"The best system is to use a simple, well understood algorithm which relies on the security of a key rather than the algorithm itself. This means if anybody steals a key, you could just roll another and they have to start all over."
-- Andrew Carol

Symmetric key cryptography

- preliminaries (operational and attacker models)
- block ciphers (basics, DES, 3DES, AES)
- block ciphers in practice (modes of operation)
- a security flaw induced by CBC padding
- stream ciphers

Large numbers

time until next ice age……………………………… 2^{39} seconds
time until the sun goes nova………………… 2^{55} seconds
age of the planet…………………………………… 2^{55} seconds
age of the Universe……………………………… 2^{59} seconds

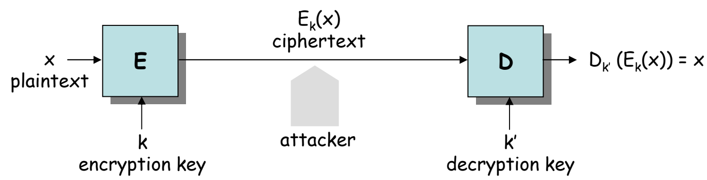
number of atoms in the planet……………… 2^{170}
number of atoms in the sun………………… 2^{190}
number of atoms in the galaxy…………… 2^{223}
number of atoms in the Universe ………. 2^{265}
(dark matter excluded)

volume of the universe…………………………… 2^{280} cm^3

(source: Schneier, Applied Cryptography, 2nd ed., Wiley 1996)

© Levente Buttyán

Operational model of encryption



- attacker's goal:
  - to systematically recover plaintext from ciphertext
  - to deduce the (decryption) key
- Kerckhoff's assumption:
  - attacker knows all details of E and D
  - attacker doesn't know the (decryption) key

© Levente Buttyán

Attack models

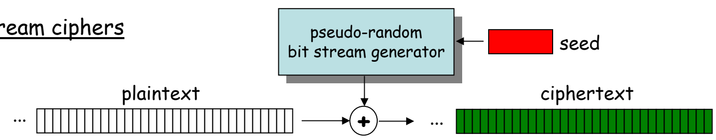
- ciphertext-only
  - only data transmitted over the ciphertext channel is available to the attacker
- known-plaintext
  - plaintext-ciphertext pairs are available to the attacker
- chosen-plaintext
  - ciphertexts are available corresponding to plaintexts of the attacker's choice
  - adaptive: choice of plaintexts may depend on previously obtained plaintext-ciphertext pairs
- chosen-ciphertext
  - plaintext-ciphertext pairs are available for some number of ciphertexts of the attacker's choice
  - adaptive: choice of ciphertexts may depend on previously obtained plaintext-ciphertext pairs
- related-key attack
  - attacker has access to the encryption of plaintexts under both the unknown key and keys known to have certain relationship with the unknown key

© Levente Buttyán

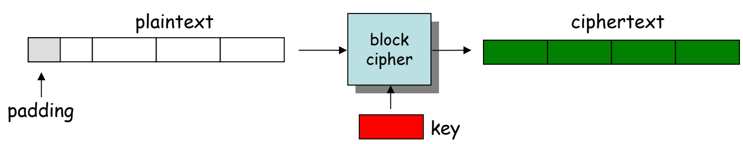
Asymmetric- vs. symmetric-key encryption

- asymmetric-key encryption
  - it is hard (computationally infeasible) to compute k' from k
  - k can be made public (public-key cryptography)
- symmetric-key encryption
  - it is easy to compute k from k' (and vice versa)
  - often k = k'
  - two main types: stream ciphers and block ciphers

stream ciphers



block ciphers



© Levente Buttyán

Theoretical vs. practical security

- one-time pad
  - a stream cipher where the key stream is a true random bit stream
  - unconditionally secure [Shannon, 1949]
  - however, the key must be as long as the plaintext to be encrypted
- practical ciphers
  - use much shorter keys (128 bits (symm.), 1024 bits (asymm.))
  - they are not unconditionally secure, but computationally infeasible to break (practically secure)
  - however, proving that a cipher is practically secure is not easy
    - not enough to consider *brute force attacks* (key size) only
    - a cipher may be broken due to weaknesses in its (algebraic) structure
  - no proofs of security exist for many ciphers used in practice
  - if a proof exists, it usually relies on assumptions that are widely believed to be true (such as $P \neq NP$)

© Levente Buttyán

Attack complexity

- best available measure of security for practical ciphers is the complexity of the best (currently) known attack
- attack complexity is the dominant of the following complexities
  - data complexity
    - expected number of input data units required for the attack (e.g., number of ciphertexts in a ciphertext-only attack)
  - storage complexity
    - expected number of storage units required
  - processing complexity
    - expected number of operations required to process input data and/or fill storage with data
    - parallelization may reduce attack time but not processing complexity

© Levente Buttyán

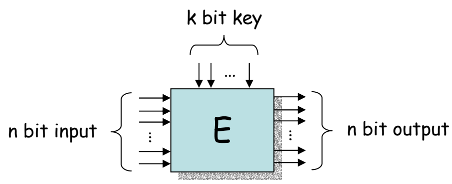
"Feistel and Coppersmith rule: Sixteen rounds and one hell of an avalanche."
-- Stephan Eisvogel in de.comp.security

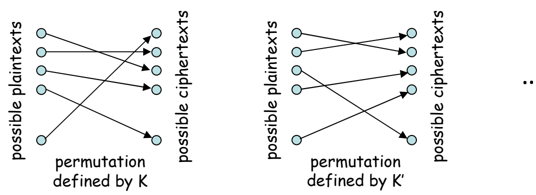
Block ciphers: DES, 3DES, and AES

- basics
- operation of DES
- cryptanalysis of DES
- multiple encryption and the 3DES
- AES

Block ciphers

- an *n* bit block cipher is a function $E: \{0, 1\}^n \times \{0, 1\}^k \rightarrow \{0, 1\}^n$, such that for each $K \in \{0, 1\}^k$, $E(X, K) = E_K(X)$
  - is an invertible mapping from $\{0, 1\}^n$ to $\{0, 1\}^n$
  - cannot be efficiently distinguished from a random permutation
- the inverse of $E_K(X)$ is denoted by $D_K(Y)$, where $Y = E_K(X)$





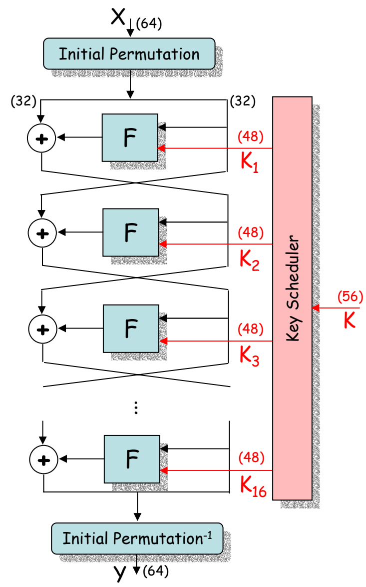
© Levente Buttyán

Exhaustive key search and key size

- given a small number of plaintext-ciphertext pairs encrypted under a key K, K can be recovered by exhaustive key search with 2^{k-1} processing complexity (expected number of operations)
  - input: (X, Y), (X', Y'), …
  - progress through the entire key space
    - for each trial key K', decrypt Y
    - if the result is not X, then throw away K'
    - if the result is X, then check the other pairs (X', Y'), …
    - if K' does not work for at least one pair, then throw away K'
  - if K' worked for all pairs (X, Y), (X', Y'), …, then output K' as the target key
  - on average, the target key is found after searching half of the key space
- if the plaintexts are known to contain redundancy, then ciphertext-only exhaustive key search is possible with a relatively small number of ciphertexts

© Levente Buttyán

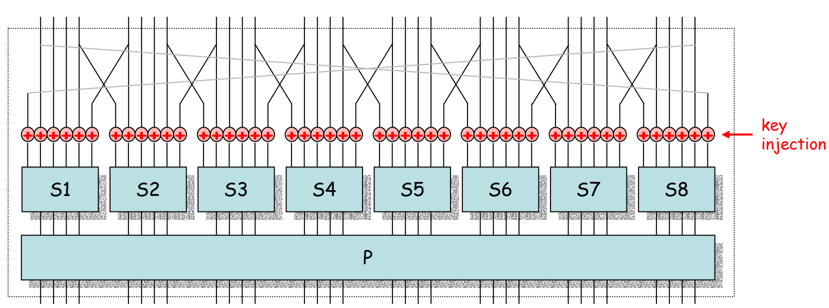
DES – Data Encryption Standard



- input size: 64
- output size: 64
- key size: 56
- 16 rounds
- Feistel structure

© Levente Buttyán

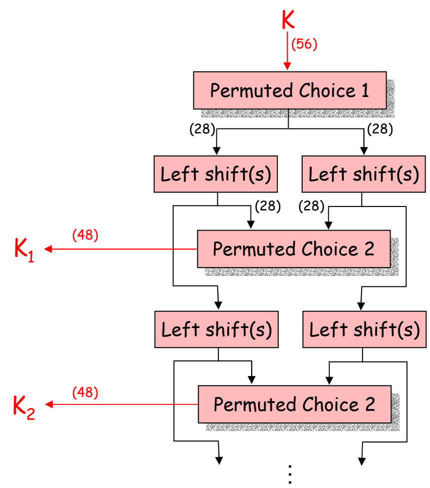
DES round function F



- Si – substitution box (S-box)
- P – permutation box (P-box)

© Levente Buttyán

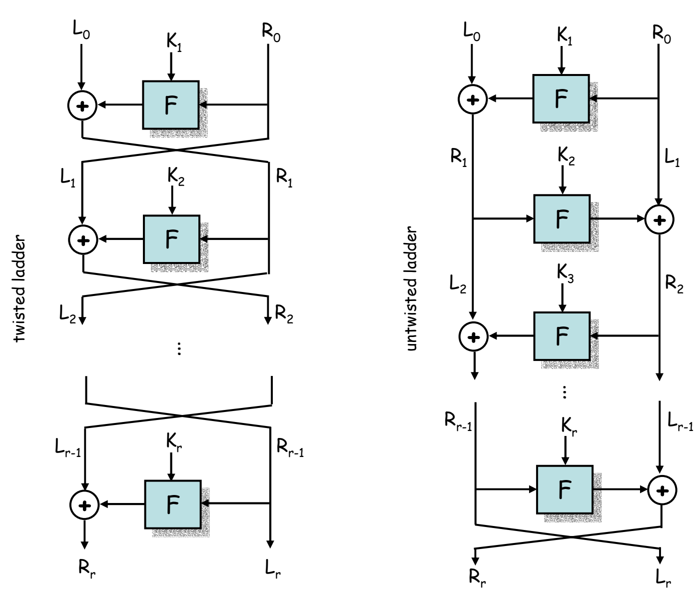
DES key scheduler



- each key bit is used in around 14 out of 16 rounds

© Levente Buttyán

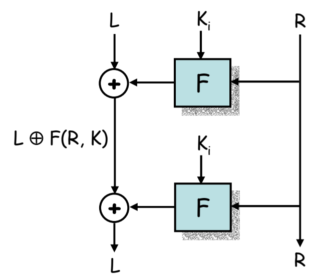
Feistel structure illustrated



© Levente Buttyán

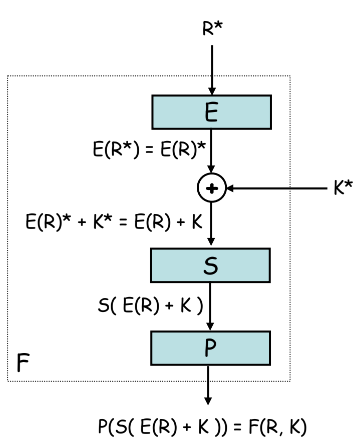
Properties of Feistel ciphers

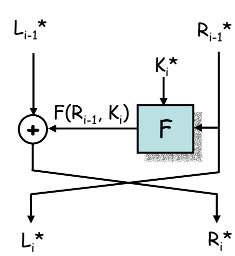
- round i maps (L_{i-1}, R_{i-1}) into (L_i, R_i) as follows:
  $L_i = R_{i-1}$
  $R_i = L_{i-1} \oplus F(R_{i-1}, K_i)$
- a Feistel cipher is always invertible even if F is not invertible:
  $R_{i-1} = L_i$
  $L_{i-1} = R_i \oplus F(R_{i-1}, K_i) = R_i \oplus F(L_i, K_i)$
- decryption can be achieved using the same r-round process with the round keys used in reverse order (K_r through K_1)



Complementation property of DES

- $Y = DES_K(X)$ implies $Y^* = DES_{K^*}(X^*)$
  - where X* denotes the bitwise complement of X





Consequences of the complementation prop.

- assume an attacker can mount a chosen-plaintext attack
- the attacker chooses a plaintext X, and obtains $Y_1 = DES_K(X)$ and $Y_2 = DES_K(X^*)$
- by the complementation property, the attacker knows that $DES_{K^*}(X) = Y_2^*$
- the attacker then runs an exhaustive key search
  - for each trial key K', he computes $Y' = DES_{K'}(X)$
    - if $Y' = Y_1$, then K' is possibly the target key (should be further tested)
    - if $Y' = Y_2^*$, then K'* is possibly the target key (should be further tested)
    - otherwise throw away **both** K' and K'*
- expected number of keys required before success is reduced from 2^{55} to 2^{54}
- still impractical as an attack

© Levente Buttyán

DES weak keys and semi-weak keys

- a weak key is a key K such that $DES_K(DES_K(X)) = X$
  - there are 4 DES weak keys:

```
0101 0101 0101 0101
FEFE FEFE FEFE FEFE
1F1F 1F1F 0E0E 0E0E
E0E0 E0E0 F1F1 F1F1
```

- a semi-weak key pair is a pair (K_1, K_2) such that $DES_{K_1}(DES_{K_2}(X)) = X$
  - there are 6 pairs of DES semi-weak keys
- why are these keys weak?
  - for each weak key K, there exist 2^{32} fix points of DES_K, i.e., plaintext X such that $DES_K(X) = X$
  - for 4 out of the 12 semi-weak keys, there exist 2^{32} anti-fix points, i.e., plaintext X such that $DES_K(X) = X^*$

© Levente Buttyán

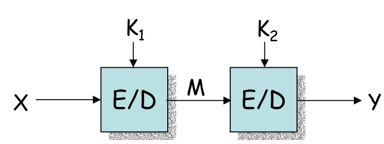
Linear and differential cryptanalysis of DES

- linear cryptanalysis (LC)
  - linear cryptanalysis is the most powerful attack against DES to date
  - requires an enormous number (~2^{43}) known plaintext-ciphertext pairs → infeasible in practical environments
  - could work in a ciphertext only model if plaintexts are redundant (e.g., contain parity bits)
- differential cryptanalysis (DC)
  - most general cryptanalytic tool to date against iterated block ciphers (including DES, FEAL, IDEA)
  - primarily a chosen-plaintext attack
  - in case of DES, it requires ~2^{47} chosen plaintext-ciphertext pairs → infeasible in practical environments
- DES was optimized against DC when it was designed
- it can, however, be improved with respect to LC (apparently the designers of DES was not aware of this attack at that time)

