Primitives

Let P_X be a probability mass function with support χ, where X is a discrete random variable taking the values in χ. The *entropy* of X is

$$H(X) = E(-\lg P_X),$$

where $E(\cdot)$ is the expected value (weighted average) function; that is,

$$H(X) = -\sum_{x\in\chi} P_X(x)\lg P_X(x). \tag{1}$$

Intuitively, the entropy of X is a measure of the number of bits of uncertainty in X. For example, suppose χ is the set of all n-bit strings, and $P_X(x) = 1/2^n$ for any $x \in \chi$; that is, every n-bit string is equally likely to be pulled from P_X. This would represent a distribution of maximum uncertainty, and it is straightforward to show that Eqn. (1) evaluates to n in this case. In fact, $H(X) = \lg|\chi|$ is an upper bound for H, where $|\chi|$ denotes the cardinality of χ.

The *minimum entropy* of a distribution P_X is defined as

$$H_\infty(X) = \min_{x\in\chi}\{-\lg P_X(x)\} \tag{2}$$

This can be understood as a measure of uncertainty for the "most probable" element in χ according to P_X. For example, if there is some element x_0 with $P_X(x_0) = 1$, then $H_\infty(X) = 0$ (there is no uncertainty in X). Suppose the most probable element x_0 has probability $P_X(x_0) = 1/2$. Intuitively, the uncertainty is unity; that is, we can guess that the next value of X will be x_0 to within a single coin flip. Indeed, evaluating Eqn. (2) for such a distribution shows that $H_\infty(X) = 1$.

Give two probability mass functions P_X and P_Y, both with support χ, the *relative entropy*, also called the Kullback-Leibler divergence, from P_X to P_Y is defined to be

$$D(P_X\|P_Y) = \sum_{x\in\chi} P_X(x)\lg\frac{P_X(x)}{P_Y(x)}. \tag{3}$$

Intuitively, the relative entropy is a measure of the difference between P_X and P_Y, although it is important to note that it is not symmetric; that is, $D(P_X\|P_Y) \neq D(P_Y\|P_X)$. However, $D(P_X\|P_Y) = 0$ if and only if $P_X = P_Y$, and increases indefinitely as P_Y diverges from P_X.

A *channel* as defined by TODO: cite is a distribution on timestamped bit sequences; i.e., a channel $\mathcal{C}$ is a distribution with support $\{(\{0,1\},t_1),(\{0,1\},t_2),\ldots\}$, where each $t_i \leq t_{i+1}$. The intent is to model communication, where not just the content but also the timing of the communication may be relevant.

Security Definitions

Hopper paper:

Hopper, et. al TODO: cite defines the security of a stegosystem in terms of a game. Give the warden W access to $M(h)$, which returns draws from $\mathcal{C}_h^b$, and an oracle $\mathcal{O}$. The oracle $\mathcal{O}$ is either SE_k or a function $O(\cdot, \cdot)$, where $O(m, h)$ simply returns draws from $\mathcal{C}_h^{|\mathsf{SE}_k(m,h)|}$. The warden also has access to randomness r. The warden's advantage against the steganographic secrecy under chosen hiddentext attack for channel $\mathcal{C}$ of stegosystem S is defined by Hopper et. al to be

$$\mathbf{Adv}_{\mathsf{S},\mathcal{C}}^{\text{ss-cha-}\mathcal{C}}(W) = \left| \Pr_{k,r,M,\mathsf{SE}}\left[W_r^{M,\mathsf{SE}_k(\cdot,\cdot)} \text{accepts}\right] - \Pr_{r,M,O}\left[W_r^{M,O(\cdot,\cdot)} \text{accepts}\right] \right|.$$

A stegosystem S is $(t, q, \ell, \varepsilon)$*-steganographically secret under chosen hiddenttext attack* for channel $\mathcal{C}$ (SS-CHA-$\mathcal{C}$) if, for any warden W making at most q queries totaling at most ℓ bits of hiddentext, and running in time at most t,

$$\mathbf{Adv}_{\mathsf{S},\mathcal{C}}^{\text{ss-cha-}\mathcal{C}}(W) \leq \varepsilon;$$

that is, the stegosystem S is insecure if an efficient warden can (with high probability) distinguish between the output of $\mathsf{SE}_k(m, h)$ and draws from $\mathcal{C}_h^{|\mathsf{SE}_k(m,h)|}$, even when given access to $\mathcal{C}_h^b$ through M.

A stegosystem S is $(t, q, \ell, \varepsilon)$*-universally steganographically secret under chosen hiddenttext attack* for channel $\mathcal{C}$ (USS-CHA-$\mathcal{C}$) if it is $(t, q, \ell, \varepsilon)$-SS-CHA-$\mathcal{C}$ for any channel $\mathcal{C}$ that satisfies $H_\infty(\mathcal{C}_h^b) > 1 \forall h$ drawn from $\mathcal{C}$.

Cachin paper: