agents were mislead by misinformation which they were being deliberately fed by an adversary[1]. Schneier, by contrast, defined semantic attacks

as ". . . attacks that target the way we, as humans, assign meaning to content. . . . Semantic attacks directly target the human/computer interface, the most insecure interface on the Internet"[7].

Denning developed a similar notion to semantic attacks, which she referred to as information warfare[8], described as a struggle over an information resource by an offensive and a defensive player. The resource has an exchange and an operational value. The value of the resource to each player can differ depending on factors related to each player's circumstances. The outcomes of offensive information warfare are: increased availability of the resource to the offense, decreased availability to the defense, and decreased integrity of the resource. Although not receiving as much attention as worms, viruses, or distributed denial of service attacks, the majority of attacks on a computer system come from insiders. Detecting such insider misuse is an important area for semantic attack countermeasures.

Computer and network security present great challenges to our evolving information society and economy. The variety and complexity of cyber security attacks that have been developed parallel the variety and complexity of the information technologies that have been deployed. Physical and syntactic attacks operate totally within the fabric of the computing and networking infrastructures. For example, the well-know Unicode attack against older, unpatched versions of Microsoft's Internet Information Server (IIS) can lead to root/administrator access. Once such access is obtained, any number of undesired activities by the attacker is possible. For example, files containing private information such as credit card numbers can be downloaded and used by an attacker. Such an attack does not require any intervention by users of the attacked system. By contrast, a *cognitive* attack requires some change in users' behavior, accomplished by manipulating their perception of reality. The attack's desired outcome cannot be achieved unless human users change their behaviors in some way. Users' modified actions are a critical link in the sequencing of a cognitive attack.

Provision of misinformation, the intentional distribution or insertion of false or misleading information intended to influence reader's decisions and/or activities, is a form of cognitive hacking. The Internet's open nature makes it an ideal arena for dissemination of misinformation. Cognitive hacking differs from social engineering, which, in the computer domain, involves a hacker's psychological tricking of legitimate computer system users to gain information, e.g., passwords, in order to launch a syntactic attack on the system.

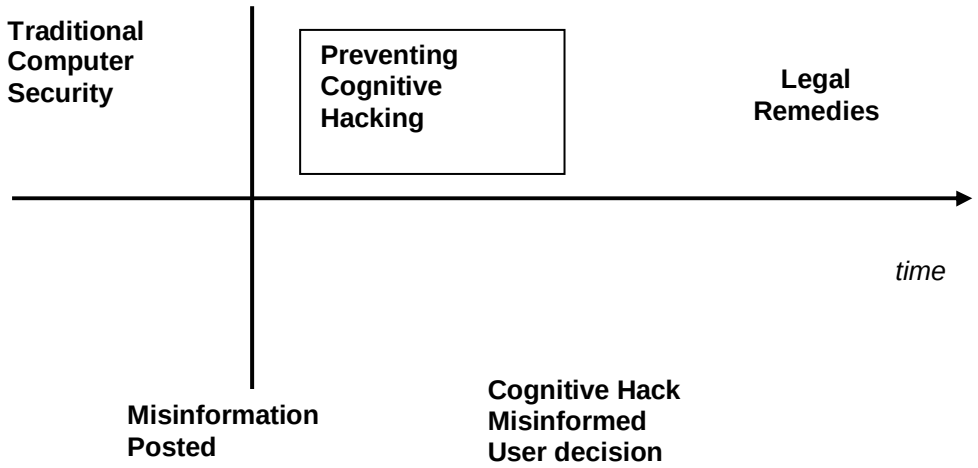

Figure 1. The sequencing of a cognitive attack

Consider the graph in figure 1. Most analyses of computer security focus on the time before misinformation is posted, i.e., on preventing unauthorized use of the system. A cognitive hack takes place when a user's behavior is influenced by misinformation. At that point the focus is on detecting that a cognitive hack has occurred and on possible legal action. Our concern is with developing tools to prevent cognitive hacking, that is, tools that can recognize and respond to misinformation before a user acts based on the misinformation. By contrast, a *cognitive* attack requires some change in users' behavior, effected by manipulating their perceptions of reality. The attack's desired outcome cannot be achieved unless human users change their behaviors in some way. Users' modified actions are a critical link in a cognitive attack's sequencing.

