

NIS/CpE 691A

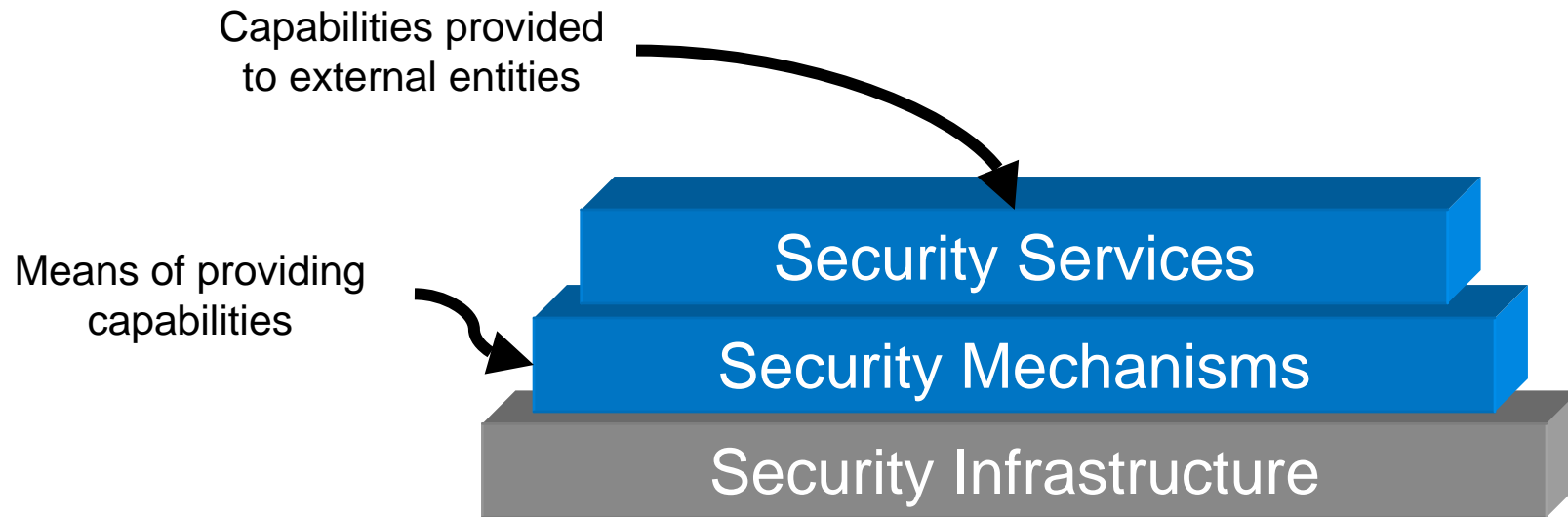
Information System Security

Class 2 – 1/25/06

What Security Issues Can Be Addressed By Cryptography and Related Techniques?

- Cryptography is NOT the solution to all security problems, but
- It does provide an enabling technology for many issues.
- If intelligently applied (balanced against other issues and needs) it can be of substantial value
- It provides a good place to start discussing detailed security technologies in an Information System

From Last Time: One Structured Way of Viewing Security



Categories of Security Mechanisms And Those That Can Be Addressed By Cryptography

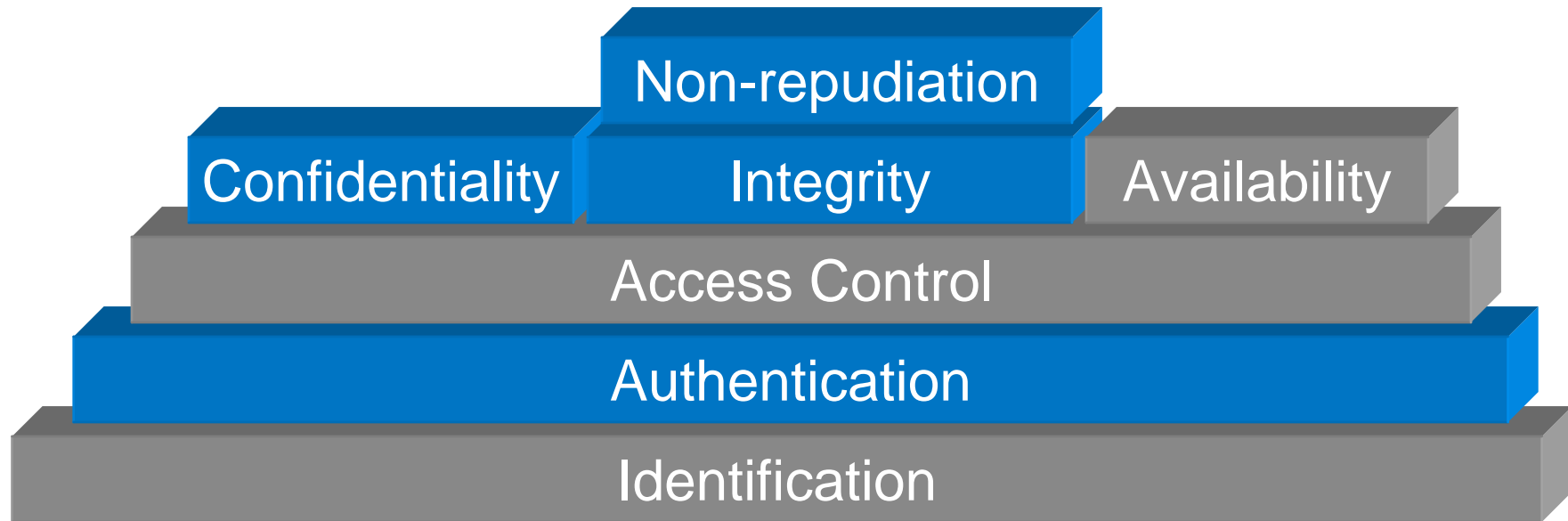
- *Prevent* occurrence of compromise
- *Detect* occurrence compromise
- *Correct* after-effects of compromise

