暗号・情報保全史特論
History of Cryptograph and Signal Security Advanced Course

第1回: 概要と簡単な古典暗号の歴史

佐藤永欣

授業ページについて

- http://www.db.soft.iwate-pu.ac.jp/~nobu-s/lecture/crypt_signal_security_history/

講義の概要と目標

- 古典的な暗号に対する攻撃手法と、古典暗号の解読事例を通じて情報システム全体を攻撃から守る方法を考える
  - 現代暗号(AES、RSAなど)は攻撃しにくい
    - 実習のようなことをしたくても現実的には無理
  - 古典的な暗号は手作業・歯車計算機・リレー式計算機で扱えるレベル
    - 現代の電子計算機なら力技でも十分行ける
- 初期の電子計算機による古典暗号の解読事例では、人によるミスや漫然とした運用が糸口を与えた
  - 現在でも情報システムのセキュリティが破られる主な原因は、漫然とした運用や無理解が多い
- 人類の歴史は情報を隠したい者と知りたい者のせめぎあいの歴史でもあった(おおげさ?)

講義の概要と目標

- 古典的な暗号
  - ざっくり、電子計算機登場以前の暗号
  - 計算を必要とする暗号もあるが、手計算・そろばんでどうにかなるレベル
  - 実運用としては変換表を用意して表を引くのがおおい
- 古典暗号に対する攻撃手法と解読事例
  - ルネサンス時代～第二次世界大戦までの事例を扱う

- 現代暗号の最初は安全だったが時間の経過とともに安全でなくなる例
  - DES: 1980年代は鉄壁だったが1999年に22時間で解読→3DESを経てAESに移行
  - RSA: 2015年ごろから鍵長1024ビットだと安全性が不足といわれだした
  - SHA-1: 2017年ハッシュ値を衝突させる攻撃が成功し始める
- 計算機の能力向上・アルゴリズムの弱点が見つかる→安全性低下
  - 古典暗号でも同じことが2000年間起きていた

授業の計画

- 計画は無計画！！！
  - 第1回 概要と簡単な古典暗号の歴史
  - 第2回 古代～中世の暗号 (手作業による暗号化と復号化，シーザー暗号，単一換字暗号)
  - 第3・4回 近世～第2次大戦期の暗号発達史と暗号解析の技法（頻度分析，クリブ，ラテン方陣，ヴィジュネル暗号，プレイフェア暗号，多表式暗号のはじまり，コードブックの利用）
  - 第5・第6回 多表式暗号と攻撃手法（機械式暗号の時代，フリードマンンテスト）
  - 第7回 多表式暗号の解読技法と暗号の運用（海軍D暗号）
  - 第8回 乱数生成・ステガノグラフィ・情報システム保全に関する哲学
- 時間が余ればステガノグラフィの実験なども計画
  - 教室で火を使って怒られないかが問題

テキスト

- 基本的に資料を配布
- 参考書など
  - 暗号大全 原理とその世界 長田順行, 講談社学術文庫, ISBN978-4-06-292439-9
  - 基礎暗号学I//II, 加藤正隆, サイエンス社(絶版), ISBN4-7819-0562-5 / ISBN4-7819-0563-3
  - 暗号技術入門第3版 結城浩、SBクリエイティブ, ISBN978-4-7973-8222-8
  - 日本陸軍暗号の敗北, 伊藤秀美, 紫峰出版, ISBN978-4-907625-19-1
  - 暗号教範 陸軍参謀本部編, 伊藤秀美・保坂廣志解説, 紫峰出版(翻刻版), ISBN978-4-907625-09-2
  - 新教程日本陸軍暗号 米陸軍数新保安部編, 伊藤秀美・保坂廣志訳, 紫峰出版(翻刻版), ISBN978-4-907625-07-8
  - 日本軍の暗号作戦 上/下, 保坂廣志, 紫峰出版, ISBN978-4-907625-08-5 / ISBN978-4-907625-12-2

