InfoSec Tutorial: Cryptography

Tony Kenyon, CEO.

Revision 1.01.

Updated: Jan 5th 2006

Ref: CNXT0003

Introduction

Definitions

- Terms Used
  - M=message. K=Key. E=Encryption transformation. D=Decryption transformation
- Definitions
  - **Block Cipher**: Plaintext message M divided into 'n' blocks of bytes or bits, then each encrypted using the same algorithm. E(M,K)=E(M1,K) E(M2,K) E(Mn,K)
  - **Cipher**: a cryptographic transformation that operates on characters or bits
  - **Ciphertext/Cryptogram**: an untigelligable (encrypted) message
  - **Clustering**: a message generates the same Ciphertext from the same algorithm buta number of different keys
  - **Codes** : a cryptographic transformation that operates on words or phrases
  - **Cryptology**: encompasses cryptography and cryptanalysis
    - **Cryptanalysis**: breaking the cyphertext (recovering a key or the plaintext)
    - **Cryptography**: hiding meaning. Greek kryptos (hidden) graphein (to write)
  - **Cryptographic Algorithm**: procedure used to encipher plaintext & decipher ciphertext
  - **Cryptosystem**: a set of transformations from a message space to a cyphertext space.
E(M)=C.   D[E(M)]=M
E(M,K)=C. D[E(M,K)]=M   // where there is a dependency on a key K
  - **Decipher & Encipher**:
  - **Key/Cryptovariable**: information or sequence that controls enciphering and deciphering:
  - **Work Function (Factor)**: dificulty of recovering plaintext/key as a function of cost/time.

History

- History
  - 3000bc Egyptions
  - 400bc Spartans (military cryptography)
  - 50bc Romans (Caesar)
  - 815ad Arabs
  - 1460ad Two rotating Cypher Disks. (Italy - Alberti)
  - 1790ad 26 rotating Cypher Disks. (Thomas Jefferson - USA)
  - 1919 Enigma Machine (Hugo Koch – Dutch)
  - 1920 Hagelin Machine M-209 (Sweden – Hagelin)
  - 1920s Herbert Yardley MI-8 Code Crackers USA (father of US cryptography)
  - 1938 Enigma Machine. Germans used 6 rotors instead of 3
  - 1940 Alan Turing, Bletchley Park UK. Breaks Enigma with Bombe
  - *Monoalphabetic* means using the same alphabet as the plaintext

Classical Cryptography

Classical Cryptographic Techniques

- Classical Ciphers
  - Substitution Ciphers
  - Transposition Ciphers
  - Vernam Cipher (One-Time-Pad)
  - Book (Running Key) Ciphers
  - Codes
  - Steganography
- Unix systems use a substitution cypher called ROT 13 (shifts alphabet 13 places)

Classical Ciphers: Substitution

- **Substitution**
  - Caesar System (C3)

    ```
    Zi = Cn(Pi)
    ```

    Zi=cyphertext characters, Cn=monoalphabetic substitution transformation, n=no of letters shifted.
  - **Monoalphabetic** means using the same alphabet as the plaintext
    - Easily attacked using **Frequency Analysis**
  - **Polyalphabetic** ciphers are stronger
    - Vigenere's cipher used 26 alphabets
    - Counters frequency analysis, but can be attacks by discovery of the **periods** (when substitution repeats)

Classical Ciphers: Transposition

- **Transposition**
  - Letters of the plaintext are permuted

    ```
    ATTACKATDAWN
    DCKAAWNATATT
    ```

  - **Columnar transposition**: plaintext written horizontally, read vertically

    ```
     THISISAN
     ENCRYPTE
     DMESSAGE     yeilds    TEDHNMICE...
    ```

  - Can be attacked using frequency analysis; but hides the properties of letter pairs and triples.