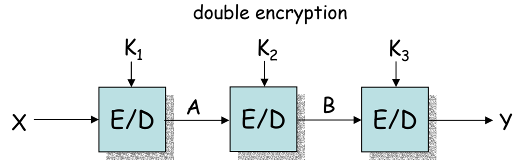
© Levente Buttyán

Multiple encryption and 3DES

- if a block cipher is susceptible to exhaustive key search (e.g., DES), then encryption of the same message more than once may increase security



double encryption



triple encryption

  - stage keys may not be independent
    - e.g., two-key 3DES: $K_1 = K_3$
  - a stage cipher may be either a block cipher or its corresponding decryption function
    - e.g., 3DES-EDE (encryption-decryption-encryption)

© Levente Buttyán

DES is not a group

- group property
  - given any two keys K1 and K2, there exists a third key K3, such that $DES_{K1}(DES_{K2}(X)) = DES_{K3}(X)$ **for all X**
- if DES was a group, then multiple encryption would be equivalent to single encryption → approaches like 3DES to strengthen DES would be useless

© Levente Buttyán

Meet-in-the-middle attack on double enc.

- a naïve exhaustive key search attack on double encryption tries all 2^{2k} keys
- a known-plaintext meet-in-the-middle attack defeats double encryption using an order of 2^k operations and 2^k storage
  - attack time is reduced at the cost of substantial space
- meet-in-the-middle attack:
  - input: known plaintext-ciphertext pairs (X, Y), (X', Y'), …
  - compute $M_i = E_i(X)$ for all possible key values $K_1 = i$ and store all (M_i, i) pairs in a table
  - compute $M'_j = D_j(Y)$ for all possible key values $K_2 = j$ and check for hits $M'_j = M_i$ against entries in the stored table
    - M'j need not be stored, it can be checked as it is generated
  - each hit identifies a candidate solution key pair (i, j)
  - using a second plaintext-ciphertext pair (X', Y'), discard false hits
  - for an L stage cascade of random ciphers, the expected number of false key hits when t plaintext-ciphertext pairs are available is 2^{Lk-tn}, where n and k are the block and key sizes, resp.

© Levente Buttyán

AES – Advanced Encryption Standard

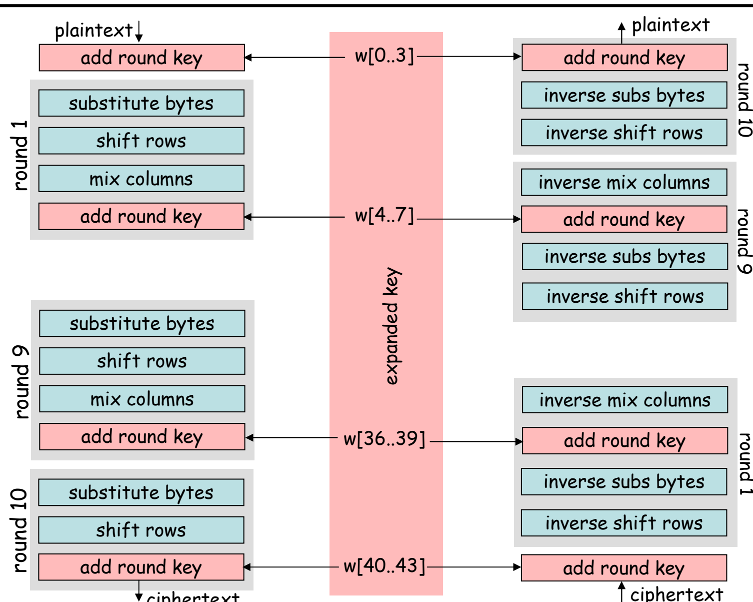
- NIST selected Rijndael (designed by Joan Daemen and Vincent Rijmen) as a successor of DES (3DES) in November 2001
- Rijndael parameters

– key size	128	192	256
– input/output size	128	128	128
– number of rounds	10	12	14
– round key size	128	128	128

- not Feistel structure
- decryption algorithm is different from encryption algorithm (optimized for encryption)
- single 8 bit to 8 bit S-box
- key injection (bitwise XOR)

© Levente Buttyán

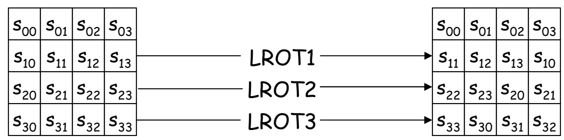
General structure of encryption/decryption