Some Security Mechanisms and the Security Services They Could Enable

Mechanisms: \ Service:	ID	Auth	AC	Conf	Integ	Avail	NR
Cryptography/Encryption		✓		✓	✓		✓
Quality of Service Controls						✓	
Audit Logs			✓ *	✓ *	✓ *	✓ *	✓
Trusted Software			✓	✓	?	?	?
Security Policies	✓	✓	✓	✓	✓	✓	✓
Biometrics	✓	✓					
Smart Cards	✓	✓	✓	✓	✓		✓
System Backups					✓		✓
Security Assessment	✓	✓	✓	✓	✓	✓	✓

List of mechanisms is not meant to be exhaustive

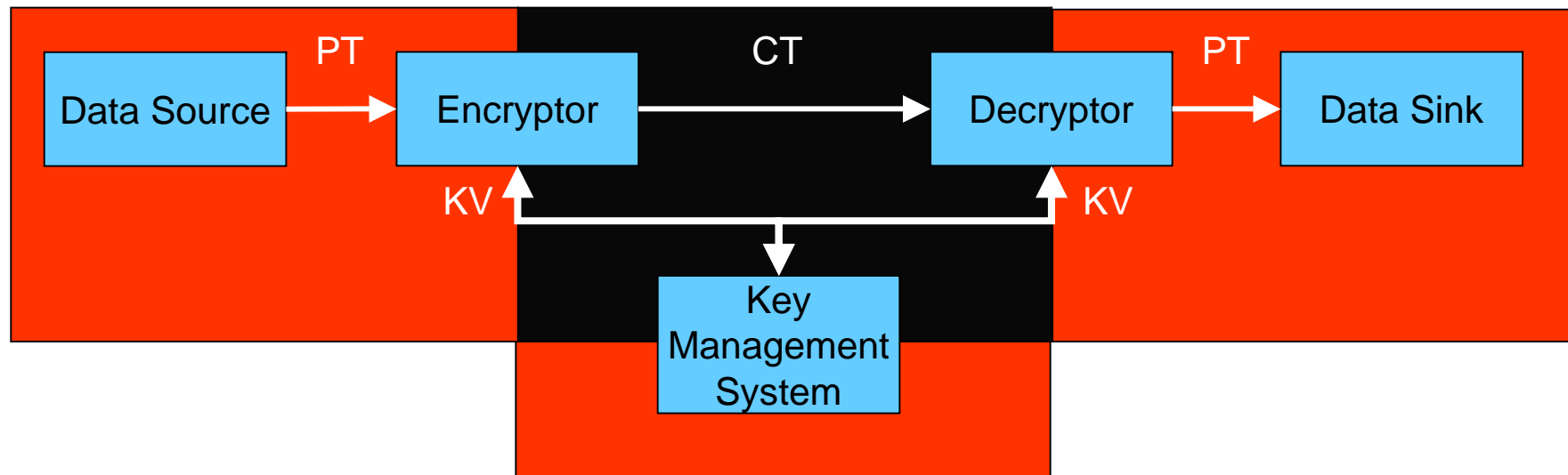
One Structured Way of Viewing Security And Security Services Addressed By Cryptograpy