Classical Ciphers: One-Time-Pad

- **One-Time-Pad (Vernam Cipher)**
  - Key is a random set of non-repeating characters
  - Each character in the key is added modulo 26 to a plaintext character

    ```
    H+X = 7+23 = 30.      30 MOD 26 is 4 = E.
    H+A = 7+0 = 7.        7 MOD 26 is 7 = H.
    ```
  - Vernam Machine developed by AT&T to use XOR of a message in Baudot code with key bits

Classical Ciphers: One Time Pads

- Advantages
  - Very secure (unbreakable). Not reliant on prime number factoring.
  - Can be very fast (basically XOR of pad + data (so can be used for data streams)
- Issues
  - Length of key char stream = message size.
    - Not really practical for large messages
    - Can be approximated by shorter random sets of characters with very long periods (but then not a true OTP)
  - Key management, storage, distribution and hierarchy
    - Needs good Random Number Generator (RNG)
    - Key Escrow needed if change keys
  - Scalability (limited typically to 30 user community in practise).
    - Both size and also the limits of trusted secret community size.
- Few vendors e.g. AlphaCipher

Classical Ciphers: Book Cipher

- **Book (Running Key) Cipher**
  - Uses text from a book (for example) to encrypt the plaintext
  - Sender and receiver must know the book, page, line no etc.
  - Text matched modulo 26
  - Eliminated periodicity but can be attacked by exploiting redundancy in the key.

Classical Ciphers: Codes

- **Codes**
  - Deal with words and phrases
  - Relate to numbersor letters
  - Example: ’Attack at Dawn’ might be coded as 526

Classical Ciphers: Steganography

- **Steganography**
  - Greek ’steganos’ meaning ’covered’ and ’graphein’ meaning ’to write’
  - Hiding the existence of a message in something else
  - Examples:
    - Microdot: compresses a message into the size of a peiod (dot)
    - Digital Watermark: a form of steganography
    - Hidden Data: redundant pixels in image files used to carry data

Secret Key Cryptography

Secret Key Cryptography

- Symmetric Key Characteristics
  - Sender & receiver have secret key used for both E & D
  - Keys must be changed frequently to maintain security
  - Encryption algorithm E is typically publically known
  - For large key sizes (>128bits) secret key systems are very hard to break
  - Systems also relatively fast for large volumes of data
  - Challenge is to share the key securely, and sender-receiver pairs need a unique key each (scalability)
  - Often teamed with Public Key cryptography to distribute keys and generate unique session keys fot peers with timesatamps (anti-replay)
- Examples
  - DES (CBC, ECB, CFB, OFB), 2DES, 3DES,
  - AES
  - IDEA
  - RC5

Symmetric Key - DES

- Background
  - 1972, derived from Horst Feistel's (IBM) Lucifer alg. US Patent #3,798,539, March, 19, 1974).
  - Adopted by FIPS 46-1 in 1977. Also ANSI X.392 1981
  - Origicanally designed for harware implementation
  - NIST to replace DES with AES
- Characteristics
  - 64-bit block size, 56-bit Key (64 – 8 bits parity)
  - Data Encryption Algorithm (DEA) uses 16-rounds of transposition & substitution
  - Based on Claude Shannon's work (confusion and diffusion)
    - **Confusion**: Non-linear substitution S-boxes (4-bit output from 6-bit input).
    - **Diffusion**: DES uses a Product Cipher 16 times for diffusion in P-boxes: `E1(M)=C1; E2(C1)=C2; ... E16(C15)=C16`
  - Brute force attack requires 2^{56} (70 quadrillion poss keys, 7.2×10^{16})
  - Large numbers of computers can be used to break this, so US Gov abandoned DES in 1998 for 3DES

Symmetric Key - DES

- DES Operates in 4 Modes:
  - Electronic Code Book (ECB)
    - Native DES block mode.
    - Used for small data (initialisation vectors, keys etc.)
    - Divides the 64-bit input vector into a 32-bit Left and Right block
    - Blocks copied to make two 48-bit blocks, which are each XORd with a 48-bit key
  - Cypher Block Chaining (CBC)
    - Plaintext block of 64-bits
    - Random 64-bit initialisation vector used to XOR block 1
    - Result encrypted with the DES key
    - Resulting Ciphertext XORd with the next 64-bit plaintext block
    - Continues until plaintext exhausted
    - Note that errors propagate in this mode
  - Cipher Feedback (CFB)
    - Stream Cipher, ciphertext used as feedback into key generation source to develop the next key stream
    - Errors will propagate in this mode
  - Output Feedback (OFB)
    - Stream cipher, XORs plaintext with the keystream.
    - Uses initialisation vector, feedback used to generate the key stream
    - Errors will not proagate in this mode.