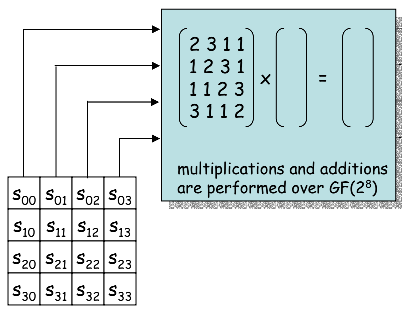
© Levente Buttyán

Shift row and mix column

shift row



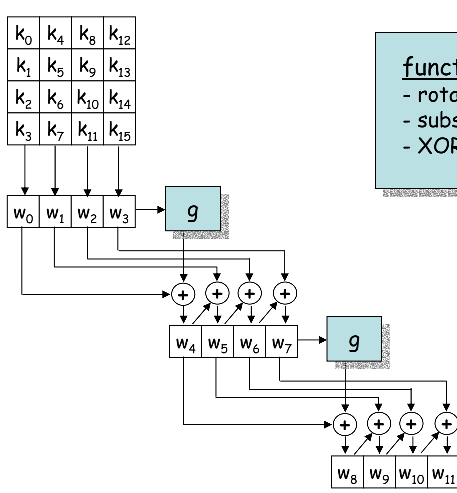
mix column



© Levente Buttyán

Block ciphers / AES (Rijndael)

Key expansion



function g
- rotate word
- substitute bytes
- XOR with round constant

...

© Levente Buttyán

Summary

- block cipher basics
  - trade-offs in block size
  - trade-offs in key size, exhaustive key search
  - product ciphers, SP networks
- DES
  - operation
  - properties (Feistel structure, complementation, weak keys)
  - differential and linear cryptanalysis
  - multiple encryption and the 3DES
  - meet-in-the middle attack on 2DES
- AES
  - operation

Using a block cipher in practice: modes of operation

- ECB
- CBC
- CFB, OFB, CTR

Block cipher modes of operation

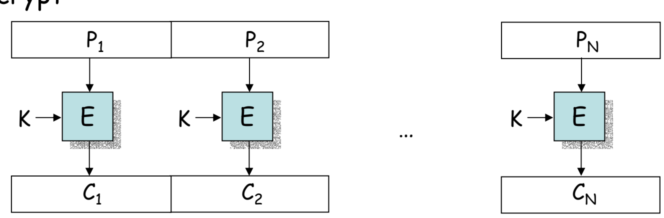
- ECB – Electronic Codebook
  - used to encipher a single plaintext block (e.g., a DES key)
- CBC – Cipher Block Chaining
  - repeated use of the encryption algorithm to encipher a message consisting of many blocks
- CFB – Cipher Feedback
  - used to encipher a stream of characters, dealing with each character as it comes
- OFB – Output Feedback
  - another method of stream encryption, used on noisy channels
- CTR – Counter
  - simplified OFB with certain advantages

Block ciphers / Modes of operation

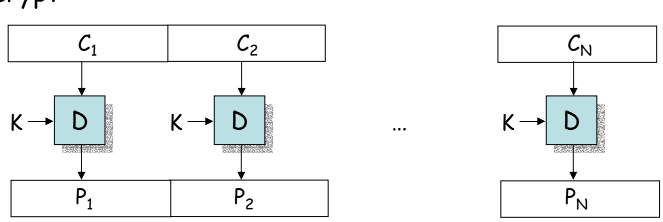
© Levente Buttyán

ECB mode

- encrypt



- decrypt



Block ciphers / Modes of operation

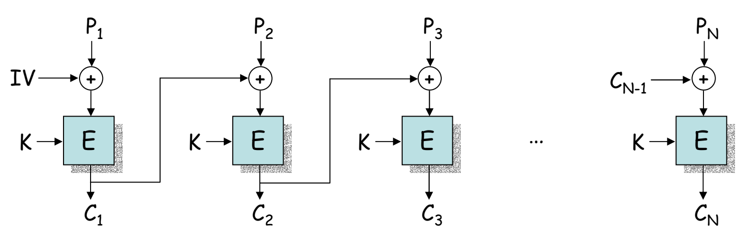
© Levente Buttyán

Properties of ECB mode

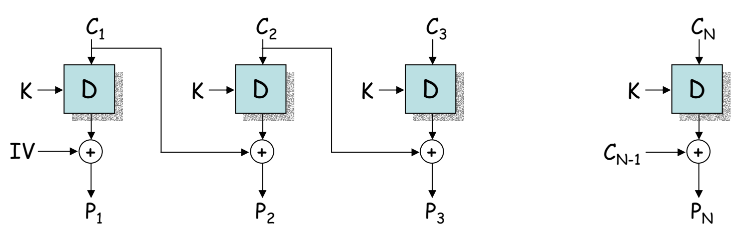
- encrypting the same plaintext with the same key results in the same ciphertext
- identical plaintext blocks result in identical ciphertext blocks (under the same key of course)
  - messages to be encrypted often have very regular formats
  - repeating fragments, special headers, string of 0s, etc. are quite common
- blocks are encrypted independently of other blocks
  - reordering ciphertext blocks result in correspondingly reordered plaintext blocks
  - ciphertext blocks can be cut from one message and pasted in another, possibly without detection
- error propagation: one bit error in a ciphertext block affects only the corresponding plaintext block (results in garbage)
- overall: not recommended for messages longer than one block, or if keys are reused for more than one block