Security Services

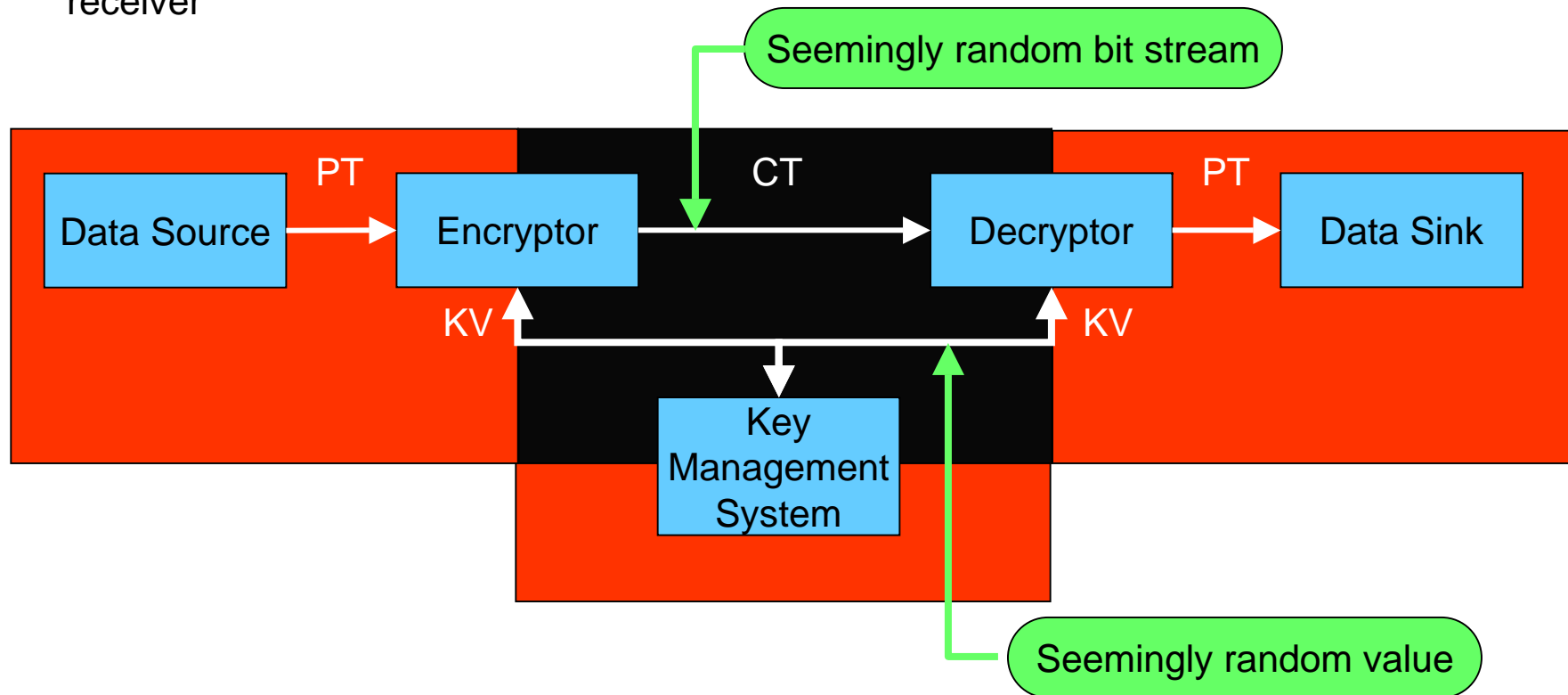
Cryptography Terminology

- Plaintext (PT) – unprotected source material (images, text, data, etc.)
- Ciphertext (CT) – Plaintext that has been enciphered (encrypted)
- Key Variable (KV) – Parameter of cryptographic system that selects, specifies, or controls key stream
- Key Management – Process for providing corresponding key variable(s) to sender and receiver



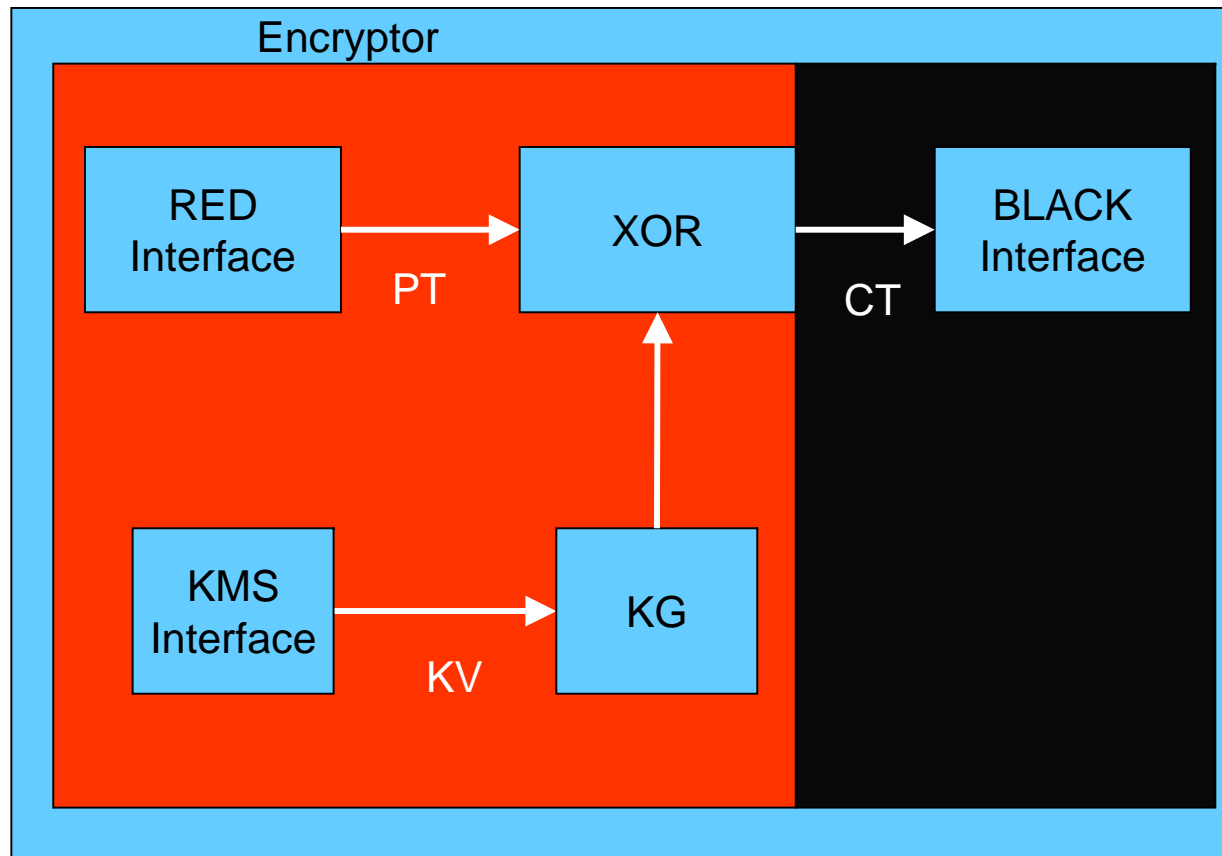
Cryptography Terminology

- Plaintext (PT) – unprotected source material (images, text, data, etc.)
- Ciphertext (CT) – Plaintext that has been enciphered (encrypted)
- Key Variable (KV) – Parameter of cryptographic system that selects, specifies, or controls key stream
- Key Management – Process for providing corresponding key variable(s) to sender and receiver



Cryptography Terminology - Continued

- Key Stream (Key Sequence) (KS) – (Pseudo)random string of symbols used to encrypt and/or decrypt plaintext
- Key Generator (KG) – Device that generates the key stream for a stream encipherment device



Miscellaneous Cryptography Terminology

- Affine:
 $F(x) = \alpha x + \beta$
- Linear:
 $F(x) = \gamma x$
 $F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$ [superposition]
- Nonlinear:
Superposition does not apply
- Permutation:
Reordering of inputs, e.g., $P(\{a,b,c,d\}) = \{c,b,a,d\}$
- Substitution:
Functional mapping, non necessarily 1-1 or onto

Information Theory 101

- In the 1970s certain large telecommunications research company used to give performance feedback in the form:
“Your performance was [above | below] average”
- Making certain assumptions about the distribution of performance (e.g., the mean was equal to the median), information content of this feedback was 1 bit.

