Enhancing the Privacy of Web-based Communication

Aleksandra Korolova
Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304, USA
korolova@alum.mit.edu

Ayman Farahat
Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304, USA
farahat@parc.com

Philippe Golle
Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304, USA
pgolle@parc.com

ABSTRACT

A *profiling* adversary is an adversary whose goal is to classify a population of users into categories according to messages they exchange. This adversary models the most common privacy threat against web based communication.

We propose a new encryption scheme, called stealth encryption, that protects users from profiling attacks by concealing the semantic content of plaintext while preserving its grammatical structure and other non-semantic linguistic features, such as word frequency distribution. Given English plaintext, stealth encryption produces ciphertext that cannot efficiently be distinguished from normal English text (our techniques apply to other languages as well).

Categories and Subject Descriptors: C.2.0 Computer-Communication Networks

General Terms: .Security and Protection

Keywords: Privacy, profiling.

1. INTRODUCTION

Unencrypted email communication offers no privacy. The contents of email messages are exposed to a number of intermediaries between senders and receivers, such as web-based email providers, Internet service providers, backbone Internet routers. Only users with strong privacy needs are willing to bear the costs of encrypting their email messages. The communication of a typical email user is rarely of any value to an adversary, and is thus unlikely to be targeted in isolation. A whole population of users, on the other hand, may in aggregate attract the interest of an eavesdropping adversary. The real threat to privacy, therefore, comes from attacks which target a population of users rather than a specific individual. We call such attacks *profiling attacks* [5].

2. RELATED WORK.

While Stealth encryption bears some resemblance to lexical steganography techniques such as Mimic functions [6] or [2], it differs from these techniques in two major ways. First, lexical steganography is much less efficient (they produce ciphertext that is much larger than the plaintext, and can therefore be easily detected by an adversary). In addition, Stealth encryption does not require any key exchange (the lexical steganography [2] require the exchange of a style file).

3. BASIC STEALTH ENCRYPTION

In stealth encryption every word of English plaintext is replaced with a word of ciphertext drawn from an English dictionary. The mapping E_k between plaintext and ciphertext words is determined by the choice of a secret encryption key k.

We observe that pronouns, determiners, adverbs, prepositions and conjunctions contain little semantic information and thus need not be encrypted. We divide the content words into κ categories according to their morphological analysis. In order to preserve a Zipfian distribution of terms, we a word w to a new word w' such that $|Rank(w) - Rank(w')| \leq \alpha$ where α is a small integer (where rank is position in list of words sorted by frequency). In other words, stealth encryption operates like a *low-inversion permutation.*

The stealth encryption function consists of one secret α-low inversion permutation per grammatical category. To encrypt a word of plaintext, we compute the grammatical category to which it belongs and let the ciphertext be the image of the plaintext by the α-low inversion permutation corresponding to that category. Thus we obtain an encryption function that maps a plaintext word to a ciphertext word that belongs to the same category and has approximately the same frequency in standard English.

3.1 Example

Plaintext: Hi friend, Its been a long time since I last wrote. How are your kids doing? I am immensely enjoying my stay here at PARC and California is even more beautiful than I had imagined!
Ciphertext: Hi young, Its been a support power since I last flew. How are your prices doing? I am immensely occupying my blood here at IBM and Kansas is even more environmental than I had illustrated!

4. APPLICATIONS

The sender of an email proceeds as follows. Let M denote the message body of the email to be sent, and let H be the header of that email (consisting of the address of the sender, the address of the recipient, the time at which the email is sent, sender's IP address, and potentially other fields). For

WWW 2005, May 10–14, 2005, Chiba, Japan.
ACM 1-59593-051-5/05/0005.

simplicity, we assume that all the fields in H are known both to the sender and recipient of the email (if that is not the case, we may define H as a subset of the header that is known to both the sender and recipient).