評価

- 課題レポートをいくつか出題→そのうち一つをえらんで提出
- 採点結果による
  - いい成績は取ってほしい

現代的な情報セキュリティから見た立ち位置

- 現代の暗号関連アルゴリズム: AES, RSA, SHA, MD5, etc.
  - 基本的に「計算量的に安全」
  - 鍵を見つける/ハッシュ値を衝突させるにはbrute force attackが基本
  - 暗号関連アルゴリズムに対する攻撃は、攻撃者から見ると実用的ではない
    - とはいえ、アルゴリズムの弱点が発見されれば利用される
- 「計算量的に安全」
  - 力技で殴れば理論的には可能だが解読完了は現実的には難しい、の意味
  - 時間がかかりすぎる and/or力技を実行するのが大変
- 「現実的には難しい」の目安
  - 解読までの計算時間の期待値として10年以上?
    - 用途にもよる。数か月でもまあまあ安全なこともある
  - 個人・会社レベルでは力技で殴る手段を用意するのが無理→すごくお金がかかる

現代的な情報セキュリティから見た立ち位置

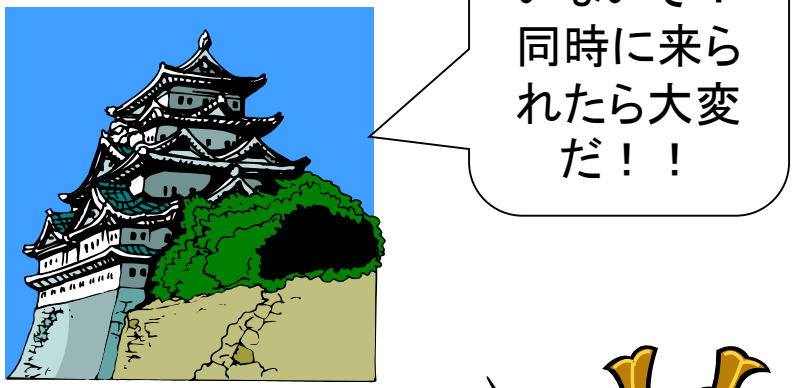
- よくあるセキュリティ事案(=色々盗んだりするのに実用的な攻撃)
  - 人によるミスを誘発する/裏切りetc.を狙うのがほとんど
    - 有害なリンクを踏ませる
    - FWの内部で外部からコントロール可能なプログラムをどうにかして送り込む
- 人によるミスを防止するためにいろいろな規則や制限を設ける
  - 規則: 一人で操作させない、通信機器持ち込み禁止、etc.
  - 制限: USBポートをふさぐ、httpとhttps以外通さないFW、etc.
- でもね、人間だもの。
  - 日常的にやることの範囲に面倒な制限があれば回避したくなる
- 人による「やらかし」を狙うのが主流
  →昔の「やらかし」に学ぶのは意味がありそう

現代的な情報セキュリティから見た立ち位置

- 昔の「やらかし」: 第2次大戦中の敗戦側の「やらかし」はわかっている
- 暗号化・復号化と通信は手作業(機械を使っていても手作業が入る)
  - 表を引いて暗号化→モールス信号で電信→表を引いて復号化
  - キーボードを打つ→光ったランプの文字をメモ→電信→キーボードを打って復号化
- (当時としては)強力な暗号を使っていても…
  - 通信手の手抜き
    - 暗号化するときの表の開始位置が毎回同じ
    - 暗号鍵が単純・数種類の使いまわし
  - 運用ミス
    - 通信文の文頭がいつも同じ
    - 同じ通信文を何種類かの暗号で送ってしまう
  - 暗号書が盗まれた→お役所的事情で盗難事件はなかったことになる

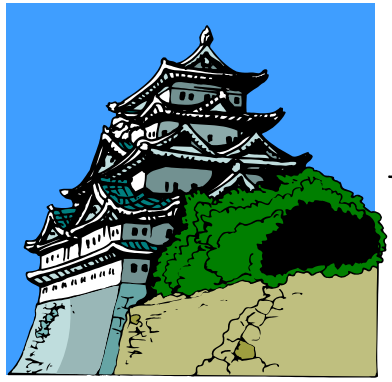
現代的な情報セキュリティから見てみる

- 秘匿
  - 情報を第3者に知られないこと、隠すこと
    - 知っている人が少ない情報ほど価値がある
  - 秘匿できなかったときの問題点
    - 個人情報流出
    - 軍事作戦上の秘密情報
      - 秘密がばれたら味方が死ぬ



