

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Measurement Model for Deception Detection in Online Chat Software

¹Ade Adi Wibowo, ¹Zarina Shukur and ²Rozmi Ismail

¹School of Information Technology,

²School of Psychology and Human Development, National University of Malaysia, Malaysia

Abstract: Previous studies in psychology and linguistics revealed that people who committed fraud in online chats would show certain cues in their text/words. For example, deceptive senders displayed higher non-immediacy and expressiveness; used more words but in less complexity and less use of self-reference and diversity in their messages compared to truthful senders. Nevertheless, previous researchers did not explain how much was deemed to be less or more, compared to truthful senders. Such previous studies only determined whether a message was deceiving or not. Therefore, in order to detect deceptive senders, a deception detection model has been developed. It was integrated into a newly developed chatting software to analyse the degree of trust in chatting partners. An experiment was conducted to test the level of user trust in the developed software. The results of the experiment indicated that the software analysis produced an 11.36% possibility that a participant has been deceived. Although there was only a 11.36% chance that a chatting partner had been deceived, this proved that fraud had been committed by the chat partner. The experiments further found that the participants believed in the given results of the analysis and this affected their decision.

Key words: Online chatting, deception detection, detecting lies, chatting dangers, deception model, computer-mediated communication

INTRODUCTION

The world of Internet today is not far away from social networks. One way to communicate in the cyber world is to use tools such as Instant Messengers (IM), social networking platforms and email. Basically, IM occur in real-time and are not supported by transaction control. However, with more advanced technology, IM have added more functions that allow users to see each other via webcams, or talk directly for free over the Internet using a microphone and headphones (McZeal, 2004).

Although, IM have many benefits, they also have some risks in their usage. These risks include security risks (e.g., IM is used to infect computers with spyware, viruses, trojans, worms, etc.), compliance risks, inappropriate use, leakage of secrets, etc., (Kim and Leem, 2005; Hindocha and Chien, 2003).

In the virtual world, people are not limited by the constraints placed on them in the real world; this causes many cases of false identity and impersonation. In many cybercrimes, the criminal uses services such as chat, forums, blogs and IM programmes, to commit crime against children, fraud and identity theft. In particular, there are many cases of people pretending to be someone else and taking advantage of the anonymity provided by the Internet to commit crimes (Kontostathis *et al.*, 2010; Blair, 2003).

In a study by Hancock *et al.* (2004) it was found that the rate of lies occurring in an IM conversation is about the same as a face-to-face conversation. Approximately one-fifth of IM conversations involve a lie. In an experiment conducted by De Turk and Miller (1985) they found six indicators that were related to fraudulent communication. The indicators were message period, response latency, adapter, pause, non-fluencies and h and signals.

The increased use of online chats has opened up greater space for people with bad intentions to commit deception. Since, detection of deception through body language cannot be used in online chats, further studies should be conducted to detect deception within the text format. Although, psychological studies in detecting deception through body language (i.e., body movements, eye movements and voice analysis) are nearly stable, there is still room for studies in deception detection through text analysis.

BACKGROUND

Definitions of deception: Many definitions of deception have been suggested by previous researchers. For example, lying is saying what you believe is not true when you believe that the following norms in the conversation

on effects: “Do not say what you believe is not true” (Grice, 1989). This means that when you are sure that something is not true, then don’t say it; if you say it, then you are lying. Wallace (1995) also said that deception included “a scheme designed to deceive”.

Someone who wants to deceive will plan their deception. That person intends to commit a fraud.

Burgoon and Buller (1996) said that deception was an intentional transmission of information that aimed to foster a false conclusion or belief in the receiver. This means that if you are successful in making a wrong conclusion in the receiver, then you have succeeded in deceiving him h^{-1} . This definition has been referred to by many researchers.

Meanwhile, Amos (2008) said that in practice, not all that was said in a conversation topic of deceit was a lie. Some are truths that are used to support the deception as a whole.