Information Theory 101

- In the 1970s certain large telecommunications research company used to give performance feedback in the form:
“Your performance was [above | below] average”
- Making certain assumptions about the distribution of performance (e.g., the mean was equal to the median), information content of this feedback was 1 bit.
- In a more enlightened time, with more detailed information given to employees, performance feedback was of the form:
“Your performance [failed to meet | partially met | fully met | exceeded | far exceeded] expectations
- How much information does this convey? More than 2 bits?

Information Theory 101

- In the 1970s certain large telecommunications research company used to give performance feedback in the form:
“Your performance was [above | below] average”
- Making certain assumptions about the distribution of performance (e.g., the mean was equal to the median), information content of this feedback was 1 bit.
- In a more enlightened time, with more detailed information given to employees, performance feedback was of the form:
“Your performance [failed to meet | partially met | fully met | exceeded | far exceeded] expectations
- How much information does this convey? More than 2 bits?

If $p(\text{FM})=.03$, $p(\text{PM})=.07$, $p(\text{FM})=.1$, $p(\text{E})=.75$, $p(\text{FE})=.05$, information content is 1.28 bits

Information Theory 101

- In the 1970s certain large telecommunications research company used to give performance feedback in the form:
“Your performance was [above | below] average”
- Making certain assumptions about the distribution of performance (e.g., the mean was equal to the median), information content of this feedback was 1 bit.
- In a more enlightened time, with more detailed information given to employees, performance feedback was of the form:
“Your performance [failed to meet | partially met | fully met | exceeded | far exceeded] expectations
- How much information does this convey? More than 2 bits?

If $p(\text{FM})=.03$, $p(\text{PM})=.07$, $p(\text{FM})=.1$, $p(\text{E})=.75$, $p(\text{FE})=.05$, information content is 1.28 bits
- Information content is not fully specified by the number of states, but also the probability of the states.

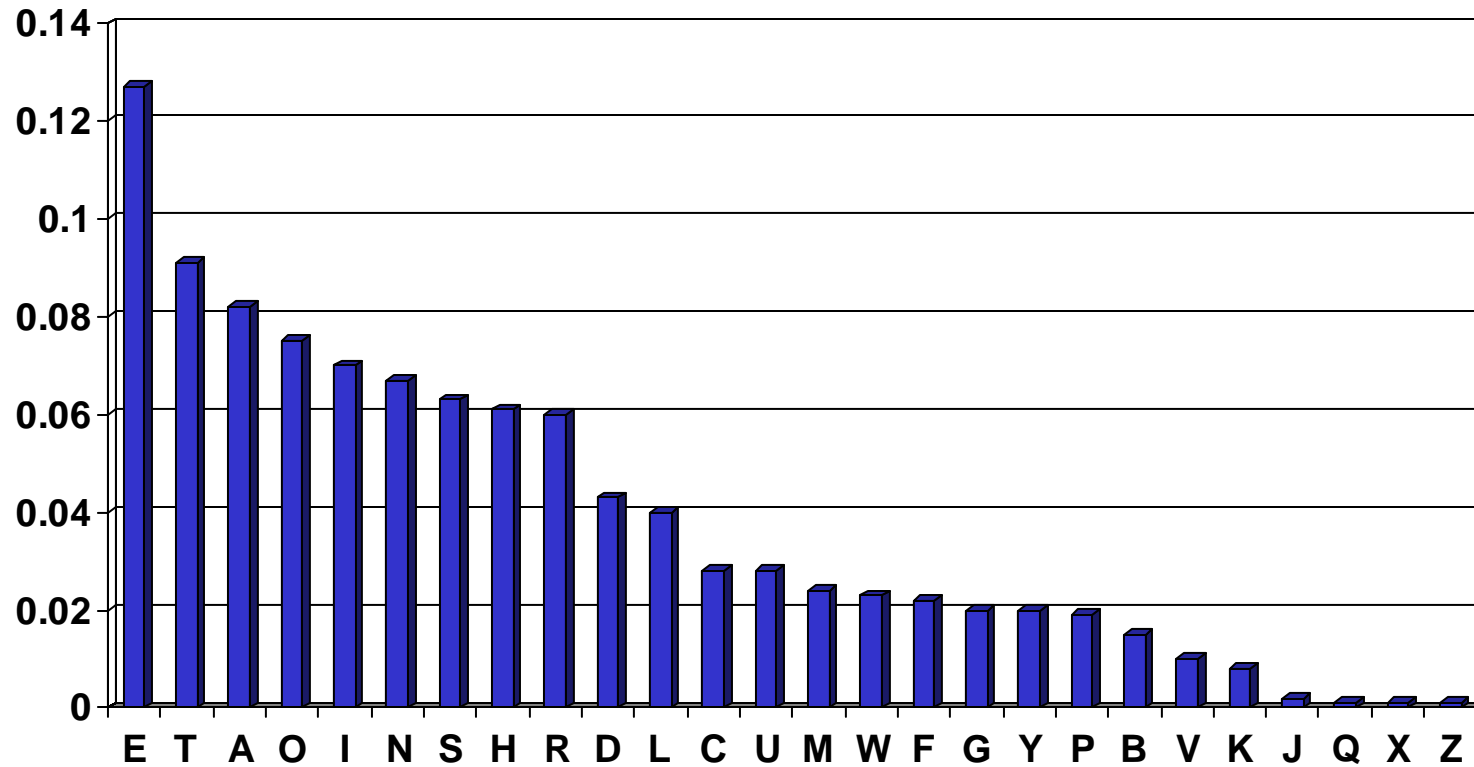
Information Theory 101

- In the 1970s certain large telecommunications research company used to give performance feedback in the form:
“Your performance was [above | below] average”
- Making certain assumptions about the distribution of performance (e.g., the mean was equal to the median), information content of this feedback was 1 bit.
- In a more enlightened time, with more detailed information given to employees, performance feedback was of the form:
“Your performance [failed to meet | partially met | fully met | exceeded | far exceeded] expectations
- How much information does this convey? More than 2 bits?