現代的な情報セキュリティから見てみる

- 暗号化すると…
  - 通信が行われているのはわかるが、受け手以外には内容はわからない（情報の秘匿）
  - 暗号化できるのは、秘密の鍵を知っている人だけのはず（送信者の認証）
  - 暗号化の代わりに、使用者が少ない言語に翻訳することもあった



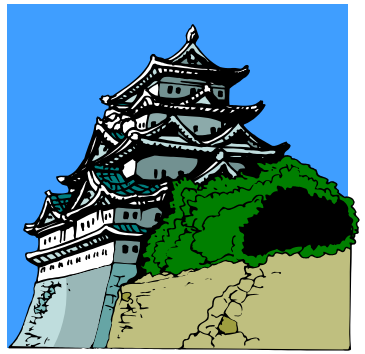
…。なんだかわからないぞ？？？

○×▲◆□◎△▽
×■●▲×△×□
◇◎○▲

◇◎×××◇▲○△▼
●□▽

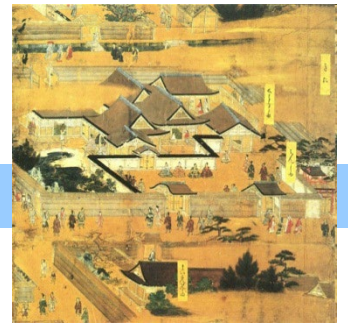
現代的な情報セキュリティから見てみる

- 暗号化以外の隠し方: ステガノグラフィ
  - 絵や文書の中に情報の受け手以外にわからないように情報を隠すこと
    - 秘密のインクで手紙をかく
    - 和歌の中に隠す



絵のやり取りですか…。

この絵は京の都の絵でござってな。



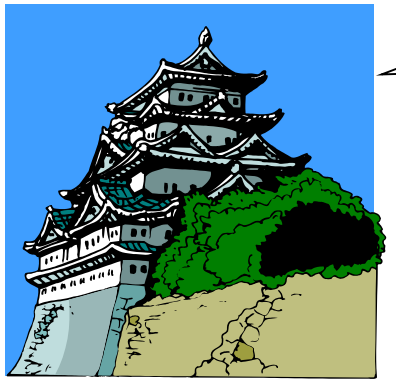
明日の正午に10000人！

明日の正午に10000人であの城を攻撃するから、10000人で攻めて！！！

ほう。これは中々良いものでござるな。

現代的な情報セキュリティから見てみる

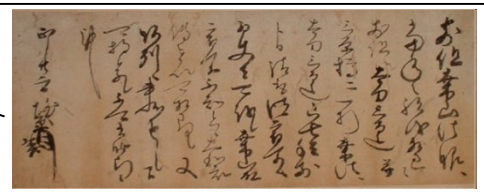
- 改竄不可能性(古典暗号の世界には概念がない)
  - メッセージ認証コード、タイムスタンプ