George and Carlson (1999) also suggested that deception messages that are sent via email, chat, or instant messages, should be more difficult to detect than those delivered by non-text media, such as telephone or face-to-face communications.

Cues of deception in the text: There are many current studies, where researchers have examined the cues of deception that can be obtained from text. These cues are summarised in Table 1.

From these findings Zhou *et al.* (2003, 2004a, b), Zhou (2005), Zhou and Zhang (2007) and Zhou and Sung (2008) clustered together nine linguistic constructs that are useful for detecting deception in text; namely, affect, complexity, diversity, expressiveness, non-immediacy, quantity, specificity, uncertainty and informality.

Based on Table 1, previous researchers did not determine how much deception was considered to be more or less when compared to truthful senders. Previous researchers only determined whether a message was deceiving or not but did not explain how much of the message was deceiving. Therefore, the author suggested an idea to give a percentage of possibility that a message was deceiving.

In order to do that, a measurement model for deception detection has been developed. A linguistic construct was adapted in the attempt to build the model. To make full use of the model, it was integrated into a newly developed chatting software. It would analyse the degree of trust in chatting partners. It was also meant to give early warnings to chatters about their chatting partners based on their communication.

Table 1: Tendency cues of deception

Cues	Tendency	Authors
Words	L>T	Zhou <i>et al.</i> (2003, 2004a), Hancock <i>et al.</i> (2005) and Humpherys (2010)
Verbs	L>T	Zhou <i>et al.</i> (2003, 2004a), Humpherys (2010)
Second person pronouns	L>T	Vrij (2000), Matthew <i>et al.</i> (2003), Zhou <i>et al.</i> (2004a) and Humpherys (2010)
Third person pronouns	L>T	Vrij (2000), Matthew <i>et al.</i> (2003) and Zhou <i>et al.</i> (2004a)
Expressivity, quantity, non-immediacy, affect	L>T	Burgoon <i>et al.</i> (2003)
Informal	L>T	Burgoon <i>et al.</i> (2003) and Zhou <i>et al.</i> (2003,2004a)
Uncertainty, details of spatio-temporal	L>T	DePaulo <i>et al.</i> (2003), Matthew <i>et al.</i> (2003), and Tiantian <i>et al.</i> (2005)
Group reference	L>T	Zhou <i>et al.</i> (2003, 2004a) and Humpherys (2010)
Other reference	L>T	Hancock <i>et al.</i> (2005)
First person singular pronouns	L<T	Wiener and Mehrabian (1968), Vrij (2000), Matthew <i>et al.</i> (2003) and Zhou <i>et al.</i> (2004a)
Sentences, share detailed information	L<T	Burgoon and Buller (1996)
Exclusion words	L<T	Vrij (2000), Matthew <i>et al.</i> (2003) and Zhou <i>et al.</i> (2004a)
Complexity	L<T	Burgoon <i>et al.</i> (2003), DePaulo <i>et al.</i> (2003), Matthew <i>et al.</i> (2003) Tiantian <i>et al.</i> (2005) and Zhou and Sung (2008)
Diversity	L<T	Burgoon <i>et al.</i> (2003), DePaulo <i>et al.</i> (2003), Matthew <i>et al.</i> (2003), Tiantian <i>et al.</i> (2005) and Humpherys (2010)
Specifications	L<T	Burgoon <i>et al.</i> (2003)

*L: Liars; T: Truthful senders

METHODS

Measurement of deception detection model: This study used “Message Feature Mining” (Adkins *et al.*, 2004) to detect deception. Message Feature Mining is a method for classifying messages based on intent by using content-independent message features, in which the message features chosen are determined by the context of the classification. The selected features should follow the proven cues of deception. The selected features were adapted from the previously described linguistic constructs. Steps to find the percentage of the possibility that a message is deceitful or not will be explained in the following section.

Process

Step 1: The first step is to determine the selected features in the text that showed the cue that someone was deceitful or not. These features were partially adapted from (Zhou *et al.*, 2003, 2004b). The selected features are grouped into five categories; as shown in Table 2.