CBC mode

- encrypt



- decrypt



Properties of CBC mode

- encrypting the same plaintexts under the same key, but different IVs result in different ciphertexts
- ciphertext block C_j depends on P_j and all preceding plaintext blocks
  - rearranging ciphertext blocks affects decryption
  - however, dependency on the preceding plaintext blocks is only via the previous ciphertext block C_{j-1}
  - proper decryption of a correct ciphertext block needs a correct preceding ciphertext block only
- error propagation:
  - one bit error in a ciphertext block C_j has an effect on the j-th and (j+1)-st plaintext block
    - P_j' is complete garbage and P_{j+1}' has bit errors where C_j had
    - an attacker may cause predictable bit changes in the (j+1)-st plaintext block
- error recovery:
  - recovers from bit errors (self-synchronizing)

© Levente Buttyán

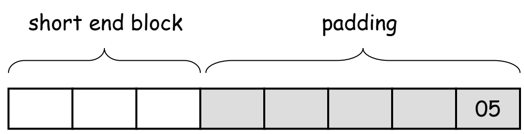
Integrity of the IV in CBC mode

- the IV need not be secret, but its integrity should be protected
  - malicious modification of the IV allows an attacker to make predictable changes to the first plaintext block recovered
- one solution is to send the IV in an encrypted form at the beginning of the CBC encrypted message

© Levente Buttyán

Padding

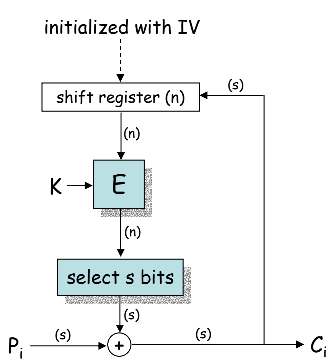
- the length of the message may not be a multiple of the block size of the cipher
- one can add some extra bytes to the short end block until it reaches the correct size – this is called padding
- usually the last byte indicates the number of padding bytes added – this allows the receiver to remove the padding



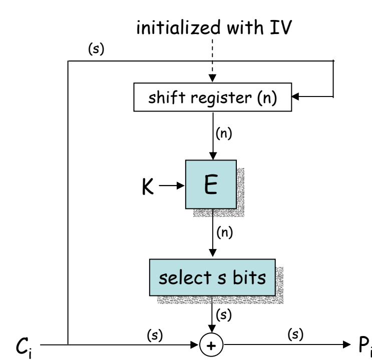
- note: if the encrypted message must have the same size as the clear message, then no padding can be used
  - encrypt the last ciphertext block again
  - select m bits and XOR them to the remaining m bits of the clear message

CFB mode

- encrypt



- decrypt

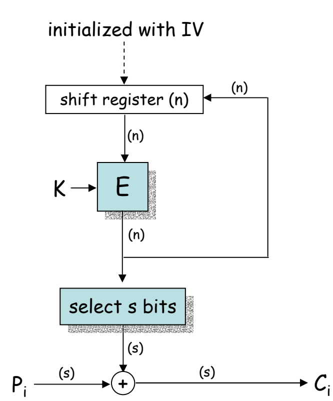


Properties of CFB mode

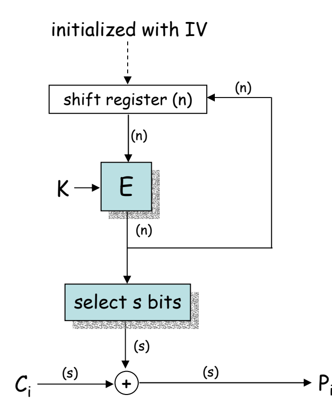
- encrypting the same plaintexts under the same key, but different IVs result in different ciphertexts
- the IV can be sent in clear
- ciphertext character C_j depends on P_j and all preceding plaintext characters
  - rearranging ciphertext characters affects decryption
  - proper decryption of a correct ciphertext character needs the preceding n/s ciphertext characters to be correct
- error propagation:
  - one bit error in a ciphertext character C_j has an effect on the decryption of that and the next n/s ciphertext characters (the error remains in the shift register for n/s steps)
    - P_j' has bit errors where C_j had, all the other erroneous plaintext characters are garbage
    - an attacker may cause predictable bit changes in the j-th plaintext character
- error recovery:
  - self synchronizing, but requires n/s characters to recover

OFB mode

- encrypt



- decrypt



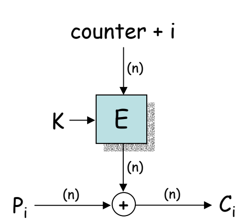
Properties of OFB mode

- a different IV should be used for every new message, otherwise messages will be encrypted with the same key stream
- the IV can be sent in clear
  - however, if the IV is modified by the attacker, then the cipher will never recover (unlike CFB)
- ciphertext character C_j depends on P_j only (does not depend on the preceding plaintext characters)
  - however, rearranging ciphertext characters affects decryption
- error propagation:
  - one bit error in a ciphertext character C_j has an effect on the decryption of only that ciphertext character
    - P_j' has bit errors where C_j had
    - an attacker may cause predictable bit changes in the j-th plaintext character
- error recovery:
  - recovers from bit errors