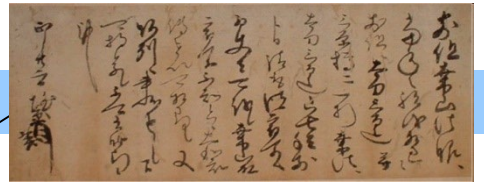


書状を書き換えよう。
これで別個に返り討ち！！

この城を明日
正午に10000人で
攻撃するでござる

明後日の正午に10000人であの城を攻撃するから、10000人で攻めて！！！





明日の正午に10000人であの城を攻撃するから、10000人で攻めて！！！

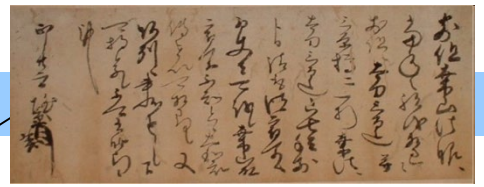
明後日か。
10000人で攻めるでござる。

現代的な情報セキュリティから見てみる

- 送信者保証（中間者攻撃の防止）(古典暗号の世界にはない)
  - 送信者が誰か確認できることを保証する
  - チャレンジアンドレスポンス

書状をすりかえよう。

この城を明日
正午に10000人で
攻撃するでござる



この書状は本物でござるか？？？
いつもと紙が違うでござる？？？

明日の正午に10000人であの城を攻撃するから、10000人で攻めて！！！

伊達政宗の花押の
鳥の目に、本物は針穴が
開いているというのは
単なる伝説らしい
(針穴が開いているのが
誰かに知られた時点で
使えなくなってしまう)

現代的な情報セキュリティから見てみる

- 送信者保証（成りすましの防止）(古典暗号の世界では個別に工夫)

情報の秘匿に対する要求

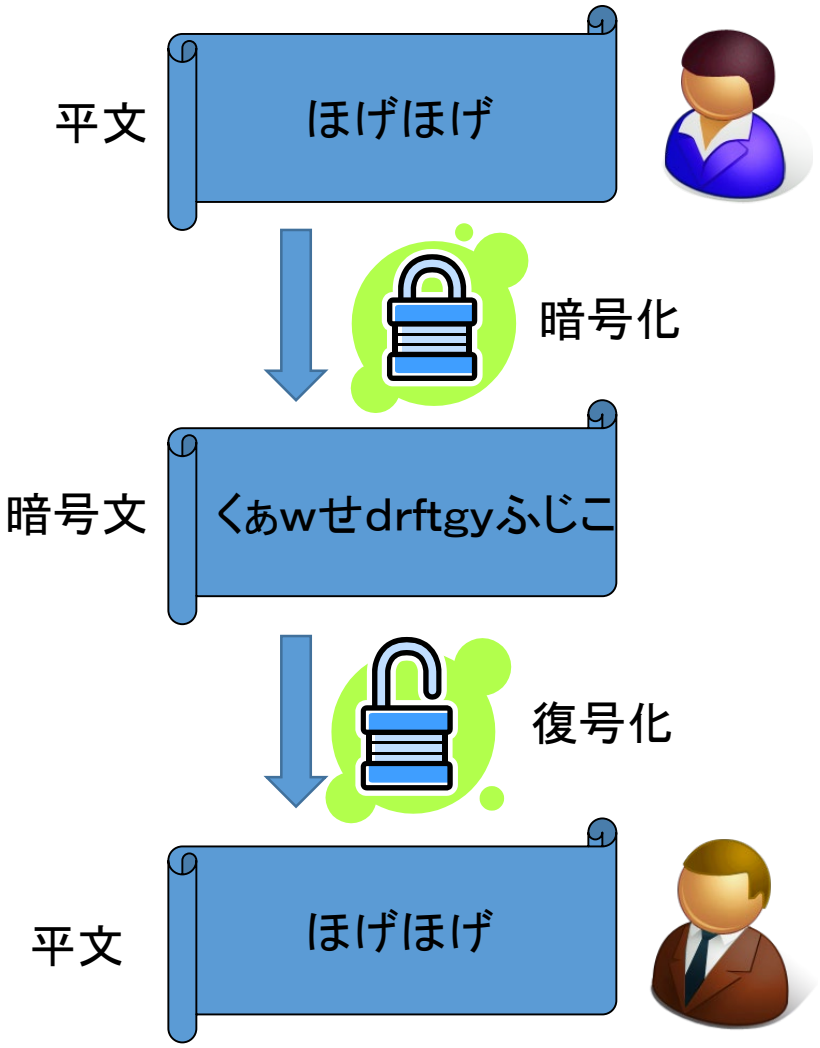
- 情報の存在そのものを秘匿
  - 通信そのものを秘匿
  - 大量の通信に紛れ込ませる
    - 通信量が急に増えると、意味が分からなくても「何かあった」と悟られてしまう
    - 普段から通信量が多ければ気づかれにくい
- 情報は盗まれても意味が分からないようにする
  - 暗号化
    - 手紙を盗まれる
    - 使者が捕らえられる
    - 使者が裏切る

