Fr. Conceicao Rodrigues College Of Engg., Bandra (W)

BEIT (SEM VII)

Project Proposal Form

1. **Name of the student(s) with roll numbers**
   - Vidhi Bhansali (Roll No. 6118)
   - Sayli Kapse (Roll No. 6137)
   - Tanvi Shah (Roll No. 6264)
2. **Title of the Project:**

   "Detecting Word Substitutions In Text"
3. **Project Category:** [Tick $\surd$ wherever applicable]

Software	$\surd$
Hardware	
Software and Hardware	
Checked By:	**Date:**

4. **Project Area:** [Tick $\surd$ wherever applicable]

Cloud Computing		**DSP OR Image Processing**	
Web Technology		**μC and Embedded Systems**	
Evolutionary Algorithms		**Micro wave OR Fiber Optics**	
Network Security		**Wireless Communication**	
Fuzzy Logic OR Genetics OR Neural Networks		**Robotics OR Instrumentation**	
Mobile Application		**Computer Networks**	
Artificial Intelligence		**Communication Systems**	
Others (Please Specify)	$\surd$	**Computer Graphics**	
Data Mining			

5. **Project Abstract:**

Searching for words on a watch list is one way in which large-scale surveillance of communication can be done, for example in intelligence and counterterrorism settings. One obvious defense is to replace words that might attract attention to a message with other, more innocuous, words. For example, the sentence the attack will be tomorrow" might be altered to the complex will be tomorrow", since 'complex' is a word whose frequency is close to that of 'attack'. Such substitutions are readily detectable by humans since they do not make sense. We address the problem of detecting such substitutions automatically, by looking for discrepancies between words and their contexts, and using only syntactic information. We define a set of measures, each of which is quite weak, but which together produce per-sentence detection rates around 90 with false positive rates around 10. Rules for combining per sentence detection into per-message detection can reduce the false positive and false negative rates for messages to practical levels. We test the approach using sentences from the Enron email and Brown corpora, representing informal and formal text respectively.

6. **Project Overview:**

   When messages may be intercepted because they contain certain words, terrorists and criminals may replace such words by other words or locutions. If the replacement words have different frequencies from the original words, techniques to detect the substitution are known. In this, we consider ways to detect replacements that have similar frequencies to the original words. We consider the frequencies of generalized n-grams that we call kgrams, the change in frequency that results from removing the word under consideration from its context, and the change in frequency that results from replacing a word by its hypernym. In our preliminary results, none of these measures are strong individually, but together they become effective.
   One way to conceal content is to encrypt the messages, but this strategy has a number of drawbacks. First, encryption draws attention to messages, making techniques such as traffic analysis easier to apply. Second, encryption is hard to use with readily available components in some settings, for example cell phone calls. Third, it is hard to be sure exactly how robust encryption is in practice, since agencies such as the U.S. NSA do not reveal their decryption capabilities and there are persistent rumors of back doors into common encryption systems. Another strategy to conceal the content of messages is to replace significant words with other words or locutions that are judged less likely to attract attention. For example, it is known that Echelon scans for a list of significant words or phrases, and terrorists would presumably wish not to use these words in their messages. The difficulty is that, while it is clear that some words must be on these lists (e.g. 'nuclear'), it is difficult to guess how long such lists are. Replacing words with more innocuous words in real time, for example during a cell phone call, is not easy, and it is likely that the replacement words will differ in obvious ways from the words they replace. For example, humans do not appear to have an intrinsic understanding of word frequencies, so it is likely that a word and its replacement would have significantly different frequencies. However, replacement of words by words of similar frequency becomes possible given access to a word-frequency table or a codebook. In this, we examine whether such substitutions are detectable by software.

7. **Methodology / Programming Techniques to be used:**

   Matlab, Java, .Net

8. **What is Novelty / Innovative in the proposed project?:**

   Terrorists are aware that their communications are likely to be intercepted by systems such as Echelon, and would like to conceal the content of these communications as much as possible. Criminals also face similar issues since their communications may be intercepted by law enforcement.

9. **Name Of Project Guide assigned:** [To be filled by Head Of the Department]

Bibliography

[1]Bilmes, J.A., Kirchhoff, K.: Factored language models and generalized parallel backoff. In: Proceedings of HLT/NACCL (2003)

[2]Fong, S.W., Skillicorn, D.B., Roussinov, D.: Detecting word substitution in adversarial communication. In: Workshop on Link Analysis, Counterterrorism and Security at the SIAM International Conference on Data Mining, to appear (2006)

[3]Lee, H., Ng, A.Y.: Spam deobfuscation using a Hidden Markov Model. In: Proceedings of the Second Conference on Email and Anti-Spam (2005)

[4]Roussinov, D., Zhao, L.: Automatic discovery of similarity relationships through web mining. In: Decision Support Systems, pp. 149166 (2003)

[5]Roussinov, D., Zhao, L., Fan, W.: Mining context specific similarity relationships using the World Wide Web. In: Proceedings of the 2005

[6]Skillicorn, D.B.: Beyond keyword filtering for message and conversation detection. In: Kantor, P., Muresan, G., Roberts, F., Zeng, D.D., Wang, F.-Y., Chen, H., Merkle, R.C. (eds.) ISI 2005. LNCS, vol. 3495, pp. 231243. Springer, Heidelberg (2005)