If $p(\text{FM})=.03$, $p(\text{PM})=.07$, $p(\text{FM})=.1$, $p(\text{E})=.75$, $p(\text{FE})=.05$, information content is 1.28 bits

- Information content is not fully specified by the number of states, but also the probability of the states.
- The maximum information content is realized when all symbols are equally likely

English Language Letter Statistics



$$\sum_{i \in \text{Alphabet}} p_i \cdot \log_2(p_i) = 4.18 \text{ bits}$$

Evolution of Cryptography

- Monoalphabetic substitution, e.g.,
 - Caesar cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{b,c,d,e,f,g,\dots,y,z,a\}$
 - Atbash cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{z,y,x,\dots,f,e,d,c,b,a\}$
 - Any permutation of the alphabet

Evolution of Cryptography

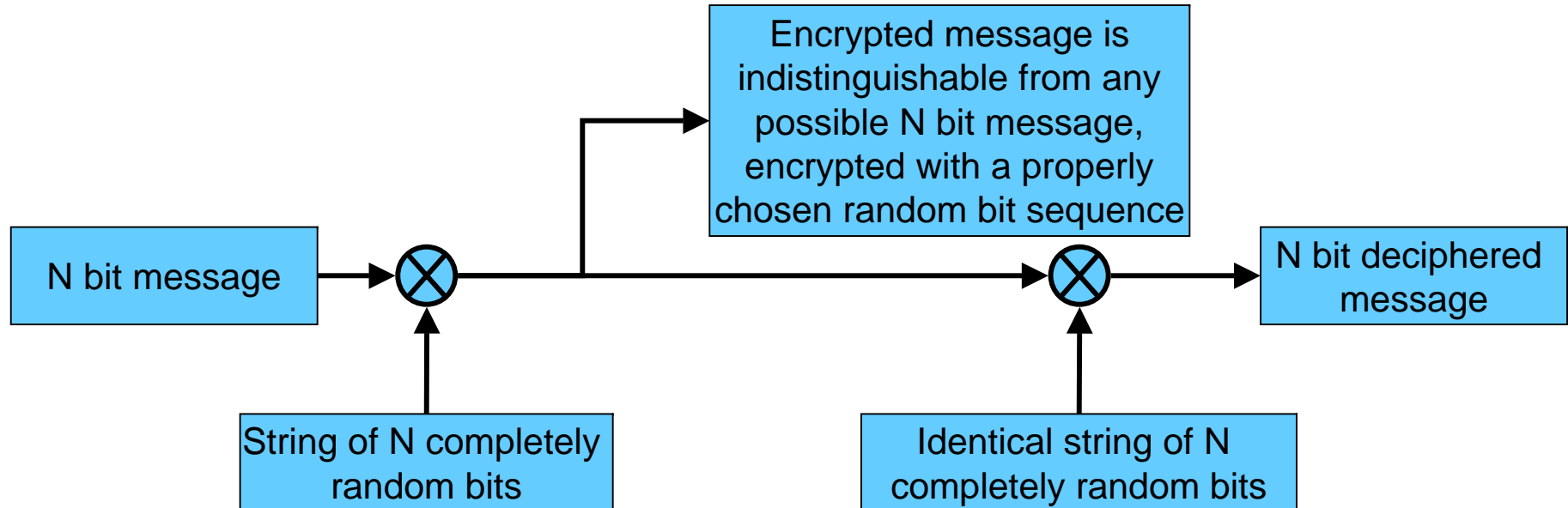
- Monoalphabetic substitution, e.g.,
 - Caesar cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{b,c,d,e,f,g,\dots,y,z,a\}$
 - Atbash cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{z,y,x,\dots,f,e,d,c,b,a\}$
 - Any permutation of the alphabet
 - Easily solved by observing single and double letter frequencies
 - English (like most other non-ideographic languages) have distinct letter frequencies over a small alphabet.
 - Encoding English letters requires $\log_2(26) \sim 4.7$ bits/letter,
 - but information content in English text is $\sum p \log_2(p) \sim 4.2$ bits/letter – each letter gives attacker .5 bits of advantage

Evolution of Cryptography

- Monoalphabetic substitution, e.g.,
 - Caesar cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{b,c,d,e,f,g,\dots,y,z,a\}$
 - Atbash cipher $\{a,b,c,d,e,f,\dots,x,y,z\} \rightarrow \{z,y,x,\dots,f,e,d,c,b,a\}$
 - Any permutation of the alphabet
 - Easily solved by observing single and double letter frequencies
 - English (like most other non-ideographic languages) have distinct letter frequencies over a small alphabet.
 - Encoding English letters requires $\log_2(26) \sim 4.7$ bits/letter,
 - but information content in English text is $\sum p \log_2(p) \sim 4.2$ bits/letter – each letter gives attacker .5 bits of advantage
- Polyalphabetic substitution:
 - `thisisamessagetobeencrypted` - plaintext
 - `badbadbadbadbadbadbadbadbad` - key stream
 - `vimujwcniuteifxqciogtztvfh` - ciphertext
 - Correlation-like techniques find the length of the key stream, k
 - Problem then reduces to solving k monoalphabetic ciphers
 - Using running text (e.g., from an agreed to book) makes solution harder, but with enough ciphertext, both the plaintext as well as the key stream are easily found

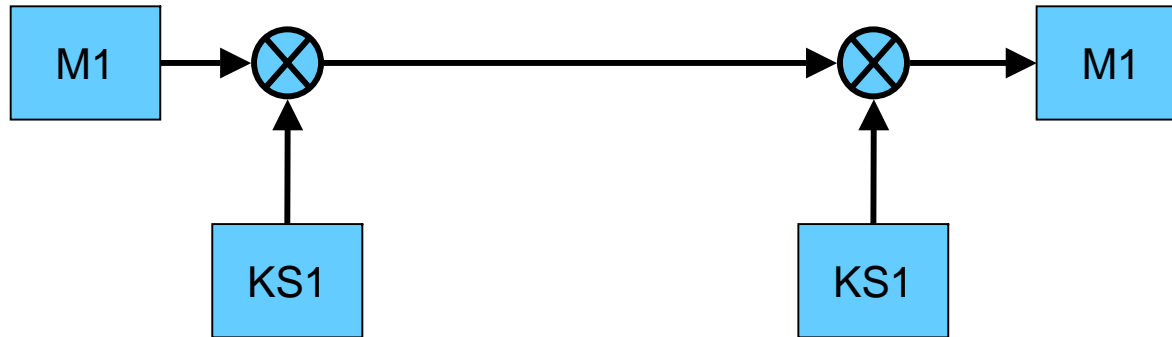
Evolution of Cryptography - 2

- Weakness of polyalphabetic cipher is repetition of the key stream
– What if it never repeated?