用語

- Cipher (サイファ)
  - 文字単位で暗号に変換する
- Code (コード)
  - 単語単位で暗号に変換する: 隠語・符丁・略語
  - 寿司屋の符丁: ガリ、ムラサキ、シャリ
- 日本語ではどちらも暗号(英語ではcryptが相当する?)
- Cipherとcodeは組み合わせて使われた
  1. Code bookに頻繁に使う単語のCodeの順引き・逆引きを用意
  2. 通信文に出てくる単語をcodeで置き換える
  3. 出来上がった暗号文をさらにcipherで暗号化
  - Code bookを更新することで頻繁に使う単語を解読者にわかりにくくした

用語

- 平文(plain text)
  - 暗号化されていない、普通のメッセージ
- 暗号文(cipher text)
  - 暗号化されたメッセージ
- 暗号化(encrypt, encipher)
  - 平文を暗号化すること
- 復号化(decrypt, decipher)
  - 暗号文を平文に戻すこと
- 解読(crypt analysis)
  - 送信者と受信者以外が暗号を平文に戻すこと



お約束

- 古典暗号の世界には独特のお約束がある
  - 平文を小文字
  - 暗号文を大文字
  - 分かち書き(空白)
    - しない
    - 空白の代わりになにか文字を使う
    - する
  - 転置: 平文の文章全体をいくつかに分割→順番を入れ替える
    - 特に転置しない
    - 長い電文だけ転置する

情報の秘匿のレベル: 手紙の類の場合

- 情報の内容が知られている
  - 公開書簡
- 情報の存在が知られている(中身は不明)
  - 外国語の手紙: 外国語を知っている人には読める
  - 暗号化された手紙: 復号化・解読の方法を知っている人には読める
- 情報の存在も知られていない
  - 手紙をある規則で読むと別の内容になる
  - 透明なインクをつかった手紙
  - 手紙そのものを隠匿する

意思の伝達手段と秘匿

- 音声言語の秘匿
  - 必要性が大きくなったのは有線電話・無線電話の発明以降
  - 歴史的には外国語・使用者の少ない言語をしゃべる、符丁を使う
  - お経(漢文の日本式音読)、祝詞(古い日本語)、ミサ(ラテン語、教会スラブ語)
    - 秘匿は意図していないが日常の言語とはかけ離れすぎていて理解できる人が少ない
- 文字言語の秘匿
  - 換字: 文字を別の文字に置き換え
  - 転置: 文字の順番を入れ替え
  - 分置: 別の文字を挿入
  - 隠語: 単語レベルで別の意味を付与して置き換え
  - 物理的な秘匿

意思の伝達手段と秘匿

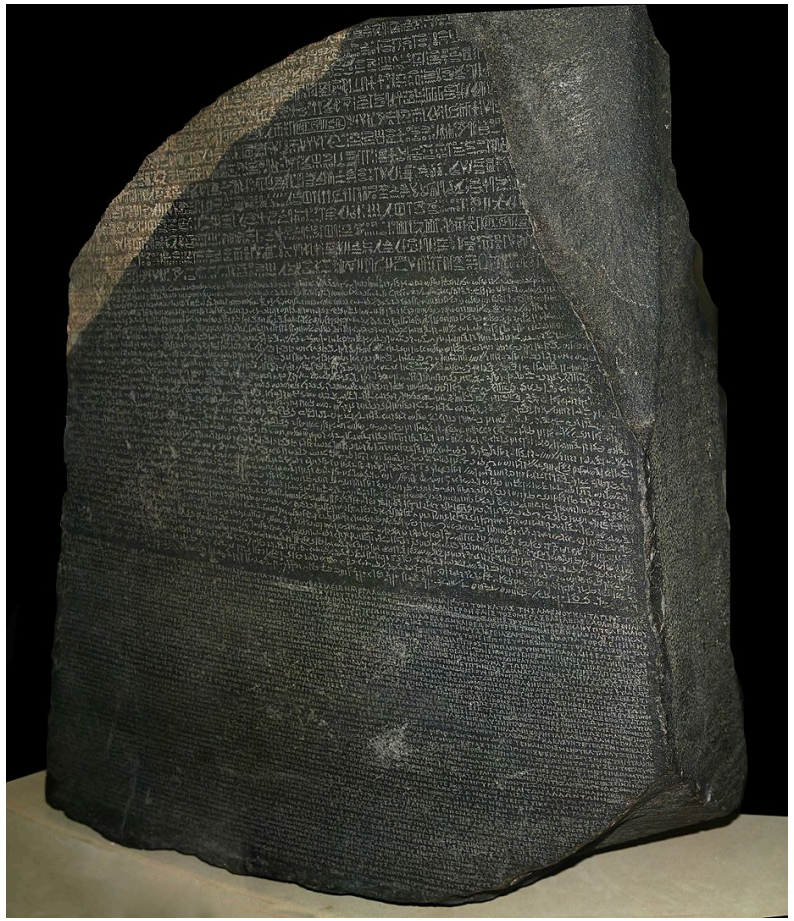
- 言語・文字の発達史の側面から見てみる
  - 最初は単なる絵、そこから絵文字・象形文字・縄の結び目のように抽象化が始まる
    - 類似の字形で意味の派生が起きたり(漢字でもその辺は同様、かつシステマチック)