Symmetric Key - 3DES

- Background
  - FIPS 46-3
- Characteristics
  - Encrypts a message 3 times
    - Merkle & Hellman have shown that encrypting plaintext twice can be broken in 2^{n+1} attempts by a Meet-In-The-Middle Attack
  - Several ways of doing this:
    - DES-EDE2 `[E{D[E(M,K1)],K2},K1]`
    - DES-EEE2 `[E{E[E(M,K1)],K2},K1]`
    - DES-EEE3 `[E{E[E(M,K1)],K2},K3]`
    - The latter DES-EEE3 is the most secure

Symmetric Key - AES

- Background
  - Block cipher to replace DES, though 3DES likley to remain approved for US Gov use
  - Rijndael Block Cipher selected on Oct 2nd 2000 by NIST for AES
  - FIPS PUB 197. AES-128, AES-192, AES-256.
- Design aims
  - Resistance against all known attacks
  - Design simplicity
  - Small code footprint and speed on multiple platforms
  - Suitable for implementations on High-Speed chips with no area restrictions, and for a compact co-processor on a smart card
- Characteristics
  - **Iterated Block Cypher**, **variable block length**
  - Key length of either 128, 192 or 256 bits with fixed block size of 128 bits

Symmetric Key - AES

- Robustness
  - 3.4×10^{38} possible 128-bit keys
  - 7.2×10^{57} possible 192-bit keys
  - 1.1×10^{77} possible 256-bit keys
  - E.g. If a computer can crack DES at 2^{56} keys per second, it would require 149 trillion years (149×10^{12}) to crack Rijndael.
    - Note that the Universe is estimated to be < 20 Billion (20×10^{9}) years old
- Rijndael employs a round transformation of 3 layers of distinct and invertible transformations (instead of a Feistel network).
  - Non Linear Layer: Parellel S-boxes, worst case non-linear perf
  - Linear Mixing Layer: high diffusion of multple rounds
  - Key Addition Layer: XOR of the Round Key to the Intermediate State
- Round key bit = BlockLength x (NoOfRounds + 1)
- NoOfRounds is a function of the key size (256=14, 192=12, 128=10)

Symmetric Key – Twofish

- Uses a **Feistel** network
  - Each round, half of the 12-bit block is fed into an F-function box which is XORd with the other half of the text in the network.
  - The half block is broken into 32-bit units, themselves broken into 4-bytes, each fed into four different S-boxes
  - Emerging 4 bytes then combined in a Maximum Distance Separable (MDS) matrix to form two 32-bit units.
  - The two 32-bit units are combined using a Pseudo-Hadamard Transform (PHT) and added to the two round subkeys.
  - The PHT is linear ($d_1=(2b_1+b_2)\mathrm{mod}\ 256$)
  - Results XORd with right half of the 64-bits of plaintext. 1-bit rotations perfomed before and after the XOR.
  - These operations are repeated for 15 more rounds
- Also does **pre** and **post-whitening** where additional subkeys areXORd with the plaintext before the 1st and after the 16th round
  - Means that the whitening keys also have to be uncovered during crptanalysis

Symmetric Key – IDEA

- Characteristics
  - 1992, developed by James Massey and Xuejia Lai
  - 64-bit Block encryption using a 128-bit key
  - Uses confusion and diffusion
  - Perform 8 rounds on 16-bit sub blocks using algebraic calculations suited to hardware implementations
    - Modulo 2^{16} addition, modulo $2^{16} + 1$ multiplication and XOR
  - Much harder to crack than DES (128-bit key)
  - Operates in the modes applied to DES, used in PGP

source: www.cyphernetix.co.uk

Symmetric Key – RC5

- Characteristics
  - Block cypher. 1994 Ronald Rivest. Patented in 1997 by RSA
  - Block length. Typical block sizes are 32,64,128-bits
  - Key size variable (0-2048)
  - No of rounds variable (0-255)
  - Encrypts using integre addition, application of bit-wise XOR, and variable rotations