The use of information, or misinformation, to affect the behavior of humans is not new. Language, or more generally communication, is used by one person to influence another. Propaganda has long been used by governments, or by other groups, particularly in time of war, to influence populations[9,10,11,12]. Although the message conveyed by propaganda, or other communication intended to influence, may be believed to be true by the propagator, it usually is presented in a distorted manner, so as to have maximum persuasive power, and, often, is deliberately misleading, or untrue. Propaganda is a form of perception management. Other types of perception management include psychological operations in warfare[13], consumer fraud, and advertising[9]. As described in section 4.1, deception detection has long been a significant area of research in the disciplines of psychology and communications.

3. INTELLIGENCE AND SECURITY INFORMATICS

Cybenko et al.[2] define cognitive hacking as gaining access to, or breaking into, a computer information system for the purpose of modifying certain behaviors of a human user in a way that violates the integrity of the overall system, including the user and the information system. The integrity of the information system would for example include correctness or validity of the information the user gets from such a system. In this context, the integrity of a computer system can be defined more broadly than the definition implicit in Landwehr's classic definition of computer security in terms of confidentiality, integrity, and accessibility[4]. In Cybenko et al.[2] the focus of semantic attacks was on manipulation of consumers on the Internet, e.g., through pump-and-dump schemes. This paper extends this focus to encompass the concepts of semantic attacks put forward by Libicki[1] and Schneier[7] and to argue that understanding semantic attacks, and the development of countermeasures against semantic attacks is of central importance for a science of intelligence and security informatics.

Intelligence and security informatics will be supported by data mining, visualization, and link analysis technology, but intelligence and security analysts should also be provided with an analysis environment supporting mixed-initiative interaction with both raw and aggregated data sets[14]. Since analysts will need to defend against semantic attacks, this environment should include a toolkit of semantic hacking countermeasures. For example, if faced with a potentially deceptive news item from the Foreign Broadcast Information Service (FBIS), an automated countermeasure might provide an alert using adaptive fraud detection algorithms[15] or through a retrieval mechanism which would allow the analyst to quickly assemble and interactively analyze related documents bearing on the potential misinformation.

Information retrieval, or document retrieval, developed historically to serve the needs of scientists and legal researchers, among others. Despite occasional hoaxes and falsifications of data in these domains, the overwhelming expectation is that documents retrieved are honest representations of attempts to discover scientific truths, or to make a sound legal argument. This assumption does not hold for intelligence and security informatics. Most information retrieval systems are based either on: a) an exact match Boolean logic by which the system divides the document collection into those documents matching the logic of the request and those that do not, or b) ranked retrieval. With ranked retrieval a score is derived for each document in the collection based on a measure of similarity between the query and the document's representation, as in the vector space model[16], or based on a probability of relevance[17,18].

Although not implemented in existing systems, a utility theoretic approach to information retrieval[19] shows promise for a theory of intelligence and security informatics. In information retrieval predicting relevance is difficult enough. Predicting utility, although more difficult, would be more useful. When information contained in, say, a FBIS document,

may be misinformation, then the notion of utility theoretic retrieval, becomes more important. The provider of the content may have believed the information to be true or false, aside from whether it was true or false in some objective sense. The content may be of great value to the intelligence or security analyst, whether it is true or false, but, in general, it would be important to know not only whether it was true or false, but also whether the provider believed it to be true or false. Current information retrieval algorithms would not take any of these complexities into account in calculating a probability of relevance.

3.1. Deception Detection

Deception of detection in interpersonal communication has long been a topic of study in the fields of psychology and communications[3,20.21]. The majority of interpersonal communications are found to have involved some level of deception. Psychology and communications researchers have identified many cues that are characteristic of deceptive interpersonal communication. Most of the this research has focused on the rich communication medium of face-to-face communication, but more recently other forms of communication have been studied such as telephone communication and computer-mediated communication[22]. A large study is underway[21] to train people to detect deception in communication. Some of this training is computer-based. Most recently a study has begun to determine whether psychological cues indicative of deception can be automatically detected in computer-mediated communication, e.g., e-mail, so that an automated deception detection tool might be built[22,23].