    - ヒエログリフや楔形文字の解読には19世紀当時の暗号解読の手法も応用された
      - 対照平文との対応関係、類似する既知の言語との対照、可能な単語のあてはめ、他
  - 表意文字: 漢字、ヒエログリフ、マヤ文字、公共の場のアイコン
  - 表音文字: 表意文字が表音的に使われ出して文字の形が変わる
    - 意味が無関係になると書きやすさが優先される模様。印刷術発明で字形が固定された
    - ܐܠܦܒܝܬ, ܐܣܛܪܢܓܠܐ, ܐܒܓܕ (シリア文字の各書体の例。時代と東西で異なる。左から古、西、東)
- 表意文字の段階まではある種のコードと考えられる

おまけ: ヒエログリフの解読について

- ヒエログリフ: 古代エジプトの象形文字
  - 
    - 左がUnicode表現; Tut ankh Amun ツタンカーメン
  - ４世紀ぐらいまでには読めなくなっていた
  - 表意文字として始まり表音文字の機能を持つようになっていた
  - 表意文字と表音文字の両方をまぜて使われた
- ロゼッタストーン: 1799年発見
  - ヒエログリフ、デモティック(民衆文字)、古代ギリシャ語で同じ文がかかれていた
  - 一番上のヒエログリフは半分ほど失われていた
  - デモティックと古代ギリシャ語も一部失われていた
    - 古代ギリシャ語は失われた部分の文章の推定ができた



写真© Hans Hillewaert, CC BY-SA 4.0
大英博物館蔵

おまけ: ヒエログリフの解読について

- 最初の仮説: 3つの言語で同じ内容ではないか?
  - ヒエログリフが表意文字だとしても、外国語由来の固有名詞は表音的に書かれているのではないか?　(例) 巧克力: チョコレート　咖啡: コーヒー
- 古代ギリシャ語部分: 欧州の学問の伝統でほぼ理解されている
  - 固有名詞を取り出す; ギリシャ人の名前はエジプトにとっては外来語
    - プトレマイオス(PTOLMES)、クレオパトラ(KLEOPATR)が最終的に取り出された
  - あまり忠実な訳ではなかったらしい(ギリシャ語と古代エジプト語は言語の系統が全く異なる)
- トマス・ヤングによる2番目の仮説: デモティックはヒエログリフの発展形ではないか
  - デモティックも読み方が失われているが、古代エジプト語の最終形に近いはずのコプト語から類推できる?
    - デモティック自体はパピルスの巻物などで存在が知られていた