- One-time-pad is the only provably secure cryptographic system
What happens if key sequence is (accidentally) reused?

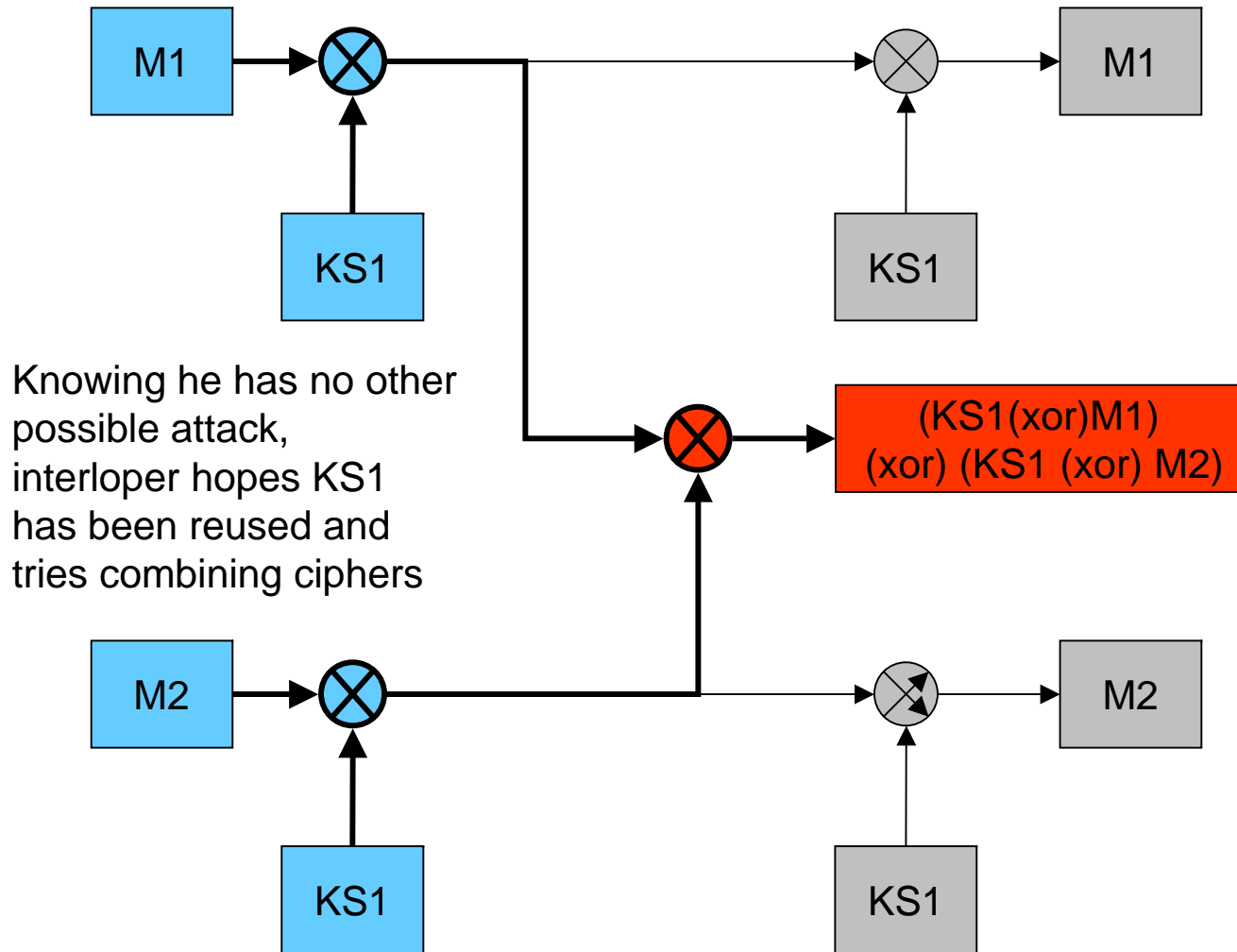
One-bit-pad Key Reuse



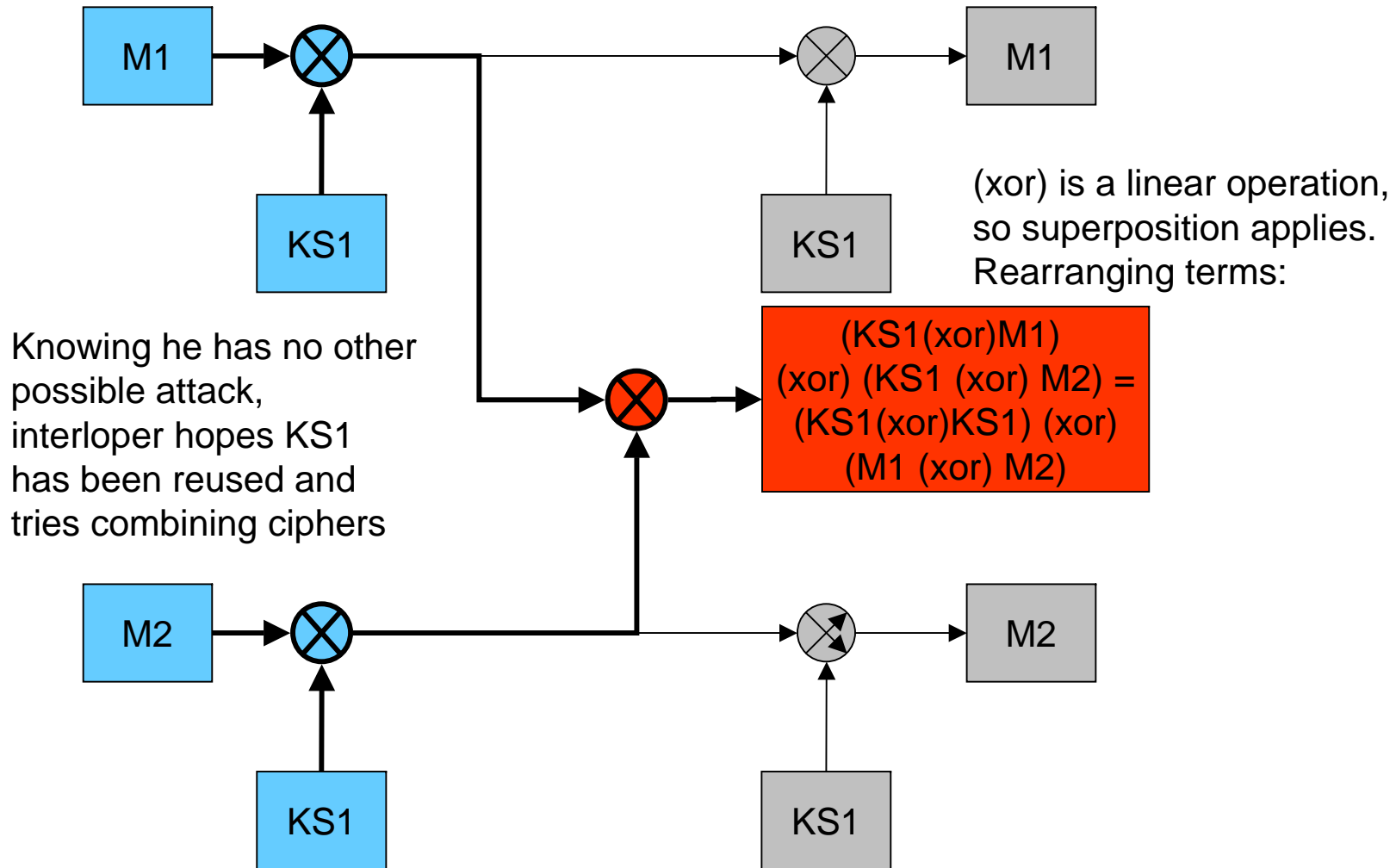
Sender (or receiver) accidentally sent M2, reusing KS1, previously used for M1