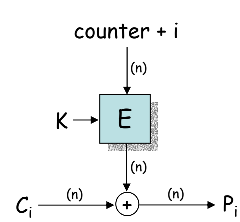
© Levente Buttyán

CTR mode

- encrypt



- decrypt



© Levente Buttyán

Properties of CTR mode

- similar to OFB
- cycle length depends on the size of the counter (typically 2^n)
- the i-th block can be decrypted independently of the others
  - parallelizable (unlike OFB)
  - random access
- the values to be XORed with the plaintext can be pre-computed
- at least as secure as the other modes

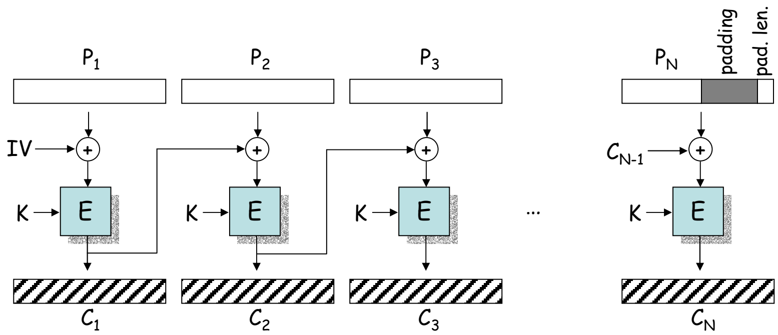
note1: in CFB, OFB, and CTR mode only the encryption algorithm is used (decryption is not needed)
- that is why Rijndael is optimized for encryption
- these modes shouldn't be used with public-key encryption algs.

note2: the OFB and CTR modes essentially make a synchronous stream cipher out of a block cipher, whereas the CFB mode converts a block cipher into a self-synchronizing stream-cipher

© Levente Buttyán

A security flaw induced by CBC padding

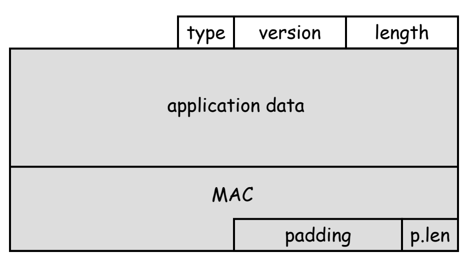
CBC encryption with padding



- TLS padding
  - last byte is the length n of the padding (not including the last byte)
  - all padding bytes have the value n
  - examples for correct padding: 0, 11, 222, 3333, …
- verification of TLS padding:
  - if the last byte is n, then verify if the last n+1 bytes are all n

Side channel

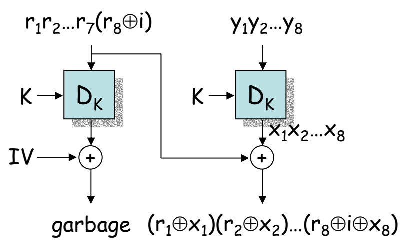
- TLS record message format



- send a random message to a TLS server
- the server will drop the message with overwhelming probability
  - either the padding is incorrect (the server responds with a DECRYPTION_FAILED alert)
  - or the MAC is incorrect with very high probability (the server responds with BAD_RECORD_MAC alert)
- if the response is BAD_RECORD_MAC, then the padding was correct
  → we get 1 bit of information !

Last byte(s) oracle

- assume we have an encrypted block $y_1y_2 \ldots y_8 = E_K(x_1x_2 \ldots x_8)$
- we want to compute x_8 (the last byte of x)
- idea:
  1. choose a random block $r_1r_2 \ldots r_8$; let i = 0
  2. send $r_1r_2 \ldots r_7(r_8 \oplus i)y_1y_2 \ldots y_8$ to the server (oracle)
  3. if there's a padding error, then increment i and go back to step 2
  4. if there's no padding error, then $r \oplus x$ ends with 0 or 11 or 222 …
     - the most likely is that $(r_8 \oplus i) \oplus x_8 = 0$, and hence $x_8 = r_8 \oplus i$



A security flaw induced by CBC padding

Last byte(s) oracle

- assume we get that $x \oplus r$ has a correct padding, but we don't know if it is 0 or 11 or 222 …
- algorithm:
  1. let j = 1
  2. change r_j and send $r_1r_2 \ldots r_8y_1y_2 \ldots y_8$ to the server again
  3. if the padding is still correct then the j-th byte was not a padding byte; increment j and go back to step 2
  4. if the padding becomes incorrect then the j-th byte was the first padding byte; $x_j \oplus r_j | x_{j+1} \oplus r_{j+1} | \ldots | x_8 \oplus r_8 = (8-j) | \ldots | (8-j)$ and hence $x_j\, x_{j+1} \ldots x_8 = r_j \oplus (8-j)\; r_{j+1} \oplus (8-j) \ldots r_8 \oplus (8-j)$

```
x = DE AD BE EF DE AD BE EF
r = 01 23 45 67 DD AE BD EC
r⊕x = DF 8E FB 88 03 03 03 03
```

j	r	r⊕x	padding
1	00 23 45 67 DD AE BD EC	DE 8E FB 88 03 03 03 03	OK
2	00 22 45 67 DD AE BD EC	DE 8F FB 88 03 03 03 03	OK
3	00 22 44 67 DD AE BD EC	DE 8F FA 88 03 03 03 03	OK
4	00 22 44 66 DD AE BD EC	DE 8F FA 89 03 03 03 03	OK
5	00 22 44 66 DC AE BD EC	DE 8F FA 89 02 03 03 03	ERROR