おまけ: ヒエログリフの解読について

- 二つの仮説に基づいて行われた作業(シャンポリオンによる)
  - デモティックから古代ギリシャ語の固有名詞に該当しそうな部分をとりだす
  - ヒエログリフとデモティックの対応関係を推定する
    - デモティックはヒエログリフと同じ言語、またはヒエログリフの娘言語にあたる言語と仮定
  - カルトゥーシュ(ヒエログリフを囲んでいる枠)の中が王の名前の固有名詞ではないかという次の仮説
  - 最終的にいくつかのギリシャ語アルファベットに対応する音価を持つヒエログリフが特定された
- そこから王の名前のヒエログリフが特定された(各地にあった碑文や壁画の文字が利用された)
  - ラムセス(Ramses )、トトメス(Tuthmosis )
- その後、カルトゥーシュ以外の部分の表音文字が特定された
  - 同音の文字が何種類かある、母音はあまり表記されない、などが判明

ステガノグラフィ: 存在そのものの秘匿

- 文字を見えなくする例
  - 使者の頭をそり上げる→秘密の通信を入れ墨→髪の毛が伸びたら出発
  - カラスの羽に墨で書く→絹の布を密着させて蒸して写し取って読む
  - 透明なインクを使用→熱・水・薬品・紫外線・赤外線により現像
    - 熱で現像(あぶり出し): 柑橘類の汁、重炭酸ナトリウム
      - 柑橘類の汁: 汁は酸性→火であぶると汁がついているところが先にこげる
      - 塩化コバルト水溶液: 熱であぶり出して乾燥させると青くなる。放置しておくと水分を吸って薄いピンクに戻る。水溶液の濃度によるが薄ピンクになると視認できない
      - (フリクションボール: 熱で消える、冷凍庫に入れて冷やすと元に戻る)
    - 水で現像: ミョウバンの粉で書く→水に浮かべる
    - 薬液・ガスで現像: ぶどう糖水溶液+硫酸アンモニウム水溶液、フェノールフタレイン+アンモニア水、硫化鉄・硫化銅+アンモニア
      - アミノピリン(大昔の解熱剤。毒性により現在は販売禁止)をジンやウォッカで溶かす方法があるらしいが現像方法が不明

ステガノグラフィ: 存在そのものの秘匿

- 文字の中に文字を隠す例
  - から衣 きつつなれにし つましあれば はるばるきぬる たびをしぞおもふ (伊勢物語)
    - かきつはた→かきつばた→燕子花
    - 和歌: 頭(語頭をとる) 沓(語尾を取る) 古今集・新古今集の時代の言葉遊び?
  - Having trouble about loudspeaker. Believe antenna connected improperly, but do whatever you can
    - Get ready to run
  - 2文字目、3文字目…・単語1個飛ばし2個飛ばしなど様々なパターンがある
- 別のもので覆ってしまう・物理的に隠す
  - 切手の下に極小文字で通信文を書く; 切手は水につけておけばはがせる
  - 手紙を細く切ってコヨリにして色々なところに編み込む

古典暗号史概観

- この授業で扱う古典暗号の定義: 現代暗号でないもの
- 現代暗号
  - 概ね1960年代以降に開発
  - 暗号化・復号化・通信に電子計算機を使用するのが前提
  - 計算量的に安全であることが暗号自体の安全性の根拠
  - 時間の経過とともに計算量的に安全ではなくなる
    - アルゴリズム上の弱点が発見される
    - 計算機の能力が向上する
- 現代暗号のよく知られているアルゴリズム: 一部はもうダメなのがあるね?
  - 共通鍵暗号: DES AES
  - 公開鍵暗号: RSA DSA ECDSA
  - 一方向ハッシュ関数: MD5 SHA-1 SHA-3 SHA256

古典暗号史概観

- 古典暗号
  - 古代から第2次世界大戦ごろまでの間に使用されていた
  - 暗号化・復号化に手作業 or 機械式スイッチなどを使用
  - 通信は手紙、狼煙、手旗、腕木、電信 (テレタイプライタも入れていいかも)
  - 現代から振り返ると、電子計算機がない世界であれば計算量的に安全だったようにも思われる(計算量的に安全という概念がなかった)
- 換字式・転置式
  - 別の文字に入れ替える・文字は同じで順番を入れ替える
- コードブック式
  - 単語ごとに別の単語に置き換える
- 物理的な秘匿: あぶりだし