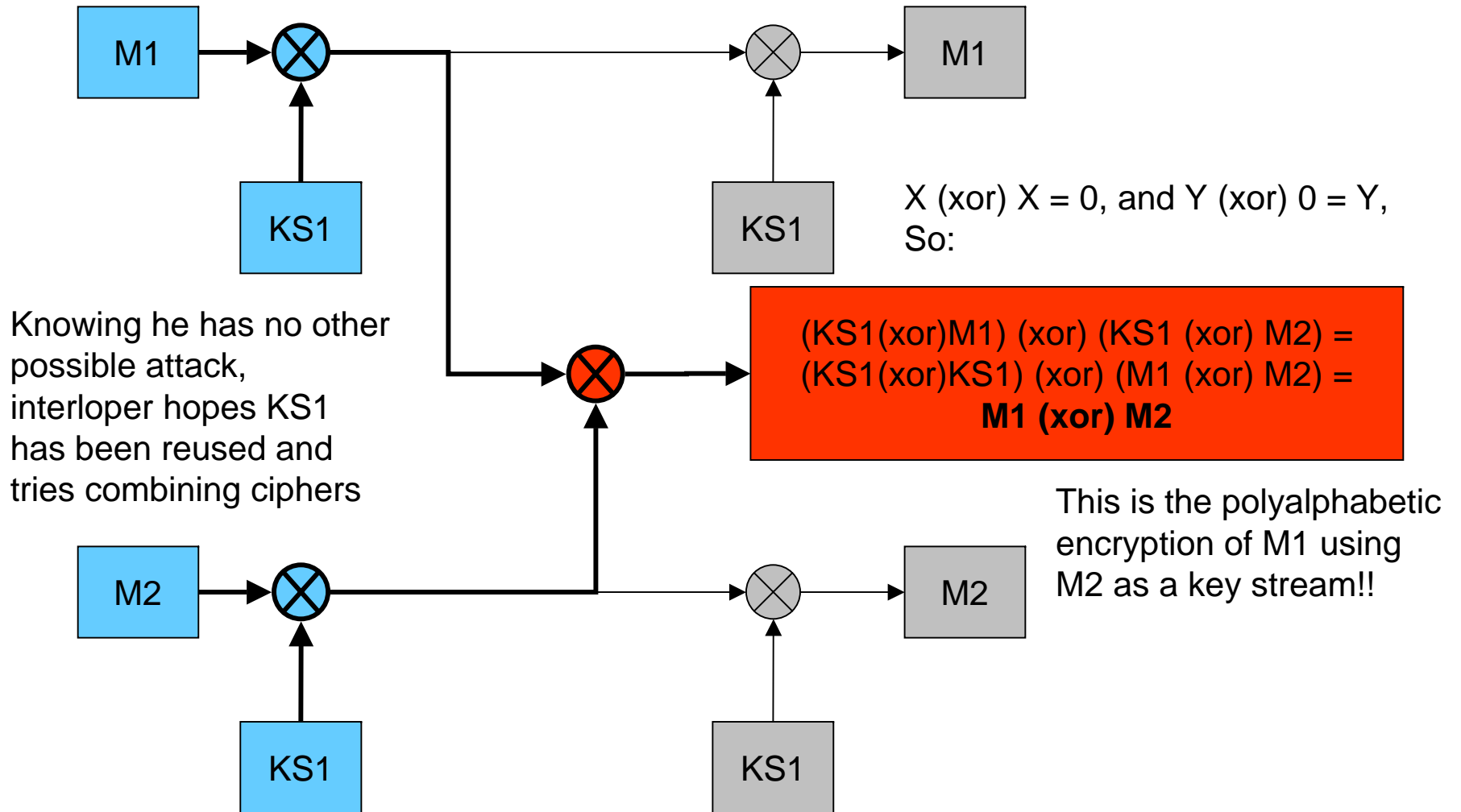
One-bit-pad Key Reuse



One-bit-pad Key Reuse

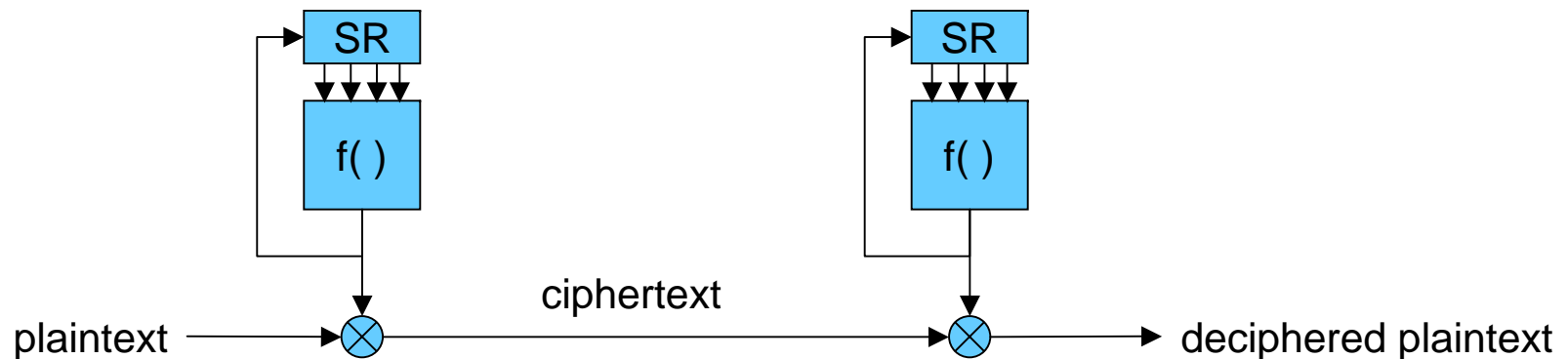