3.2. Predictive modeling in intelligence and security informatics

Predictive modeling using the concepts of cognitive hacking and utility-theoretic information retrieval can be applied in two intelligence and security informatics settings which are mirror images of each other, i.e., the user's model of the system's document content and the system's model of the user as a potential malicious insider. Consider an environment where an intelligence analyst accesses sensitive and classified information from intelligence databases. The accessed information itself may represent cognitive attacks coming from the sources from which it has been gathered, e.g., FBIS documents. As discussed above, each of these documents will have a certain utility for the analyst, based on the analyst's situation, based on whether or not the documents contain misinformation, and, if the documents do contain misinformation, whether, or not, the analyst can determine that the misinformation is present. On the other hand, the analyst might be a malicious insider engaged in espionage. The document system will need to have a cost model for each of its documents and will need to build a model of each user, based on the user's transactions with the document system and other external actions. These user models could be expressed as Hidden Markov Models[26].

Denning's theory of information warfare[8] and an information theoretic approach to the value of information[24,25] can be used to rank potential risks given the value of each document held by the system. Particular attention should be paid to deception on the part of the trusted insider to evade detection, using cognitive hacking countermeasures. Modeling the value of information to adversaries will enable prediction of which documents are likely espionage targets and will enable development of hypotheses for opportunistic periods and scenarios for compromise. These models will be able to detect unauthorized activity and to predict the course of a multi-stage attack so as to inform appropriate defensive actions.

4. COGNITIVE HACKING COUNTERMEASURES

Cognitive hacking on the internet is an evolving and growing activity, often criminal and prosecutable. Technologies for preventing, detecting and prosecuting cognitive hacking are still in their infancies. Given the variety of approaches to and the very nature of cognitive hacking, *preventing* cognitive hacking reduces either to preventing unauthorized access to information assets (such as in web defacements) in the first place or detecting posted misinformation before user behavior is affected (that is, before behavior is changed but possibly after the misinformation has been disseminated). The latter may not involve unauthorized access to information, as for instance in "pump and dump" schemes that use newsgroups and chat rooms. By definition, *detecting* a successful cognitive hack would involve detecting that the user behavior has already been changed. Detection in that sense is not the focus of this research.

Discussion of methods for preventing cognitive hacking will be restricted to approaches that could automatically, or semi-automatically, alert users to problems with their information source or sources (information on a web page, newsgroup, chat room and so on). Techniques for preventing unauthorized access to information assets fall under the general category of computer and network security and will not be considered here. Similarly, detecting that users have already modified their behaviors as a result of the misinformation, namely that a cognitive hack has been successful, can be reduced to detecting misinformation and correlating it with user behavior.

The cognitive hacking countermeasures discussed here will be primarily mathematical and linguistic in nature. The use of linguistic techniques in computer security and information assurance, more generally, has been pioneered by Raskin and colleagues at Purdue University's Center for Education and Research in Information Assurance and Security (CERIAS)[27]. Their work, however, has not addressed cognitive hacking countermeasures. The discussion of countermeasures will be broken into two categories: countermeasures for situations in which there is a single source of information (which may be misinformation), and situations where there are multiple sources. The countermeasures to be deployed do not necessarily divide evenly between these two situations. For example, name attribution countermeasures could be useful in both circumstances.

4.1. Single source cognitive hacking countermeasures

Here a few possible approaches for the single source problem are considered. By single source, is meant situations in which redundant, independent sources of information about the same topic are not available. An authoritative corporate personnel database would be an example. More examples of single source countermeasures are discussed in Cybenko et al.[2]

4.1.1. Authentication of source

This technique involves due diligence in authenticating the information source and ascertaining its reliability. Various relatively mature certification and PKI technologies can be used to detect spoofing of an information server. Additionally, reliability metrics can be established for an information server or service by scoring its accuracy over repeated trials and different users. In this spirit, Lynch[28] describes a framework in which trust can be established on an individual user basis based on both the identity of a source of information, through PKI techniques for example, and in the behavior of the source, such as could be determined through rating systems. Such an approach will take time and social or corporate consensus to evolve.

4.1.2. Information "Trajectory" Modeling

This approach requires building a model of a source based on statistical historical data or some sort of analytic understanding of how the information relates to the real world. For example, weather data coming from a single source (website or environmental sensor) could be calibrated against historical database (from previous years) or predictive model (extrapolating from previous measurements). A large deviation would give reason for hesitation before committing to a behavior or response. Modeling information sources is something that can be done on a case-by-case basis as determined by the availability of historical data and the suitability of analytic modeling.

4.1.3. Linguistic Countermeasures with Single Sources: Genre Detection and Authority Analysis

A careful human reader of some types of misinformation, e.g., exaggerated pump-and-dump scheme postings on the Web about a company's expected stock performance, can often detect the misinforming posting from other legitimate postings, even if these legitimate postings are also somewhat hyperbolic. Since Mosteller and Wallace's seminal work on authorship attribution in 1964[29], statistical linguistics approaches have been used to recognize the style of different

writings. In Mosteller and Wallace's work this stylistic analysis was done to determine the true author of anonymous Federalist papers, where the authorship was disputed. Since then Biber[30,31] and others[32] have analyzed the register and genre of linguistic corpora using similar stylistic analysis. Kessler et al.[33] have developed and tested algorithms based on this work to automatically detect the genre of text.

4.1.4. Psychological Deception Cues

The approach to genre analysis taken, e.g., by Biber and Kessler et al., is within the framework of corpus linguistics, i.e., based on a statistical analysis of general word usage in large bodies of text. The work on deception detection in the psychology and communications fields is based on a more fine-grained analysis of linguistic features, or cues. Psychological experiments have been conducted to determine which cues are indicative of deception. To date this work has not led to the development of software tools to automatically detect deception in computer-mediated communication, but researchers see the development of such tools as one of the next steps in this line of research[23].

4.2. Multiple source cognitive hacking

In this section possible approaches to preventing cognitive hacking when multiple, presumably redundant, sources of information are available about the same subject of interest are discussed. This is clearly the case with financial, political and other types of current event news coverage. More examples of multiple source countermeasures are discussed in Cybenko et al.[2]

Several aspects of information dissemination through digital, network media, such as the Internet and World Wide Web, make cognitive hacking possible and in fact relatively easy to perform. Clearly, there are enormous market pressures on the news media and on newsgroups to quickly disseminate as much information as possible. In the area of financial news, in particular, competing news services strive to be to the first to give reliable news about breaking stories that impact the business environment. Such pressures are at odds with the time consuming process of verifying accuracy. A compromise between the need to quickly disseminate information and the need to investigate its accuracy is not easy to achieve in general. Automated software tools could in principle help people make decisions about the veracity of information they obtain from multiple networked information systems. A discussion of such tools, which could operate at high speeds compared with human analysis, follows.

4.2.1. Source reliability via collaborative filtering and reliability reporting

The problem of detecting misinformation on the Internet is much like that of detecting other forms of misinformation, for example in newsprint or verbal discussion. Reliability, redundancy, pedigree and authenticity of the information being considered are key indicators of the overall "trustworthiness" of the information. The technologies of collaborative filtering and reputation reporting mechanisms have been receiving more attention recently, especially in the area of on-line retail sales[34]. This is commonly used by the many on-line price comparison services to inform potential customers about vendor reliability. The reliability rating is computed from customer reports. Another technology, closely related to reliability reporting is collaborative filtering[35]. This can be useful in cognitive hacking situations that involve opinions rather than hard objective facts.

Both of these approaches involve user feedback about information that they receive from a particular information service, building up a community notion of reliability and usefulness of a resource. The automation in this case is in the processing of the user feedback, not the evaluation of the actual information itself.

4.2.2. News Verifier: a countermeasure for misinforming news stories

Consider the following scenario. An end user is examining a posting to the business section of Google News[36]. The document purports to provide valuable news about a publicly traded company that the user would like to act on quickly by purchasing, or selling stock. Although this news item might be reliable, it might also be misinformation being fed to unwary users by a cognitive hacker as part of a pump-and-dump scheme, i.e., a cognitive hacker's hyping of a company by the spread of false, or misleading information about the company and the hacker's subsequent selling of the stock as the price of its shares rise, due to the misinformation. The end user would like to act quickly to optimize his or her gains, but could pay a heavy price, if this quick action is taken based on misinformation.