古典暗号史概観: 古代～中世

- シーザー暗号
  - 文字をアルファベット順にn文字ずらす
    - hello world → IFMMP XPSME
    - 復号化の時は逆方向にn文字ずらす
    - 暗号鍵の数は25種類
- 単一換字暗号
  - 文字を別の文字・記号に1対1変換する: 変換表をつかう
    - 暗号鍵の種類は26!種類
- 15～16世紀ぐらいまではこのレベルの時代が続く
  - (あまり記録が残ってない)
  - そもそも文字を読み書きできる人が少なかった
    - ローマ時代は庶民でも一応読み書きできたらしい
  - 10～12世紀ぐらいは貴族層でも自分の名前が書けないことが多かった

古典暗号史概観: 中世末期～近世

- 平文の特徴を使った暗号解読法が生まれた
  - 文字の出現頻度: 英語ではETAOINSHRDLCUMWFGXPBVKJQYZの順
  - 音韻・綴字法的な特徴: qの次はほぼu、thの順も頻出、ドイツ語だとschが頻出
  - 母音と子音が概ね交互に来る
    - 日本語の場合は特に
    - アラビア語、ベブライ語のように基本的に母音を表記しない言語もある
- 頻度分析
  - 大量の暗号文を入手すれば文字→文字への1対1変換をしている暗号は統計的な攻撃が可能になる
  - 平文の言語が既知or推定できる場合は非常に有効

古典暗号史概観: 中世末期～近世

- 秘匿したい側でも当然対抗する
  - 秘匿度数式暗号
    - 出現頻度の高い文字を複数文字に変換
    - 低い文字は数文字まとめて同じ文字に変換
  - 多表式暗号
    - 変換表を複数使う
    - 1文字ごとに次の変換表に切り替え
- 一方で外交用には隠語による暗号表が19世紀に入るまで使われていた
  - 隠語の選び方によってはステガノグラフィのように機能する
    - Ex)ロシア大使館にはほぼ必ず毛皮商人がいた→毛皮関係の用語を隠語につかう
- 基本的に物理媒体による通信しかなかった

古典暗号史概観: 19世紀後半～

- 有線電信･無線電信の実用化
  - モールス電信機の開発: 1836-38年
  - 大西洋横断ケーブル実用化: 1866年
  - 大西洋横断無線電信: 1901年
    - 初期の無線電信は火花式。数十Hz～数十kHzまでの電波を火花放電で作って発射
  - 電信: モールス符号etcでアルファベット、数字、カナを送受信
    - 基本的に人間が介在する
    - 1930年ごろからテレタイプライタを使用したテレックスが使われ始める
- 電信で利用できる暗号が発展する
  - 物理媒体による通信は媒体を盗まれたり複写したりされなければ問題ない
  - 無線電信は電波が届く範囲ならだれでも受信できる。有線電信も中継を繰り返すので途中で人目に付きやすい

古典暗号史概観: 19世紀後半～

- 明治維新後～台湾出兵ごろまでの日本陸軍は単一換字式暗号を常用
  - 物理的な紙による通信、有線電信だけではあまり問題にならなかった
    - 物理的な紙: 伝令将校、伝令兵、伝書鳩
- 工業レベルが上がってくると、初歩的な暗号機械が出現
- 工業レベルが上がる前は精密な機械は作れてもすべて手作り→当然高い
  - 機械式タイプライタの発明: 1860年ごろ
  - ミシンの原形の発明が1810年ごろ
  - ガソリンエンジンの発明が1801年(2ストローク。実用は1858年)と1862年
  - この辺を作れるようになると、暗号関連も機械の導入が視野に入る

古典暗号史概観: 20世紀前半

- 第1次世界大戦期
  - タイプライターのような構造の機械式暗号機が発明される
- 第2次世界大戦期
  - 第1次大戦期に開発された機械式暗号とその発展形を使用
  - Enigma (ドイツ)
  - Purple (外務省九七式印字機) Jade (海軍九七式印字機), Coral (海軍97式印字機三型) (日本)
  - M-209 (アメリカ)
  - Fialka (ソ連)
- 電子計算機によるクリブ、鍵の力任せの探索の実現により終焉を迎えた