J. Cryptology (1988) 1: 129–131

Journal of Cryptology

A Generalization of Hellman's Extension to Shannon's Approach to Cryptography[1]

Pierre Beauchemin and Gilles Brassard[2]
Département IRO, Université de Montréal,
C.P. 6128, Succ. "A", Montréal, Québec, Canada H3C 3J7

Abstract. In his landmark 1977 paper [2], Hellman extends the Shannon theory approach to cryptography [3]. In particular, he shows that the expected number of spurious key decipherments on length n messages is at least $2^{H(K)-nD} - 1$ for *any* uniquely encipherable, uniquely decipherable cipher, as long as each key is equally likely and the set of meaningful cleartext messages follows a uniform distribution (where $H(K)$ is the key entropy and D is the redundancy of the source language). Here we show that Hellman's result holds with no restrictions on the distribution of keys and messages. We also bound from above and below the key equivocation upon seeing the ciphertext. The results are obtained through very simple purely information theoretic arguments, with no need for (explicit) counting arguments.

Key words. Cryptography, Information theory, Key equivocation, Shannon theory, Spurious decipherments.

1. Definitions and Review of Hellman's Results

Consider any uniquely encipherable, uniquely decipherable, secret-key cryptosystem. Let K be the key space and let $H(K)$ be the key entropy. Let M and C be the cleartext and ciphertext message spaces, respectively. Assume that the cryptosystem is endomorphic (i.e., $M = C$) or at least that M and C have the same number of elements. For each $k \in K$, let $E_k: M \to C$ and $D_k: C \to M$ be the corresponding enciphering and deciphering functions.

For any $m \in M$, let $p(m)$ be the *a priori* probability of cleartext message m. For any $k \in K$, let $p(k)$ be the probability that key k will be used. From the above two distributions, we can easily define for any $c \in C$ the probability $p(c)$ that ciphertext c will be produced, the key and message equivocations $H(K|c)$ and $H(M|c)$ upon seeing c, and the set $K_c = \{k \in K | p(D_k(c)) > 0\}$ of possible keys that can account for c. Finally, the key equivocation is defined by $H(K|C) = \sum_c p(c)H(K|c)$.

Hellman uses $l(c)$ to denote $\# K_c$, the number of elements in K_c. He also denotes $\max\{(l(c) - 1), 0\}$ by $n_k(c)$, which is the number of spurious key decipherments that are possible upon seeing ciphertext c. The expected number of spurious key decipherments offered by the cryptosystem is given by $\bar{n}_k = \sum_c p(c)n_k(c)$.

[1] This work was first reported at the CRYPTO 87 workshop, Santa Barbara, CA, August 1987.

[2] Supported in part by Canada NSERC Grant A4107.

Let us now assume that the cryptosystem is used to send messages of length n. Let D stand for the redundancy of the source language. In order to get his results, Hellman assumes that each key is equally likely and that the cleartext messages split between those that are impossible and the meaningful ones, the latter being all equally likely. Under these assumptions, Hellman proves that $\bar{n}_k$ is very close to $2^{H(K)-nD}$ for a so-called random cipher and (which is much more interesting) that it is always at least $2^{H(K)-nD} - 1$ for *any* uniquely encipherable, uniquely decipherable cipher.

2. Our Result

It is very nice that $\bar{n}_k \geq 2^{H(K)-nD} - 1$ holds not only for random ciphers but for every cipher. However, we feel that the assumptions used by Hellman on the key and cleartext distributions are too strong. Although it is good practice to choose the keys with uniform distribution, we know that this is not always done (consider what happened with the ill-fated Enigma's so-called telegram keys [1]). Moreover, using $H(K)$ in this context makes the formula look unnecessarily esoteric: Hellman's theorem could be reformulated more simply as $\bar{n}_k \geq \# K/2^{nD} - 1$. As for the assumption on the meaningful message distribution, it is obviously unreasonable.

The point of this paper is that these assumptions are not needed to prove that $\bar{n}_k \geq 2^{H(K)-nD} - 1$.

Theorem.

$$H(K) - nD \leq H(K|C) \leq \log_2(\bar{n}_k + 1).$$

Proof. (i) In his paper [3], Shannon proves by purely information theoretic arguments that

$$H(K|C) = H(K) + H(M) - H(C).$$

By definition of redundancy, $H(M) = (\log_2 \# M) - nD$. By a general theorem about entropy, $H(C) \leq \log_2 \# C$. By assumption, $\# M = \# C$. It follows immediately that $H(K|C) \geq H(K) - nD$.

(ii) Consider any cryptogram c. Recall that $K_c = \{k \in K | p(D_k(c)) > 0\}$. Assume that $p(c) > 0$, hence $\# K_c > 0$. Recall that $n_k(c) = \# K_c - 1$. Again by a basic property of the entropy, $H(K|c) \leq \log_2 \# K_c = \log_2(n_k(c) + 1)$. Hence

$$\begin{aligned} H(K|C) &= \sum_c p(c) H(K|c) \\ &\leq \sum_c p(c) \log_2(n_k(c) + 1) \\ &\leq \log_2\left(\sum_c p(c)(n_k(c) + 1)\right) \qquad \text{(by Jensen's lemma)} \\ &\leq \log_2(\bar{n}_k + 1). \end{aligned}$$

□

Corollary. *The inequality* $\bar{n}_k \geq 2^{H(K)-nD} - 1$ *holds for any uniquely encipherable, uniquely decipherable, endomorphic, secret-key cryptosystem.*

Note. Using similar techniques, we can also show that:

(1) $H(K|C) = H(K) - nD$ if and only if the probability distribution on the ciphertexts is uniform.

(2) $H(K|C) = \log_2(\bar{n}_k + 1)$ if and only if there exists a positive integer q such that, for all $c \in C$ such that $p(c) > 0$,

(i) $\# K_c = q$, and

(ii) the conditional key probability is uniform: $p(k|c) = 1/q$ for each $k \in K_c$.

References

[1] Garlinski, J., *Intercept, the Enigma War*, Dent, London, 1979.

[2] Hellman, M. E., An extension of the Shannon theory approach to cryptography, *IEEE Transactions on Information Theory*, vol. IT-23, 1977, pp. 289–294.

[3] Shannon, C. E., Communication Theory of Secrecy Systems, *Bell System Technical Journal*, vol. 28, 1949, pp. 656–715.