Step 2: As shown in Table 2, there is a need to find the total number of words, sentences, characters, clauses, words in a noun phrase, unique words, passive verbs, second-person pronouns, first-person singular pronouns,

Table 2: Selected cues to deception

Features	Cues	Formula
Complexity	Average sentence length	Total No. of words
		Total No. of sentences
	Average word length	Total No. of characters
		Total No. of words
	Average No. of clauses	Total No. of clauses
		Total No. of sentences
	Average length of noun phrase	Total No. of words in noun phrases
		Total No. of noun phrases
Diversity	Lexical diversity	Total No. of unique words
Non-immediacy	Passive voice ratio	Total No. of passive verbs
	Your reference ratio	Total No. of verbs
		Total No. of second person pronouns
		Total No. of words
	Self reference ratio	Total No. of first person plural pronouns
		Total No. of words
	Group reference ratio	Total No. of third person pronouns
		Total No. of words
	Other reference ratio	Total No. of third person pronouns
		Total No. of words
Expressivity	Emotiveness	Total No. of adjectives+
Informality	Typographical error ratio	Total No. of adverbs
		Total No. of nouns+
		Total No. of verbs
		Total No. of misspelled words
		Total No. of words

first-person plural pronouns, third-person pronouns, adjectives, adverbs, nouns, verbs and misspelled words in the conversation. Then the values of each cue are calculated using the formula shown in Table 2.

Step 3: As stated earlier, previous researchers did not explain how much was said to be more and how much was said to be less, compared to those who are truthful persons. Therefore, this study will determine if the person who told lies is less complex than those who are truthful. Then, the complexity of each cue is determined as less than 50% and vice versa, if the person who speaks lies is more informal than those who are truthful. Then, the value of informality is determined as more than 50%.

To determine whether each cue is greater or less than 50%, we used the formula shown in Table 3. The cues for an average sentence length, average word length, average number of clauses and average number of noun phrases, would be compared with the same cues from their chat partners.

Step 4: As shown in Table 2, there are 12 cues of deception in the text. Assuming that all cues are important; the weight of each cue is 8.33% (100%/12).

Table 3: Cues percentage equation

Cues	Equation
Average sentence length	Average sentence length (sender)×100% Average sentence length (receiver)
Average word length	Average word length (sender)×100% Average word length (receiver)
Average No. of clause	Average No. of clause (sender)×100% Average No. of clause (receiver)
Average No. of noun phrase	Average No. of noun phrase (sender)×100% Average No. of noun phrase (receiver)
Lexical diversity	Lexical diversity×100%
Self reference ratio	Self reference ratio×100%
Passive voice ratio	Passive voice ratio×100%
Your reference ratio	Your reference ratio×100%
Group reference ratio	Group reference ratio×100%
Other reference ratio	Other reference ratio×100%
Emotiveness	Emotiveness×100%
Typographical error ratio	Nisbah kesilapan menaip×100%

Lexical Diversity: Lexical Diversity×100%, Self Reference Ratio: Self Reference Ratio×100%, Passive Voice Ratio: Passive. Voice Ratio×100%
Your Reference Ratio: Your Reference Ratio×100%, Group Reference Ratio: Group Reference Ratio×100%, Other Reference Ratio: Other Reference Ratio×100%, Emotiveness, Emotiveness×100%, Typographical Error Ratio: Nisbah Kesilapan Menaip×100%

After the percentage of each cue is calculated using the equation shown in Table 3, the weight percentage of each cue obtained is measured using the equation shown in Table 4.

Step 5: After the percentage of each cue is found, they are summed to calculate the total. The sum total becomes the percentage of the possibility that the conversation partner is deceitful.