$x_5\, x_6\, x_7\, x_8$ = DD⊕03 AE⊕03 BD⊕03 EC⊕03 = DE AD BE EF

Block decryption oracle

- assume we have an encrypted block $y_1 y_2 \ldots y_8 = E_K(x_1 x_2 \ldots x_8)$ and we know the value of $x_j x_{j+1} \ldots x_8$ (using the last byte(s) oracle)
- we want to compute x_{j-1}
- algorithm:
  1. choose a random block $r_1 r_2 \ldots r_8$ such that $r_j = x_j \oplus (9-j)$; $r_{j+1} = x_{j+1} \oplus (9-j)$; … $r_8 = x_8 \oplus (9-j)$;
  2. let i = 0
  3. send $r_1 r_2 \ldots r_{j-2} (r_{j-1} \oplus i) r_j \ldots r_8 y_1 y_2 \ldots y_8$ to the server (oracle)
  4. if there's a padding error then increment i and go back to step 3
  5. if there's no padding error then $x_{j-1} \oplus r_{j-1} \oplus i = 9-j$ and hence $x_{j-1} = r_{j-1} \oplus i \oplus (9-j)$

x = DE AD BE EF DE AD BE EF

r = 01 23 45 67 DA A9 BA EB

r⊕x = DF 8E FB 88 04 04 04 04

i	r	r⊕x	padding
0	01 23 45 67 DA A9 BA EB	DF 8E FB 88 04 04 04 04	ERROR
1	01 23 45 66 DA A9 BA EB	DF 8E FB 89 04 04 04 04	ERROR
…	…	…	…
140	01 23 45 EB DA A9 BA EB	DF 8E FB 04 04 04 04 04	OK

x_4 = EB⊕04 = EF

Decryption oracle

- assume we have a CBC encrypted message $(C_1, C_2, \ldots, C_N)$ where
  - $C_1 = E_K(P_1 \oplus IV)$
  - $C_i = E_K(P_i \oplus C_{i-1})$ (for 1 < i < N)
  - $C_N = E_K([P_N|pad|plen] \oplus C_{N-1})$
- we want to compute $P_1, P_2, \ldots P_N$
- algorithm:
  - decrypt C_N using the block decryption oracle and XOR the result to C_{N-1}; you get $P_N|pad|plen$
  - decrypt C_i using the block decryption oracle and XOR the result to C_{i-1}; you get P_i
  - decrypt C_1 using the block decryption oracle and XOR the result to IV; you get P_1 (if IV is secret you cannot get P_1)

complexity of the whole attack:

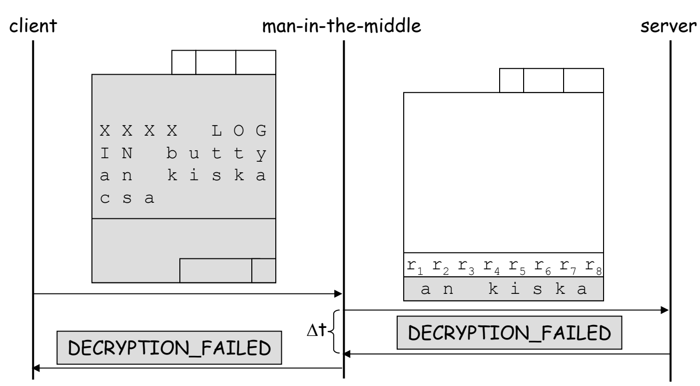
on average we need only $\frac{1}{2}$*256*8*N = 1024*N oracle calls !

Application

- vulnerable protocols: SSL, TLS, WTLS, IPsec, …
- some notes:
  - problems (and solutions) with TLS
    - alert messages are encrypted → BAD_RECORD_MAC and DECRYPTION_FAILED cannot be distinguished
      - measure timing between oracle call and oracle response
      - BAD_RECORD_MAC takes more time than DECRYPTION_FAILED
    - BAD_RECORD_MAC and DECRYPTION_FAILED are fatal errors → connection is closed after one oracle call
      - a password can still be broken if it is sent periodically to a server using TLS (a different session is used each time the password is sent)

Example: IMAP over TLS

- Outlook Express checks for new mail on the server periodically (every 5 minutes)
- each time the same password is sent for every folder
  XXXX LOGIN "username" "password"<0D><0A>
- it is possible to uncover the password using the attack as follows:



Fixes

- randomize response time after an error occurred (measuring timing of alert messages won't work)
- use random padding bytes
- put the padding before the MAC !

© Levente Buttyán