5 Random Keys

All cryptanalytic methods collapse when the key is a random letter sequence, chosen in an independent way for each plaintext, and never repeated. In particular all the letters in the ciphertexts occur with the same probability. Or in other words, the distribution of the ciphertext letters is completely flat.

This encryption method is called **One-Time Pad** (OTP). Usually Gilbert VERNAM (1890–1960) is considered as the inventor in the World War II year 1917. But the idea of a *random* key is due to MAUBORGNE who improved VERNAM's periodic XOR cipher in this way. The German cryptologists KUNZE, SCHAUFFLER, and LANGLOTZ in 1921—presumably independently from MAUBORGNE—proposed the "individuellen Schlüssel" ("individual key") for running-text encryption of texts over the alphabet $\{\mathtt{A}, \ldots, \mathtt{Z}\}$.

In other words: The idea "was in the air". In 2011 Steve Bellovin discovered a much earlier proposal of the method by one Frank MILLER in 1882 who however was completely unknown as a crypologist and didn't have any influence on the history of cryptography.

Steven M. Bellovin. *Frank Miller: Inventor of the One-Time Pad.* Cryptologia 35 (2011), 203–222.

Uniformly Distributed Random Variables in Groups

This subsection contains evidence for the security of using random keys. The general idea is:

"Something + Random = Random" or "Chaos Beats Order"
(the Cildren's Room Theorem)

We use the language of Measure Theory.

Theorem 1 *Let G be a group with a finite, translation invariant measure μ and Ω, a probability space. Let $X, Y : \Omega \longrightarrow G$ be random variables, X uniformly distributed, and X, Y independent. Let $Z = X * Y$ (where $*$ is the group law of composition). Then:*

(i) *Z is uniformly distributed.*

(ii) *Y and Z are independent.*

Comment The independency of X and Y means that

$$P(X^{-1}A \cap Y^{-1}B) = P(X^{-1}A) \cdot P(Y^{-1}B) \quad \text{for all measurable } A, B \subseteq G.$$

The uniform distribution of X means that

$$P(X^{-1}A) = \frac{\mu(A)}{\mu(G)} \quad \text{for all measurable } A \subseteq G.$$

In particular the measure P_X on G defined by $P_X(A) = P(X^{-1}A)$ is translation invariant, if μ is so.

Remark Z is a random variable because $Z = m^{-1} \circ (X, Y)$ with $m = *$, the group law of composition. This is measurable because its g-sections,

$$(m^{-1}A)_g = \{h \in G \mid gh \in A\}$$

are all measurable, and the function

$$g \mapsto \mu(m^{-1}A)_g = \mu(g^{-1}A) = \mu(A)$$

is also measurable. A weak form of FUBINI's theorem gives that $m^{-1}A \subseteq G \times G$ is measurable, and

$$(\mu \otimes \mu)(m^{-1}A) = \int_G (m^{-1}A)_g \, dg = \mu(A) \int_G dg = \mu(A)\mu(G).$$

Counterexamples We analyze whether the conditions of the theorem can be weakened.

1. What if we don't assume X is uniformly distributed? As an example take $X = \mathbf{1}$ (unity element of group) constant and Y arbitrary; then X and Y are independent, but $Z = Y$ in general is not uniformly distributed nor independent from Y.
2. What if we don't assume X and Y are independent? As an example take $Y = X^{-1}$ (the group inverse); the product $Z = \mathbf{1}$ in general is not uniformly distributed. Choosing $Y = X$ we get $Z = X^2$ that in general is not uniformly distributed nor independent from Y. (More concrete example: $\Omega = G = \mathbb{Z}/4\mathbb{Z}, X =$ identity map, $Z =$ squaring map.)

General proof of the Theorem

(For an elementary proof of a practically relevant special case see below.)
Consider the product map

$$(X, Y) \colon \Omega \longrightarrow G \times G$$

and the extended composition

$$\sigma \colon G \times G \longrightarrow G \times G, \quad (g, h) \mapsto (g * h, h).$$

For $A, B \subseteq G$ we have (by definition of the product probability)

$$(P_X \otimes P_Y)(A \times B) = P_X(A) \cdot P_Y(B) = P(X^{-1}A) \cdot P(Y^{-1}B);$$

because X and Y are independent we may continue this equation:

$$\begin{aligned} &= P(X^{-1}A \cap Y^{-1}B) = P\{\omega \mid X\omega \in A, Y\omega \in B\} \\ &= P((X,Y)^{-1}(A \times B)) = P_{(X,Y)}(A \times B). \end{aligned}$$