Experiment design: This experiment used the modified problem scenario of life in the desert (Desert Survival Problem by Lafferty and Eady (1974)). This scenario has been widely used by previous researchers (Adkins *et al.*, 2004; Fuller *et al.*, 2009; Twitchell, 2005; Zhou *et al.*, 2003, 2004a, b; Zhou and Zhang, 2007; Zhou *et al.*, 2008; Zhou and Sung, 2008) to collect deception data within text. This shows that this scenario is popular in collecting deception data and therefore aligned with the needs of this study. Since this is an early stage, the experiment used a small sample size to test the current prototype. If the prototype reaches stability, a larger sample size would be used to test the prototype in the future.

Participants: Participants (N = 20) were students from the Faculty of Economics, University of North Sumatera (Male = 7, Female = 13), who received additional marks for participating in the experiment.

Procedure: The entire experiment took place in an Economic Faculty classroom, which had a wireless Internet connection. Each participant was required to bring their own laptop.

Table 4: Cue weight percentage equation

Features	Cues	Condition	Equation
Complexity	Average sentence length	If the value of this cue is < 50%, then do the equation besides, if otherwise, then the weight percentage of these cues are 0%	$((50\% \times \text{Cues})) / (50\% \times 8.33\%)$
	Average Word Length		
Diversity	Average No. of Clause	If the value of this cue is > 50%, then do the equation besides, if otherwise, then the weight percentage of these cues are 0%	$((\text{Cues})) / (100\% \times 8.33\%)$
	Average No. of Noun Phrase		
Non-immediacy	Lexical Diversity		
Non-immediacy	Self Reference Ratio		
Expressivity	Passive Voice Ratio		
	Your Reference Ratio		
	Group Reference Ratio		
	Other Reference Ratio		
Informality	Emotiveness		
	Typographical error ratio		

The participants completed the experiment in the classroom by logging into the developed chat software. Randomly paired, participants were randomly given the roles of either “Sender” or “Receiver”. They communicated with their friends using text only via., the chat software. The experiment session took between 50-80 min to complete. After that, participants were asked to complete questionnaires.

Case scenario: Participants were informed that they would participate in the study of decision making that was modified from the problems of life in the desert (i.e., the Desert Survival Problem). This study asked the participants to imagine that they had a jeep which crashed in the desert of Kuwait and that there was no clean water and there were some items that could be repaired. The items were determined by the researchers.

Role of the receiver: Participants who played the role of the Receiver were told to define the rankings from the 12 items (e.g., mirror, compass, knife) for their survival. They were asked to rank from number 1-12, with 1 as the most important item to number 12 as the least important item to have. Before interacting with their partners (i.e., the Sender), all participants were asked to read a detailed document on what was needed to survive in the desert. This information was taken from military field manuals (available at: <http://rk19-bielefeld-mitte.de/survival/FM/13.htm>). They were used to form the basis of the ranking and the discussion.

First, participants (Receiver) were asked to determine their initial position freely, following which they started to discuss with their partners (Sender) on why their position (Receiver) was like that, until at its final position to reach an agreement. Participants (Receiver) would also be informed that their chatting partner might intend to deceive and they were told that they could use the tools provided in the chatting software to measure the possibility that their chat partner was being deceitful. Then they completed their final ranking freely and completed the questionnaires given in the chatting software.

Role of the sender: Participants who were given the role of “Sender” were tasked to deceive, mislead and give invalid information to the “Receiver”. After reading the instructions of the general task, “Sender” (or liar) received additional instructions as follows:

“In addition, the main objective of this study is to learn how people can detect false, misleading and deceptive information in communication. There are often situations where it is NOT in the best interests of a person to tell “the truth, the whole truth and nothing but the truth”-for example, to avoid unpleasant circumstances, to protect your loved ones, or to protect your country. A certain level of communication skill is necessary to be able to adapt to this situation. Your instructor will ask you to mislead your partner so that the instructor can determine how well your partner can detect your deception. If you have any objections to performing this role, please notify the experimenter at this time”.

After reading the documents, they were asked to list the ranking “contrary to what is recommended by the experts and contrary to what you believe is true and correct information”. To strengthen this manipulation, they were told that their task was to provide incorrect, misleading and deceptive information to their partners and that they should argue that communication tools such as a mirror and a flashlight were useless in the desert and propose to leave all unnecessary clothing (e.g., rain-equipment) and equipment to make walking easier.