News Verifier, a prototype cognitive hacking countermeasure, allows an end user to effectively retrieve and analyze documents from the Web that are similar to the original news item. When the end user receives a news item that he, or she, suspects, may represent a cognitive attack, i.e., contain deliberate misinformation, the user can run the News Verifier. First, a query is automatically generated from the text of the news item. This query is then sent automatically to an API for Google News. Then, a set of documents is retrieved by the Google News clustering algorithm. The Google News ranking of the clustered documents is generic, not necessarily optimized as a countermeasure for cognitive attacks. News Verifier uses a combination process in which several different search engines are used to provide alternative rankings of the documents initially retrieved by Google News. The ranked lists from each of these search engines, along with the original ranking from Google News, are combined using the Combination of Expert Opinion algorithm[37,38,39] to provide a more optimal ranking. Relevance feedback judgments from the end user are used to train the constituent search engines. It is expected that this combination and training process will yield a better ranking than the initial Google News ranking. This is an important feature in a countermeasure for cognitive hacking, because a victim of cognitive hacking will want to detect misinformation as soon as possible in real time.

4.3. Linguistic countermeasures with multiple sources: authorship attribution

Authorship attribution using stylometry is a field of study within statistics and computational linguistics with a long history. In 1964, Mosteller and Wallace[29] resolved a longstanding debate on the authorship of certain of the Federalist Papers. More recently, a principal components analysis approach to authorship attribution has been pioneered by Burrows[40] in the field of literary and linguistic computing. Rao and Rohatgi[41] have shown that Burrows' techniques can be employed even more successfully with text taken from the Internet. Forensic linguistics has become a recognized research discipline with professional societies including the International Association of Forensic Linguists and professional journals such as *Forensic Linguistics*[42]. Contemporary experts such as Donald Foster have assisted law enforcement with investigations such as the Unabomber and the murder of JonBenet Ramsey[43]. There are general linguistic text analysis software packages, e.g., Pennebaker et al.[22], that can be used for forensic linguistic analysis, but tools specifically tailored for forensic linguistic authorship attribution do not yet exist. A recent account of research on authorship attribution is given by Harold Love[4]; while works on forensic linguistics include Rieber and Stewart[45], McMenamin and Choi[46], Shuy[47], and Grant[48].

Stylometry techniques can be used to determine the likelihood that two documents of uncertain authorship are written by the same author, or that a document of unknown authorship is written by an author from whom sample writings are available. Similarly, given a set of documents with several authors, it is possible to partition the documents into subsets of documents all written by the same author. There are two parameters in such techniques: a) the data requirements per pseudonym, and b) the discriminating power of the technique. Using only semantic features, Rao and Rohatgi[19] demonstrated that anonymity and pseudonymity cannot preserve privacy. Rao and Rohatgi did some exploratory research to confirm that inclusion of syntactic features, e.g., misspellings or other idiosyncratic features much more prevalent in web, as opposed to published, documents, could provide stronger results. Authorship attribution is an important countermeasure for semantic attacks. For example, in a pump-and-dump scheme one, or a small number of authors, often use many more pseudonymous user accounts to make it appear that a company is being viewed favorably by a large number of people. Authorship attribution techniques could be used to recognize that in fact there were one, or a few authors, rather than hundreds, making favorable postings.

5. CONCLUSION

Misinformation, or semantic hacking, plays a much more prominent role in intelligence and security informatics than it has played in traditional scientific informatics. The status of content as information, or misinformation, in turn, influences its utility for users. This paper suggests the need for tools to detect and defend against cognitive and semantic attacks along side data mining and information retrieval, as an important foundational technology for intelligence and security informatics.

ACKNOWLEDGEMENTS

Support for this research was provided by a Department of Defense Critical Infrastructure Protection Fellowship grant with the Air Force Office of Scientific Research, F49620-01-1-0272; Defense Advanced Research Projects Agency projects F30602-00-2-0585 and F30602-98-2-0107; and the Office of Justice Programs, National Institute of Justice, Department of Justice award 2000-DT-CX-K001 (S-1). The views in this document are those of the authors and do not necessarily represent the official position of the sponsoring agencies or of the US Government.