Let φ be a slow one-way function [4], i.e. a publicly known function that is moderately costly to evaluate (say, on the order of a few seconds of computation) and very hard to invert. The sender computes $\varphi(H)$ by applying the slow one-way function to the header H, then generates a key k from the seed $\varphi(H)$ and encrypts the body of the email with E_k. The message sent is $H||E_k(M)$. Upon receiving this message, the recipient recovers the key k by computing $k = \varphi(H)$, then uses k to decrypt the encrypted body $E_k(M)$ of the email.

Our protocol dispenses with key exchange. Anyone can compute the decryption key from the header of the message. Whereas the intended recipient can decrypt a small number of messages at a relatively low computational cost, a profiling adversary attempting to decrypt a large number of messages would incur a tremendous computational cost. Furthermore, since our encryption scheme produces ciphertext that is machine indistinguishable from plaintext, the adversary can not separate encrypted communication from the rest.

5. ANALYSIS OF STEALTH ENCRYPTION

A parser is a sophisticated tool that uses grammatical information to identify a grammatically correct syntactic structure. An attacker could attempt to distinguish plaintext from ciphertext by running a parser on all sentences of a given message and comparing the number of sentences with no complete parse with the expected number of such sentences for a plaintext message. A sophisticated parser equipped with a good grammar incorporates most of the linguistic and statistical knowledge available about the syntax and semantics of the language and is a benchmark for many other possible attacks,

We performed our experiments on 20,000 postings to an Internet Newsgroup on the topics of computers, motors, politics, religion, sports, science, and "for sale". Parsing was performed using the XLE parser [1].

To identify whether the use of a low-inversion permutation and morphology-based categorization for encryption increases the probability of completely parsing the sentences in the ciphertext, we performed the encryption with different values of the low inversion parameter α and number of categories κ.

Table 1: Parsing Results

	Stat	Ciphertext		Plaintext
		$\alpha = 20,000$	$\alpha = 10$	
		$\kappa = 1$	$\kappa = 38$	
% sentences	avg.	74.50	57.51	49.42
with no parse	stdev	14.89	17.66	18.87

While presently plaintext and ciphertext can be distinguished using a parser-based analysis, doing so reliably would require substantial computational resources. Using the log likelihood ratio test [3], we estimate that to automatically distinguish a user sending emails from a user sending stealth encrypted emails , one would need to parse at least 500 sentences.

Table 2: Classification Results

	percent classified correctly	
	plaintext	ciphertext
average	77.61	25.36

To see whether stealth encryption inhibits and/or complicates profiling, we attempted to establish, whether a classifier that is reasonably successful in classifying plaintext messages would perform as well on the stealth encrypted versions of those messages. The classifier built was a decision-tree, that used words as its features.

As can be seen from Table 2, stealth encryption achieves its goal - while the plaintext is classified with 77.61% accuracy on average, the ciphertext is classified with 25.36% accuracy.

We compared the performance of Stealth encryption to the "nicetext" lexical steganography system [2]. The average size of the cyphertext generated by nicetext was 30 times the size of the orignial text while that generated by Stealth encryption was about the same size as the original text.

6. REFERENCES

[1] R. Kaplan, S. Riezler,, T.H. King, J.T. Maxwell, and A. Vasserman. Speed and accuracy in shallow and deep stochastic parsing. In *HLT-NAACL*, 2004.

[2] M. Chapman and G. Davida. Hiding the hidden: A software system for concealing ciphertext in innocuous text. In *ICIS*, volume 1334, Beijing, China, 11–14 1997.

[3] T. E. Dunning. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74, 1993.

[4] C. Dwork and M. Naor. Pricing via processing or combatting junk mail. In *CRYPTO 92 Proceedings*, volume 740 of *Lecture Notes in Computer Science*, 1992.

[5] P. Golle and A. Farahat. Defending email communication against profiling attack. In *Workshop on Privacy in the Electronic Society , WPES-2004*, October 2004.

[6] P. Wayner. Mimic functions. *CRYPTOLOGIA*, 16(3):193–214, July 1992.