They were also told that there were many ways to be dishonest, including telling blatant lies, exaggerating, vague, indirect, unclear and ambiguous messages, or leave and avoid discussing the relevant information. In short, they could use their own techniques and communication style to deviate from the “truth, the whole truth and nothing but the truth”. Finally, they were told that their team members did not know that they received special instructions and it was important not to disclose this information to their team member both during and after the discussions.

RESULTS

Demographic results: As explained previously, there were seven men and 13 women who participated in this experiment. From the questionnaires completed by the participants, all participants were aged between 20-30 years old. The frequency of participants using online communication with other people in one day is summarised in Table 5.

The table shows that the amount of time that the majority of participants use online communication with others was not more than two hours per day.

Deceiver motivation: To confirm that the participants, who were given the role as “Sender”, were actually deceiving and to test the participants' motivation to deceive, four questionnaires that were rated on a scale of 1-10 were given. The average score of three categories were determined-low (1.0-3.3), moderate (3.4-6.73) and high (6.74-10.0).

The reliability test of the questionnaire showed that the Cronbach's Alpha (Cronbach, 1951) from the four questions was 0.933. Because the value was above 0.7, this questionnaire was considered as reliable.

The result of this experiment show that participants who played the role of “Sender” had a strong motivation to succeed in deceiving their chat partner (average score = 6.90).

Table 6 shows the average scores of the four questionnaires used to test the motivation of participants to deceive their chat partner during the discussion.

Deception detection information by receiver: To examine the level to which participants (Receiver) could determine whether their respective chat partner was trying to

Table 5: Frequency of using online communications

Item	Frequency	Percent
Less than 1 h	9	45.0
More than 1 h but less than 2 h	8	40.0
More than 3 h but less than 4 h	2	10.0
More than 6 h	1	5.0
Total	20	100.0

How much time a day that you use to communicate online with others?

Table 6: Deceiver Motivation

Item	Average
On a scale of 0 to 10, how important was it for you to give convincing answers to your partner during the discussion about the best items to salvage and the best course of action to survive in the desert?	6.80
On a scale of 0 to 10, how important was it for you to keep your partner from becoming suspicious during the discussion?	6.70
On a scale of 0 to 10, how important was it for you to succeed in making your partner believe you?	7.20
On a scale of 0 to 10, how truthful were you in representing your true rankings and reasons to your partner?	6.90
Average score	6.90

deceive them or not, the researchers gave them nine questionnaires to complete. The result of Cronbach's Alpha from the nine questions was 0.894 and because the value was above 0.7, this questionnaire was deemed as reliable.

Table 7 indicates the value of the average scores of the nine questions. For questions numbered 1 and 5, they were rated on a scale of 1-2. Question numbers 2, 3 and 4 were rated on a scale of 1-10, while question numbers 6-9 were rated on a scale of 1-7.

In the questionnaires that had a value on a scale of 1-7, the researchers determined the average score of three categories which was low (1.0-2.33), moderate (2.34-4.67) and high (4.68-7.0).

Table 7 shows that overall, the participants in this experiment believed in the results given from the analysis provided by the application (average = 5.80) and had some impact on their decision (average = 6.60). However, this table also shows that overall, they were not able to determine whether their chat partner was deceiving (average = 1.90) and they were very confident (mean = 4.80) with their decision. Overall, this discussion went smoothly (average = 4.40) and participants were very comfortable during the discussion (5.00) and understood their chat partner (4.20).

Table 8 shows how many participants, who were given the role of “Receiver” answered the correct answer in sorting the 12 items from the most important to the least important item.

Table 8 shows that the participants could not sort the items from the most important to the least important (Average = 1.00). This proves that the participants who were given the role of “Sender” had been successful in deceiving their partners who were given the role of “Receiver”.