REFERENCES

1. M. Libicki, The mesh and the Net: Speculations on armed conflict in an age of free silicon National Defense University McNair Paper 28, 1994. http://www.ndu.edu/ndu/inss/macnair/mcnair28/m028cont.html
2. G. Cybenko, A. Giani, and P. Thompson, Cognitive Hacking: A Battle for the Mind, *IEEE Computer* vol. 35, no. 8, August 2002, p. 50-56.
3. D. Buller and J. Burgoon, Interpersonal deception theory, *Communication Theory*, vol. 6, no. 3, special issue, *Interpersonal deception: theory and critique*, 1996, p. 203-242.
4. H. Love, *Attributing Authorship: An Introduction* Cambridge University, Cambridge, U.K., 2002.
5. *NSF / NIJ Symposium on Intelligence and Security Informatics*, *Lecture Notes in Computer Science*, Berlin: Springer-Verlag, 2003.
6. C. Landwehr, Formal models of computer security, *Computing Surveys*, vol. 13, no. 3, 1981.
7. B. Schneier, Semantic attacks: The third wave of network attacks *Crypto-gram Newsletter* October 15, 2000. http://www.counterpane.com/crypto-gram-0010.html
8. D. Denning, *Information warfare and security*, Addison Wesley, Reading, Massachuesetts., 1999.
9. J. Combs and D. Nimmo, *The new propaganda: The dictatorship of palaver in contemporary politics*, Longman, New York, 1993.
10. L. Doob, *Propaganda: Its psychology and technique*, Holt, New York, 1935.
11. J. Ellul, *Propaganda* translated from the French by K. Kellen and J. Lerner, Knopf, New York, 1966.
12. A. Pratkanis and E. Aronson, *Age of propaganda: The everyday use and abuse of persuasion*, Freeman, New York, 1992.
13. Information Warfare Site, 2001, http://www.iwar.org.uk/psyops/index.htm
14. P. Thompson, Semantic Hacking and Intelligence and Security Informatics, *NSF / NIJ Symposium on Intelligence and Security Informatics*, *Lecture Notes in Computer Science*, Berlin: Springer-Verlag, June 1-3, Tucson, Arizona, 2003.
15. T. Fawcett, and F. Provost, Fraud detection, In W. Kloesgen and J. Zytkow (eds.) *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, 2002.
16. G. Salton, and M. McGill, *Introduction to Modern Information Retrieval* McGraw-Hill, New York, 1983.
17. M. Maron, and J. Kuhns, On relevance, probabilistic indexing and information retrieval, *Journal of the ACM* vol. 7 no. 3, 1960, p. 216-244.
18. C. van Rijsbergen, *Information Retrieval*, 2d. edition, Buttersworth, London, 1979.
19. W. Cooper, and M. Maron, Foundations of Probabilistic and Utility-Theoretic Indexing, *Journal of the Association for Computing Machinery*, vol. 25, no. 1, 1978, p. 67-80.
20. K. Cornetto, Identity and illusion on the Internet: interpersonal deception and detection in interactive Internet environments, Ph.D. thesis, University of Texas at Austin, 2001.

