Cryptanalytic Time-Memory Tradeoff

Many problems in computer science have time-memory tradeoffs, in which the computation time can be reduced for the price of using larger memory (or the amount of memory can be decreased for the price of slowing the attack).

Can we find an attack in which the computation time can be reduced for the price of using larger memory (or the amount of memory can be decreased for the price of slowing the attack)?

Yes, we can. This is called a cryptanalytic time-memory tradeoff.

The attack requires large resources, even against DES.

Known Attacks

Denote the key size of a block cipher by $k$.

Then, there are two universal very simple attacks against any cipher:

1. Exhaustive search: $T = 2^k$ time, $S = 1$ memory

2. Table attack: $T = 1$ time, $S = 2^k$ memory (but $2^k$ precomputation, assumes knowing the plaintext in advance)

Both attacks require large resources, even against DES.

The Case of a Block Cipher

The attack assumes that the plaintext is known in advance, since it is required for the precomputation. Thus, this is basically a chosen plaintext attack, but it can also be applied as a known plaintext or even a ciphertext only attacks if there is some sufficiently frequent plaintext block (such as "Password").

If $f$ is a permutation (i.e., 1-1 and onto), then each value $x$ has a unique predecessor $f^{-1}(x)$, and the values are divided into cycles.

Let $t = \sqrt[k]{|T|}$.

In each cycle, select a starting point $x_0$, and let $p$ be the period of the cycle.

Assume without loss of generality that $p | t$. Compute:

$$x_1 = f(x_0), x_2 = f(x_1), \ldots, x_p = f(x_{p-1}).$$

Keep all the $x_0$'s of all the cycles in a table, where for each $x_0$ it is easy to access the respective $x_{p-1}$. The generation of this table costs $O(n)$.

Then, given $y'$, compute $f$ iteratively on $y'$ till the result appears in the table as some $x_{p-1}$. Jump to $x_{p-2}$ and iteratively compute $f$ again, till we get $y'$. The previous value is $f^{-1}(y')$. The complexity is $O(k\sqrt[k]{n})$.

The General Case: $f$ is Not a Permutation

In this case, there are no cycles, but a forest of many large connected subgraphs, with a small single cycle at the center of each.

Therefore, the technique used for permutations cannot be directly applied.
The Idea

Choose \( m \) starting points (keys) \( S_1, \ldots, S_m \), and apply \( f \) iteratively \( t \) times (we will choose \( m \) and \( t \) later):

\[
X_1 = Y_1 = f(Y_1) = f(f(Y_1)) = \ldots = f^{t-1}(Y_1) = E_1
\]

where for any \( i \) and \( j \), \( X_{i,j} = f(X_{i,j-1}) \), and

\[ E_i = f^i(S_i). \]

To save space, we will not keep the intermediate values of \( X_{i,j} \). We will keep only \((S_i, E_i)\) sorted by \( E_i \).

The Attack (cont.)

If \( Y_1 \) is not one of the \( E_i \)'s, or if a false alarm was found:

We compute \( Y_2 = f(Y_1) \) and check if it is one of the \( E_i \)'s.

If it is, \( X_{r+1} \) is a candidate for \( K \).

And so on, for \( Y_2, Y_3, \ldots \).

Analysis

If all the \( X_{i,j} \)'s are different, and if \( K \) is chosen at random, then the probability of success is:

\[ P(S) = \frac{m}{N} \]

where \( N = 2^q \).

For a particular probability, we can choose various values for \( m \) and \( t \) as long as their product remains unchanged.

This allows us to apply a tradeoff between time and memory: \( t \) time, \( m \) memory.

Can we choose \( m = t = 2^{\sqrt{2}} \)? No.

The Attack (cont.)

In general, we apply:

\[ Y_j = R(C_j) = f(K_j) \]

for \( j = 1 \) to \( t \) do:

If \( Y_j \notin \{E_i\} \) (say \( E_i = E_k \))

compute \( K' = X_{r-j+1} = f^{-i}(S_i) \)

Check if \( K' \) is a false alarm

If \( K' \) is not a false alarm

Output \( K = K' \)

\[ Y_{j+1} = f(Y_j) \]

Behavior of Random Functions

Denote the key space by \( K = \{0, \ldots, 2^q - 1\} \).

\( f \) is a function \( f : K \rightarrow K \).

Are All Values in \( K \) in the Range of \( f \)?

No. If \( f \) is random, then we can view the outputs as of throwing \( N \) balls into \( N \) holes.

About

\[
(1 - \frac{1}{N})^N \approx \frac{1}{e}
\]

of the holes remain empty, i.e., \( 1/e \) of the values in \( K \) are not in the range of \( f \).

Thus, unless these values are starting points (and most of them are not), we cannot find these keys by the table built from \( f \).

Also, many values have more than one preimage.
**How Long can we Walk Along $f$?**

Start from some $K_0$. Define $K_1 = f(K_0)$, $K_2 = f(K_1)$, etc.

How many times can we apply $f$ before we get back to somewhere we where before?

---

**If $f$ is a Permutation**

If $f$ would be a permutation, then $f$ would be invertible, and thus each value would have exactly one preimage.

Going along the path from $K_0$, we have probability $1/N$ to collide with $K_0$ in each step.

We cannot collide with any other value in the path.

Thus, the expected period of the cycle of $K_0$ is about $O(N)$ (more accurately, $N/2$).

In this case, we can deterministically choose the starting points, by walking along the cycles, and choose a starting point every $\sqrt{N}$ applications of $f$. In this case the complexity is really $t = m = \sqrt{N}$, and every key can be found by the attack using one table.

---

**Discussion**

If $f$ would define one huge cycle, we could divide the cycle into $T$ distinct paths of length $S = N/T$.

The best tradeoff would be $T = S = \sqrt{N}$.

On the other hand, if $f$ defined $N$ cycles of length one, $f$ would be the identity function, but our attack would not work at all.

In our case, after about $\sqrt{N}$ applications of $f$ we get a collision, i.e., a cycle, and continuing the path do not add any new value to the path. Thus, we should limit $t$ to be $t < \sqrt{N}$.

Moreover, there is no point to increase $mt^2$ over $mt^2 = N$, since the gain in the probability of success is very small, due to overlaps.

If we choose $m = t = N^{1/3}$ we get $P(S) = \frac{mt}{N} = N^{-1/3}$. (Hellman claims that this choice is optimal.)

---

**Rate of False Alarms (cont.)**

Proof of the Theorem

Let $F_j$ be the random variable of having a false alarm when processing $Y_j = E_i$.

Then,

$$E(F) = \sum_{i=1}^{N} \frac{1}{t} \cdot P(F_{ij}).$$

Look at $F_{ij}$. False alarm in $F_{ij}$ can be achieved in $t - j + 1$ different (and independent) cases:

1. if $f(K)$ is already in the path from $S_j$ to $E_i$.
2. if $f(K)$ is not in the path from $S_j$ to $E_i$ but $f(K) = f(K)$ is.
3. if neither $f(K)$ nor $f(K)$ are not in the path from $S_j$ to $E_i$ but $f(K)$ is.

4. etc.
Rate of False Alarms (cont.)

Each of them has probability $1/N$ to occur, thus

$$P(F_{i,j}) = (t - j + 1)/N.$$ 

In total

$$E(F) \leq \sum_{i=1}^{t} \sum_{j=1}^{t} (t - j + 1)/N = \sum_{i=1}^{t} \sum_{j=1}^{t} j/N = \frac{\text{int}(t+1)}{2N} \approx 1/2.$$ 

QED