Table 9 shows a comparison between the questions numbered 4, 5 and 6 (Table 7) with the percentage of possibilities that their chat partner was deceiving which were produced by the developed application. This is aimed to compare the users' perception towards their chat

Table 7: Information deception detection

Item	Average	Results
Have you ever used deception detection applications in chat before? (1) Yes (2) No	1.80	No
The level you believe in the analysis provided by the application?	5.80	Moderate
The effectiveness of the analysis provided by the application to your decision making?	6.60	Moderate
The level you believe that your partner is an honest person in the discussion?	5.60	Moderate
If you have to make decisions, your friend: (1) Lying (2) Truthful	1.90	Truthful
I am confident in my judgment whether my partner is an honest person or not?	4.80	High
This discussion went smoothly?	4.40	Moderate
I feel comfortable during this discussion?	5.00	High
I fully understand my partner in this discussion?	4.20	Moderate

Table 8: A score of sorting 12 items

Username	Score
Receiver 1	2/12
Receiver 10	2/12
Receiver 11	2/12
Receiver 15	2/12
Receiver 17	1/12
Receiver 19	2/12
Receiver 3	0/12
Receiver 6	0/12
Receiver 9	0/12
Trial receiver 1	1/12
Average	1.00

Table 9: Comparison of user's perceptions with analysis produced by the applications

Username	Q 4	Q 5	Q 6	System results
Receiver 1	9 (High)	2 (Truthful)	6 (High)	13.90% (Low)
Receiver 10	5 (Moderate)	1 (Liar)	3 (Moderate)	16.30% (Low)
Receiver 11	6 (Moderate)	2 (Truthful)	7 (High)	15.03% (Low)
Receiver 15	5 (Moderate)	2 (Truthful)	3 (Moderate)	9.70% (Low)
Receiver 17	5 (Moderate)	2 (Truthful)	5 (Moderate)	0.0% (Low)
Receiver 19	1 (Low)	2 (Truthful)	4 (Moderate)	10.03% (Low)
Receiver 3	10 (High)	2 (Truthful)	7 (High)	10.70% (Low)
Receiver 6	6 (Moderate)	2 (Truthful)	4 (Moderate)	12.20% (Low)
Receiver 9	4 (Moderate)	2 (Truthful)	4 (Moderate)	14.53% (Low)
Trial	5	2	5	11.16%
Receiver1	(Moderate)	(Truthful)	(High)	(Low)
Average	5.60	1.90	4.80	11.36%
	(Moderate)	(Truthful)	(Moderate)	(Low)

partner was being deceitful and the applications' perceptions on which level of their chat partner was being deceiving.

This table shows that the participant with username 'receiver 1' strongly believed (value = 9) that their chat partner was an honest person and had decided that the chat partner was an honest person (value = 2) and he was very confident of his judgment (value = 6). This was consistent with the results shown by the analysis provided from the application that their chat partner had lied at a low level (percentage = 13.90%).

Overall, participants' perception of their chat partner was of someone who was honest (average = 5.60) and they had to make decisions that their chat partner was someone who was honest (average 1.90) and they were confident in their judgment (Mean = 4.80). The percentage provided by the application also showed that the level of their chat partners being deceitful was shown at a low level (average = 11.36%).

DISCUSSION AND CONCLUSION

Based on the results obtained, participants who were given the role of "Sender" or given the task to deceive, mislead and give invalid information to the participants given the role of "Receiver" was a high motivation (average = 6.90) to convince them. This shows that the participants did a good job. However, from the perceptions of participants who were given the role of "Receiver" it appeared that they were not able to determine whether their chat partner was someone who was deceitful. This is justified by the low average value of 12 for participants who could sort the most important items to the least important (average = 1.00).

Overall, the analysis produced by the application shows that the percentage of the possibility a chat participant (Receiver) was deceiving amounted to 11.36%. When compared to a medical field, if stated that the patient has a 10% risk of cancer, this means that despite a 10% chance of cancer, it may be very risky for the patient. This is the same as the analysis produced by the application that even though an 11.36% chance of a chat partner was being deceitful; this proves that there was a lie committed by the chat partner. The possibility that the chat partner was deceitful is indicated by the percentage of 11.36%.

However, the analysis produced by the application shows that the level of possibilities that the chat partner was being deceitful was at a low level (average percentage = 11.36%). This indicates that this application could not accurately detect the possibility that the chat partner was being deceitful.

This probably occurs due to the lack of decisions to determine how much is said to be more and how much is said to be less, compared to people who are honest. Previous studies only stated that a message is deceiving or not (Zhou *et al.*, 2003; 2004a, b; Zhou, 2005; Zhou and Zhang, 2007; Zhou and Sung, 2008; Twitchell, 2005). Thus, the model to measure this deception determined that when the cue was said to be more than that of honest people, then the percentage value was more than 50% and vice versa, if the cue was said to be less than someone who was honest, then the percentage value was less than 50%. The value of this percentage may not be suitable for each cue. Therefore, it needs further study to determine the value of percentage of this deception detection to measure more accurately.

Previous studies only tested the effectiveness of a technique in detecting deception in text (Burgoon *et al.*, 2003; Depaulo *et al.*, 2003; Fuller *et al.*, 2009; Humpherys, 2010; Twitchell, 2005; Zhou and Sung, 2008; Tiantian *et al.*, 2005) but did not test how effective

the technique was that was applied to the software (e.g., online chat software) could help users to make decisions.

Therefore, the results from this experiment showed that the users believed in the analysis provided and affected their decision. This indicates that the facilities provided in this chat software prototype could help them to make the best decision.

REFERENCES

- Adkins, M., D.P. Twitchell, J.K. Burgoon and J.F. Nunamaker, 2004. Advances in automated deception detection in text-based computer-mediated communication. Proceedings of the Enabling Technologies for Simulation Science VIII, August 13, 2004, Orlando, FL, USA. .
- Amos, B., 2008. Are you lying now? A linguistic examination of deceptive utterances in online conversation. Honors Thesis, College of Agriculture and Life Sciences.
- Blair, J., 2003. New breed of bullies torment their peers on the internet. *Educ. Week*, 22: 6-9.
- Burgoon, J.K. and D.B. Buller, 1996. Interpersonal deception theory. *Commun. Theory*, 6: 311-328.
- Burgoon, J.K., J.P. Blair, T. Qin and J.F. Nunamaker Jr., 2003. Detecting deception through linguistic analysis. *Intell. Secur. Inform.*, 2665: 91-101.
- Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16: 297-334.
- De Turck, M.A. and G.R. Miller, 1985. Deception and arousal: Isolating the behavioral correlates of deception. *Hum. Commun. Res.*, 12: 181-201.
- DePaulo, B.M., J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton and H. Cooper, 2003. Cues to deception. *Psychol. Bull.*, 129: 74-118.
- Fuller, C.M., D.P. Biros and R.L. Wilson, 2009. Decision support for determining veracity via linguistic-based cues. *Decis. Support Syst.*, 46: 695-703.
- George, J.F. and J.R. Carlson, 1999. Group support systems and deceptive communication. Proceedings of the 32nd Annual Hawaii International Conference on System Sciences, Jun 5-8, 1999, Maui, HI..
- Grice, P., 1989. *Studies in the Way of Words*. Harvard University Press, Cambridge, ISBN: 9780674852716, Pages: 394.
- Hancock, J.T., J. Thom-Santelli and T. Ritchie, 2004. Deception and design: The impact of communication technologies on lying behavior. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 24-29, 2004, Vienna, Austria, pp: 129-134.
- Hancock, J.T., L. Curry, S. Goorha and M. Woodworth, 2005. Automated linguistic analysis of deceptive and truthful synchronous computer-mediated communication. Proceedings of the 38th Hawaii International Conference on System Sciences, January 3-6, 2005, Big Island, HI, USA., pp: 22-23.
- Hindocha, N. and E. Chien, 2003. Malicious threats and vulnerabilities in instant messaging. <https://www.symantec.com/avcenter/reference/malicious.threats.instant.messaging.pdf>
- Humpherys, S., 2010. A system of deception and fraud detection using reliable linguistic cues including hedging, disfluencies and repeated phrases. Ph.D. Thesis, The University of Arizona.
- Kim, S. and C. Leem, 2005. Implementation of the security system for instant messengers. Proceedings of the 1st International Symposium on Computational and Information Science, December 16-18, 2004, Shanghai, China, pp: 739-744.
- Kontostathis, A., L. Edwards and A. Leatherman, 2010. Text Mining and Cybercrime. In: *Text Mining: Applications and Theory*, Berry, M.W. and J. Kogan (Eds.). John Wiley and Sons, Ltd., Chichester, UK.
- Lafferty, J.C. and P.M. Eady, 1974. The desert survival problem. Plymouth, MI: Experimental Learning Methods. <http://beatruiusbox.files.wordpress.com/2008/02/desert-survival.pdf>
- Matthew, N.L., J.W. Pennebaker, D.S. Berry and J.M. Richards, 2003. Lying words: Predicting deception from linguistic styles. *Personality Soci. Psychol. Bull.*, 29: 665-675.
- McZeal Jr. A., 2004. Multifunctional world wide walkie talkie, a tri-frequency cellular-satellite wireless instant messenger computer and network for establishing global wireless VOLP Quality of Service (QoS) communications, unified messaging, and video conferencing via the internet. USPTO Patent Full-Text and Image Database, U.S., Patent 6763226, Application No. 10210480.
- Tiantian, Q., J.K. Burgoon, J.P. Blair and J.F. Nunamaker, 2005. Modality effects in deception detection and applications in automatic deception detection. Proceedings of the 38th Annual Hawaii International Conference on Systems Sciences, January 3-6, 2005, University of Arizona.
- Twitchell, D.P., 2005. Automated analysis techniques for online conversations with application in deception detection. Ph.D. Thesis, The University of Arizona.