Public Key Cryptography

Public Key Cryptosystems

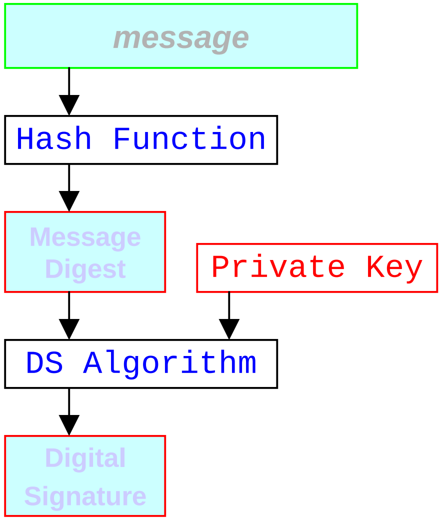
- One way Functions
- Public Key Algorithms
  - RSA
  - Diffie-Helman
  - El Gamal
  - Merkle-Hellman Knapsack
  - Elliptic Curve (EC)
  - Digital Signatures
    - DSS and SHS
    - MD5
    - HMAC

PKC – Digital Signatures

- Digital Signatures are used to:
  - **Integrity**: Detect unauthorised modifications of the message
  - **Authenticity**: Authenticate the signatory
  - **Non-repudiation**: Prove to a 3rd party that the signature was generated by the signatory

PKC – Digital Signatures

- Sending a message with a Digital Signature
  - Sender uses **Hash Function** on message to generate **Message Digest**
  - Use **Digital Signature Algorithm** (DSA) on the digest to produce a **Digital Signature**
    - Senders **Private Key** is used to encrypt the MD, and the resulting DS is attached to the message
  - Receiver uses the sender's **public key** to decrypt the Message Digest
    - Then uses the same **hash function** on the message to check that the two MDs are consistent
    - Receiver then knows that the message is intact, and that the sender is who they say they are.



PKC – DSS & SHS

- Digital Signature Standards and Secure Hash Standard
  - NIST FIPS 186-1 specifies use of RSA DSA
  - RSA DSA is based on El Gamal, both use SHA-1 (FIPS-180)
  - SHA-1 generates a fixed length (160-bit) digest from variable length message, which is then processed by DSA to generate or verify a signature
  - For short messages padding is used to create 512-bit blocks

PKC – MD5

- **MD algorithm** developed by Ronald Rivest in 1991.
- Takes variable length message, creates a **128-bit message digest**
- Uses **512-bit** blocks in 4 rounds

PKC – Hash Algorithms

- Properties
  - H(M1) != H(M2) else could have a **birthday attack**
- HMAC
  - **Hash algorithm** that uses a key to generate a Message Authentication Code (MAC)
  - This is a type of **checksum**, generated before message is sent

Review of Message Digest Algorithms

Message Digest Algorithms

- MD5, Snefru, MD4, MD2, SHA, and Haval
  - message-digest algorithms (also known as one-way hash functions, fingerprinting routines, message authentication codes, or manipulation detection codes).
  - use cryptographic techniques to ensure that any small change in input stream results in immediate and widely diverging output. Therefore, any unauthorized, malicious, or accidental change will be rapidly discovered.
  - since these algorithms use a 128-bit (or larger) signature, a brute-force attack to introduce a deliberate change in the file while trying to keep the same signature becomes computationally infeasible.
- CRC algorithms
  - use simple **polynomial division** to generate checksums.
  - Very fast, and the maths is simple and well-understood.
  - since the signature space is so small (usually 16 or 32 bits), a brute-force search for a CRC collision is well within the capabilities of most workstations.
  - currently several programs in the public domain that can, for any given input file, provide a different output file with the same CRC signature in 30 seconds or less.

source: www.cyphernetix.co.uk

Message Digest Algorithms: MD5

- Overview
  - Owned by RSA Data Security, Inc.
  - Currently considered by many to be a state-of-the-art signature algorithm, and a proposed data authentication standard.
  - MD5 attempts to address potential security risks found in the speedier, but less secure MD4. It is a more conservative algorithm from the perspective of of cryptanalytic risk.
  - Generates a 128-bit signature, and uses four rounds to ensure pseudo-random output.
- Further Details
  - See Internet Working Draft RFC 1321, via anonymous FTP from NIC.DDN.MIL or from RSA.COM as ~/pub/md5.doc.

Message Digest Algorithms: CRC

- Cyclic Redundancy Checks are long-established facto error detection algorithm standards.
- Algorithms are simple, fast, robust, and provide reliable detection of errors associated with data transmission.
- **CRC-16**
  - CRC-16 is the predecessor to CRC-32, using only 16 bits to store to the remainder of the data and the generator polynomial.
  - CRC-16 is typically at the link level, usually done in hardware to detect transmission errors.
  - A little faster than CRC32 (around two fifths of MD4 performance)
- **CRC-32**
  - CRC-32 is provided as a fast and speedy alternative to the slower message-digest algorithms.
  - It has been shown that CRC-32 has a minimum distance of 5 for block lengths of less than 4K. However, this decreases as the size of the blocks increases. Therefore, using CRC-32 on long files is a misapplication of this signature algorithm.
  - Fast (about one third of MD4 performance)

Message Digest Algorithms: MD4

- Overview
  - Developed by RSA Data Security, Inc.
  - Predecessor to MD5.
  - Submitted as a standard data authentication algorithm, and is described in the Internet Working Draft 1320.
  - The MD4 algorithm was designed to exploit 32-bit RISC architectures to maximize throughput.
  - Very fast.
- Further Details
  - MD4 can be obtained via anonymous FTP from RSA.COM in ~/pub.

Message Digest Algorithms: MD2

- Overview
  - Developed by RSA Data Security, Inc.
  - Created as part of the Privacy Enhanced Mail package - designed to authenticate and increase security of email.
  - MD2 generates a 128-bit signature.
  - The license for MD2 specifically states its use is exclusive to the Privacy Enhanced Mail package.
  - Provisions have been made with RSA Data Security, Inc. for its inclusion and use in Tripwire in its present form.
  - Quite slow. Unclear if using this slower algorithm instead of MD5 brings any comparative advantage.
- Further Details
  - MD2 is not in the public domain.

Message Digest Algorithms: SHA/SHS

- Overview
  - The NIST Digital Signature Standard, called the Secure Hash Standard.
  - SHA is about one-half as fast as MD5. It has been noted that SHS appears to be largely based on MD4 with several key enhancements, not all implemented in MD5.
  - In mid-1994 it was patched at the behest of NSA
    - raised questions about the overall strength of both SHA and MD4 in the minds of some cryptographers.
- Further Details
  - See NIST FIPS 180.

Message Digest Algorithms: Haval

- Overview
  - Developed by Yuliang Zheng at the University of Wollongong.
  - Configured as 128 bit signature using four passes
  - Pretty fast (similar to CRC32, approx 30% faster than MD5)
- Further Details
  - See Y. Zheng, J. Pieprzyk and J. Seberry: “HAVAL - a one-way hashing algorithm with variable length of output”, Advances in Cryptology - AUSCRYPT'92, Lecture Notes in Computer Science, Springer-Verlag, 1993.

Public Key Infrastructure (PKI)

- Includes
  - Digital Certificates (X.509)
  - Certificate Authority (CA)
  - Registration Authorities 9RA)
  - Policies and Procedures
  - Certificate Revocation
  - Non-Repudiation support
  - Time stamping
  - LDAP
  - Security-enabled applications
  - Cross-certification
- See PKI Forum www.pkiforum.org

Key management & Escrow

Key Management

- Key Management Issues
  - Key control measures
  - Key recovery
  - Key storage
  - Key retirement/destruction
  - Key change
  - Key generation
  - Key theft
  - Frequency of use

Escrowed Encryption

- NIST FIPS-185 EES 1994. Implemented as Clipper Chip
- Fair Cryptosystems (MIT 1992)
  - Public/Private keys split and escrowed in many parts
  - Patents purchased by Bankers Trust

Escrow - Clipper

- NIST FIPS-185 EES
  - Divides key into 2 parts and escrow each part to separate trusted bodies
  - Law enforcement organisation require a court order to get these two parts
  - EES is built into the US Gov’s **Clipper Chip**, using tamper proof hardware
    - Uses **Skipjack** Secret Key algorithm
    - Each chip has a unique **serial no**, and a **80-bit** unique **Unit** (secret) key
    - Unit key divided into **two parts** to be stored separately, with the serial no
    - Organisations can exchange info using a shared **session key** Ks, via DH or RSA key exchange, along with a Law Enforcement Access Field (**LEAF**)
      - LEAF is encrypted with the Clipper **family key**, and holds:
        - » Ks,encrypted with the secret key u
        - » Serial number of sending Clipper chip
        - » An authentication string
    - 80-bit Unit (secret) key is seen as weak and has raised concerns

Applications

Random Number Generation

Random Number Generation (RNG)

- A good RNG has three properties:
  - Evenly distributed numbers
  - Values are unpredictable
  - Long and complete cycle
- Don’t
  - use linear congruential functions, such as the CRT rand() function. rand() fails the second test miserably!

- Win32 offers CryptGenRandom() - see WinCrypt.h
  - Satisfies the first two requirements
  - Uses FIPS 186 pseudorandom number generator using SHA-1 as a G function
  - Uses system entropy (current PID, current Thread ID, GetTickCount(), GetLocalTime(), QueryPerformanceCounter(), MD4 hash of users environment block, High precision internal CPU counters (RDTSC, RDMSR, RDPMC on x86 platforms) low level system info, system exception info, system lookaside info, system interrupt info, system process info).
  - Bytestream hashed with SHA-1 to produce a 20-byte value used to generate RN according to FIPS 186-2 App 3.1

Internet Security Applications

- MAC of the FIMAS
- Secure Electronic Transaction (SET)
- SSL/TLS
- Internet Open Trading Protocol IOTP)
- MONDEX
- IPSec
- S-HTTP
- SSH-2
- Wireless Security
  - WAP
  - IEEE 802.11

Email Security

- Objectives
  - Non-repudiation
  - Messages read by intended recipients
  - Message integrity
  - Authentication of source
  - Verification of delivery
  - Labelling of sensitive info
  - Control of access
- Standards
  - S/MIME
  - MIME Object Security Service (MOSS)
  - Privacy Enhanced Mail (PEM)
  - Pretty Good Privacy (PGP)

Threats and Attacks

Cryptographic Attacks

- Obtaining plaintext from a ciphertext - Common attacks
  - **Brute force**:
  - **Known plaintext**: attacker has a copy of the plaintext
  - **Chosen plaintext**: chosen plaintext is encrypted and the output ciphertext used
  - **Adaptive chosen plaintext**: the text selected is altered based on previous results
  - **Ciphertext only**: only the ciphertext is available
  - **Chosen ciphertext**: portions of the ciphertext selected for encrypted with a decrypted plaintext
  - **Adaptive chosen ciphertext**: portions of the ciphertext chosen for encryption are based on prior results
  - **Meet-In-The-Middle**: applies to double encryption schemes, where known plaintext is encrypted with each possible key and compared with results in the middle, with decryption of the ciphertext using each possible key
  - **Man-In-The-Middle**:intercepting messages in a store and forward system
  - **Differential Cryptanalysis**:
  - **Differential Linear Cryptanalysis**:analysing key pairs in private key crypto
  - **Factoring**:mathematically determining prime factors or large numbers
  - **Statistical**:exploiting lack of randomness in key generation

Key Length & Performance

Example Key Lengths

- **Classical**
  - One Time Pad (OTP) key length >= message size
- **Symetric Encryption**
  - DES - 56 bit + 8 bit parity
  - 3DES - 112 bit + 16 bit parity
  - IDEA - tba
  - Blowfish - 128 bit
  - RC4 - 40 to 128 bit
  - RC2 - 40 to 128 bit
  - AES - 128, 192, 256 bit
- **Asymetric Encryption**
  - RSA - 512 to 2048 bit
  - DSA - tba
  - ElGamal - tba
  - Elliptic Curve - 160 bit
- **Hash Algorithms**
  - MD5 - tba
  - MD4 – tba
  - MD2 – tba
  - SHA/SHS – tba
  - Snefru - tba
  - HAVAL - tba

Cryptography: Performance

- **Hash Algorithm Performance**
  - Speed (Kilo Bytes per Sec)

MD4	- 332
CRC16	- 131
CRC32	- 111
HAVAL	- 100
MD5	- 70
SHA/SHS	- 35
Snefru	- 31
MD2	- 3

  - Source: Tripwire
    - All observed timing measures were performed on a Sequent Symmetry with ten 16 MHz 80386 processors. The numbers provided are simply an informal gauge of throughput, rather than any authoritative metric.
    - For MD4 on a Sun SparcStation, throughput rates of over 1.4 Mbytes/second are achieved.

Questions?