統計学の観点から見た暗号解読に関する定義について

神奈川工科大　竹田　裕一　　中央大・理工　藤井　光昭
中央大・理工　渡邉　則生　　中央大・理工　鎌倉　稔成

1. はじめに

コンピュータが発達し、様々な情報を送信する際にはセキュリティの観点から、送信者および受信者以外の第 3 者にはその情報を読むことができないように暗号文に変換を行う必要がある。セキュリティの分野では暗号として、主に代数的方法と乱数を用いる方法が考案されているが、本研究では乱数を用いる方法について議論する。

2. 暗号化と解読可能・解読不可能について

暗号については、暗号文を送信者および受信者以外の第 3 者が手に入れた場合、暗号文から元のメッセージを解読することができないことが重要である。本研究ではでは次のような表現を用いることとする。

メッセージはある文字により表現されているとし、各文字を 0 と 1 の組み合わせのパターン（各要素が 0 か 1 かである m 次元ベクトル）$\mathbf{\Xi}$ で表し、これに各要素が 0 か 1 の乱数ベクトル $\mathbf{Z}$ を要素毎に加えて送信信号 $\mathbf{X}$ を作る。これは、各要素を 2 進法での和を用いて $\mathbf{X} = \mathbf{\Xi} + \mathbf{Z}$ と表すことができる。メッセージ $\mathbf{\Xi}$ と乱数 $\mathbf{Z}$ を独立な確率変数と考えるため、送信信号 $\mathbf{X}$ も確率変数と考える。また、$\mathbf{\Xi}, \mathbf{Z}, \mathbf{X}$ の取る値をそれぞれ $\mathbf{a}, \mathbf{z}, \mathbf{x}$ とする。さらに前提条件として、各文字の出現確率 $P(\mathbf{\Xi} = \mathbf{a})$ が既知であるものとし、それを手掛かりに暗号の解読を行う場合を考える。このとき

$$P(\mathbf{X} = \mathbf{x}) = \sum_{\mathbf{a}} P(\mathbf{\Xi} = \mathbf{a}) P(\mathbf{Z} = \mathbf{x} - \mathbf{a})$$

が成立しており、解読を行う第 3 者は $P(\mathbf{X} = \mathbf{x})$ のみを知ることができ、$P(\mathbf{\Xi} = \mathbf{a})$ と $P(\mathbf{Z} = \mathbf{x} - \mathbf{a})$ は未知である。

セキュリティの分野では、乱数 $\mathbf{Z}$ に関する議論が行われ、乱数性の検定に強い関心が払われている。実際、米国国立標準技術研究所 (NIST) では 15 個の検定法が提案されている。このとき用いられている帰無仮説としては「0 と 1 がそれぞれ確率 1/2 で出現し、さらに各ビットの値は独立に出現する」が用いられているが、対立仮説に関しては示されていない。しかし、暗号化については暗号が解読されないことが重要であることから、帰無仮説として「$\mathbf{x}$ から $\mathbf{a}$ が解読されない」とすべきと考える。しかしながら、「解読されない」の数学的定義が現在までに明確に示されていないため、本報告では「解読されない 0,1 乱数列 $\mathbf{Z}$」の数学的定義を試みた研究結果の報告を行う。ここで「解読不可能」は次の場合を表現することとする。

> 「$P(\mathbf{X} = \mathbf{x})$ が得られたとき、$\forall \mathbf{a}$ について $P(\mathbf{\Xi} = \mathbf{a})$ として 0 も 1 も取り得て、
> さらに無限個の値を取る可能性もある」

例として、NIST が用いている帰無仮説を満たす場合は解読不可能になる。また、$\mathbf{Z}$ が周期 m を持つ場合は、解読不可能の条件を満たさない等が示せる。このことに関し、シミュレーション実験を行い、その結果も当日報告する。