Therefore $P_{(X,Y)} = P_X \otimes P_Y$, and for $S \subseteq G \times G$ we apply FUBINI's theorem:

$$P_{(X,Y)}(S) = \int_{h \in G} P_X(S_h) \cdot P_Y(dh).$$

Especially for $S = \sigma^{-1}(A \times B)$ we get

$$\begin{aligned} S_h &= \{g \in G \mid (g * h, h) \in A \times B\} = \begin{cases} A * h^{-1}, & \text{if } h \in B, \\ \emptyset & \text{else,} \end{cases} \\ P_X(S_h) &= \begin{cases} P_X(A * h^{-1}) = \frac{\mu(A)}{\mu(G)}, & \text{if } h \in B, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

Therefore

$$\begin{aligned} P(Z^{-1}A \cap Y^{-1}B) &= P\{\omega \in \Omega \mid X(\omega) * Y(\omega) \in A, Y(\omega) \in B\} \\ &= P((X,Y)^{-1}S) = P_{(X,Y)}(S) \\ &= \int_{h \in B} \frac{\mu(A)}{\mu(G)} \cdot P_Y(dh) = \frac{\mu(A)}{\mu(G)} \cdot P(Y^{-1}B). \end{aligned}$$

Setting $B = G$ we conclude $P(Z^{-1}A) = \frac{\mu(A)}{\mu(G)}$, which gives (i), and from this we immediately conclude

$$P(Z^{-1}A \cap Y^{-1}B) = P(Z^{-1}A) \cdot P(Y^{-1}B)$$

which proves also (ii). $\diamond$

Proof for countable groups

In the above proof we used general measure theory, but the idea was fairly simple. Therefore we repeat the proof for the countable case, where integrals become sums and the argumentation is elementary. For the cryptographic application the measure spaces are even finite, so this elementary proof is completely adequate.

Lemma 1 *Let G, Ω, X, Y, and Z be as in the theorem. Then*

$$Z^{-1}(A) \cap Y^{-1}(B) = \bigcup_{h \in B} [X^{-1}(A * h^{-1}) \cap Y^{-1}h]$$

for all measurable $A, B \subseteq G$.

The proof follows from the equations

$$\begin{aligned}
Z^{-1}A &= (X,Y)^{-1}\{(g,h) \in G \times G \mid g * h \in A\} \\
&= (X,Y)^{-1}\left[\bigcup_{h \in G} A * h^{-1} \times \{h\}\right] \\
&= \bigcup_{h \in G} (X,Y)^{-1}(A * h^{-1} \times \{h\}) \\
&= \bigcup_{h \in G} [X^{-1}(A * h^{-1}) \cap Y^{-1}h], \\
Z^{-1}A \cap Y^{-1}B &= \bigcup_{h \in G} [X^{-1}(A * h^{-1}) \cap Y^{-1}h \cap Y^{-1}B] \\
&= \bigcup_{h \in B} [X^{-1}(A * h^{-1}) \cap Y^{-1}h].
\end{aligned}$$

Now let G be countable. Then

$$\begin{aligned}
P(Z^{-1}A \cap Y^{-1}B) &= \sum_{h \in B} P[X^{-1}(A * h^{-1}) \cap Y^{-1}h] \\
&= \sum_{h \in B} P[X^{-1}(A * h^{-1})] \cdot P[Y^{-1}h] \quad \text{(because } X, Y \text{ are independent)} \\
&= \sum_{h \in B} \frac{\mu(A * h^{-1})}{\mu(G)} \cdot P[Y^{-1}h] \quad \text{(because } X \text{ is uniformly distributed)} \\
&= \frac{\mu(A)}{\mu(G)} \cdot \sum_{h \in B} P[Y^{-1}h] \\
&= \frac{\mu(A)}{\mu(G)} \cdot P\left[\bigcup_{h \in B} Y^{-1}h\right] \\
&= \frac{\mu(A)}{\mu(G)} \cdot P(Y^{-1}B).
\end{aligned}$$

Setting $B = G$ we get $P(Z^{-1}A) = \frac{\mu(A)}{\mu(G)}$, which gives (i), and immediately conclude

$$P(Z^{-1}A \cap Y^{-1}B) = P(Z^{-1}A) \cdot P(Y^{-1}B),$$

which proves (ii). $\diamond$

Discussion

The theorem says that a One-Time Pad encryption results in a ciphertext that "has nothing to do" with the plaintext, in particular doesn't offer any lever for the cryptanalyst.

Why then isn't the One-Time Pad the universally accepted standard method of encryption?

- Agreeing upon a key is a major problem—if we can securely transmit a key of this length, why not immediately transmit the message over the same secure message channel? Or if the key is agreed upon some time in advance—how to remember it?

- The method is suited at best for a two-party communication. For a multiparty communication the complexity of key distribution becomes prohibitive.

- When the attacker has known plaintext she is not able to draw any conclusions about other parts of the text. But she can exchange the known plaintext with another text she likes more: *The integrity of the message is at risk.*