- Vrij, A., 2000. *Detecting Lies and Deceit: The Psychology Of Lying and its Implications for Professional Practice*. 1st Edn., John Wiley and Sons, Chichester, ISBN-13: 9780471853169, Pages: 276.
- Wallace, W.A., 1995. *Auditing*. South-Western College Publishing, Cincinnati, OH..
- Wiener, M. and A. Mehrabian, 1968. *Language Within Language: Immediacy, A Channel in Verbal Communication*. Ardent Media, New York, Pages: 214.
- Zhou, L., D.P. Twitchell, T. Qin, J.K. Burgoon and J.F. Nunamaker, 2003. An exploratory study into deception detection in text-based computer-mediated communication. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, January 6-9, 2003, Baltimore, MD, USA..
- Zhou, L., J.K. Burgoon, D.P. Twitchell, T. Qin and J.F. Nunamaker, 2004a. A comparison of classification methods for predicting deception in computer-mediated communication. *J. Manage. Inform. Syst.*, 20: 139-165.
- Zhou, L., J.K. Burgoon, F. Jay, J.R. Nunamaker and D. Twitchell, 2004b. Automated linguistics based cues for detecting deception in text-based asynchronous computer-mediated communication: An empirical investigation. *Group Decis. Negotiation*, 13: 81-106.
- Zhou, L., 2005. An empirical investigation of deception behavior in instant messaging. *IEEE Trans. Prof. Commun.*, 48: 147-160.
- Zhou, L. and D. Zhang, 2007. Typing or messaging? modality effect on deception detection in computer-mediated communication. *Decis. Support Syst.*, 44: 188-201.
- Zhou, L., Y. Shi and D. Zhang, 2008. A statistical language modeling approach to online deception detection. *IEEE Trans. Knowl. Data Eng.*, 20: 1077-1081.
- Zhou, L. and Y.W. Sung, 2008. Cues to deception in online Chinese groups. *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, January 7-10, 2008, IEEE, pp: 146.