21. J. Cao, J. Crews, M. Lin, J. Burgoon, J. Nunamaker, Designing Agent99 Trainer: a learner-centered, Web-based training system for deception detection, *NSF / NIJ Symposium on Intelligence and Security Informatics*, *Lecture Notes in Computer Science*, Berlin: Springer-Verlag, June 1-3, Tucson, Arizona, 2003.
22. L. Zhou, J. Burgoon, D. Twitchell, A longitudinal analysis of language behavior of deception in e-mail, *NSF / NIJ Symposium on Intelligence and Security Informatics*, *Lecture Notes in Computer Science*, Berlin: Springer-Verlag, June 1-3, Tucson, Arizona, 2003.
23. L. Zhou, D. Twitchell, T.Qin, J. Burgoon, J. Nunamaker, An exploratory study into deception in text-based computer-mediated communication, *Proceedings of the Hawaii International Conference on Systems Science,* 2003.
24. T. Cover, and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
25. G. Cybenko, A. Giani, and P. Thompson, Cognitive Hacking and the Value of Information, *Workshop on Economics and Information Security,* Berkeley, California, May 16-17, 2002.
26. P. Thompson, Weak models for insider threat detection, *Proceedings of the Defense and Security Symposium*. April, Orlando, Florida, 2004.
27. M. Attallah, C. McDonough, V. Raskin, and S. Nirenberg, Natural language processing for information assurance and security: An overview and implementation, In: M. Schaefer (ed.), *Proceedings. New Security Paradigm Workshop,* September 18th-22nd, Ballycotton, County, Cork Ireland, ACM Press, New York, 2000, p. 51-65.
28. C. Lynch, When documents deceive: trust and provenance as new factors for information retrieval in a tangled Web, *Journal of the American Society for Information Science & Technology*, vol. 52, no. 1, 2001, p. 12-17.
29. F. Mosteller and D. Wallace, *Inference and Disputed Authorship: The Federalist*, Addison-Wesley, Reading, Massachusetts, 1964.
30. D. Biber, *Dimensions of Register Variation: a Cross-Linguistic Comparison,* Cambridge University Press, Cambridge, England, 1995
31. D. Biber, Spoken and written textual dimensions in English: resolving the contradictory findings *Language*, vol. 62, no. 2, 1986, p. 384-413
32. J. Karlgren and D. Cutting, Recognizing text genres with simple metrics using discriminant analysis, *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94)*, 1994.
33. B. Kessler, G. Nunberg, and H. Schuetze, Automatic detection of genre, *Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, 1997.
34. C. Dellarocas, Building trust on-line: The design of reliable reputation reporting mechanisms for online trading communities, *Center for eBusiness@MIT* paper 101, 2001.
35. J. Thornton, Collaborative Filtering Research Papers, 2001, http://jamesthornton.com/cf/.
36. Google News, http://news.google.com/.
37. P. Thompson, A combination of expert opinion approach to probabilistic information retrieval, part 1: the conceptual model." *Information Processing and Management*, vol. 26, no. 3, 1990, p. 371-382.
38. P. Thompson, A combination of expert opinion approach to probabilistic information retrieval, part 2: mathematical treatment of CEO model 3." *Information Processing and Management,* vol. 26, no. 3, 1990, p. 383-394.
39. G. Mateescu, M. Sosonkina, P. Thompson, A New Model for Probabilistic Information Retrieval on the Web, *Second SIAM International Conference on Data Mining (SDM 2002) Workshop on Web Analytics*, Arlington, Virgina, 2002.
40. J. Burrows, Word Patterns and Story Shapes: The Statistical Analysis of Narrative Style, *Literary and Linguistic Computing*, vol. 2, 1987, p. 61-70.
41. J. Rao and P. Rohatgi, Can pseudonymity really guarantee privacy? *Proceedings of the 9th USENIX Security Symposium* Denver, Colorado August, 2000, p. 14-17.
42. International Association of Forensic Linguists <http://www.iafl.org/> and *Forensic Linguistics: The International Journal of Speech, Language and the Law*, Birmingham, UK: University of Birmingham.
43. D. Foster *Policing Anonymity* available at <http://www.policefoundation.org/pdf/foster_anonymity.pdf.
44. J. Pennebaker, M. Francis, and R. Booth, *Linguistic Inquiry Word Count (LIWC): LIWC2001*, Erlbaum, Mahwah, New Jersey, 2001.
45. R. Rieber, and W. Stewart (eds.), *The Language Scientist as Expert in the Legal Setting,* Annals of the New York Academy of Sciences vol. 606, New York: The New York Academy of Sciences, 1990.
46. G. McMenamin and D. Choi (eds.), *Forensic Linguistics: Advances in Forensic Stylistics*, CRC, Boca Raton, Florida, 2002.

47. R. Shuy, *The Language of Confession, Interrogation, and Deception*, SAGE Publications, Thousand Oaks, California, 1998.
48. T. Grant, Ph. D. thesis, Forensic Section, School of Psychology University of Leicester, 2004 (upcoming publication).