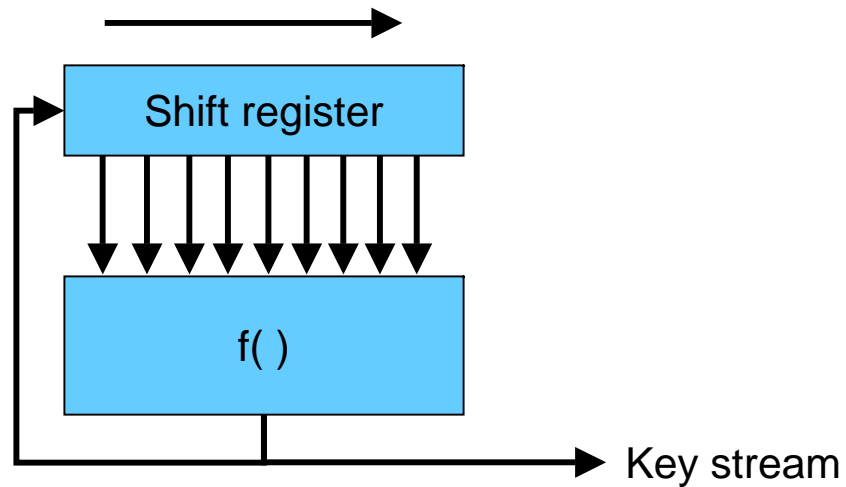


One-bit-pad Key Reuse



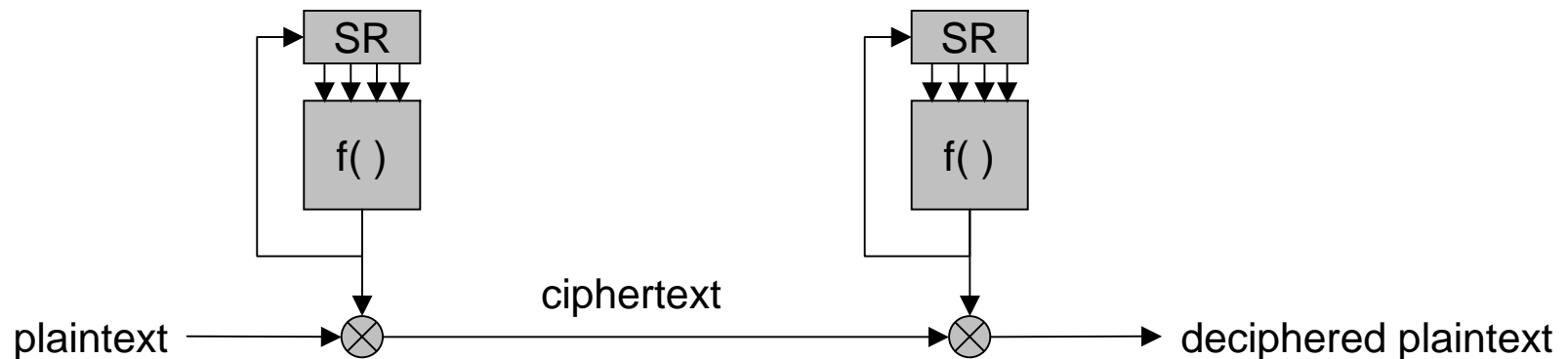
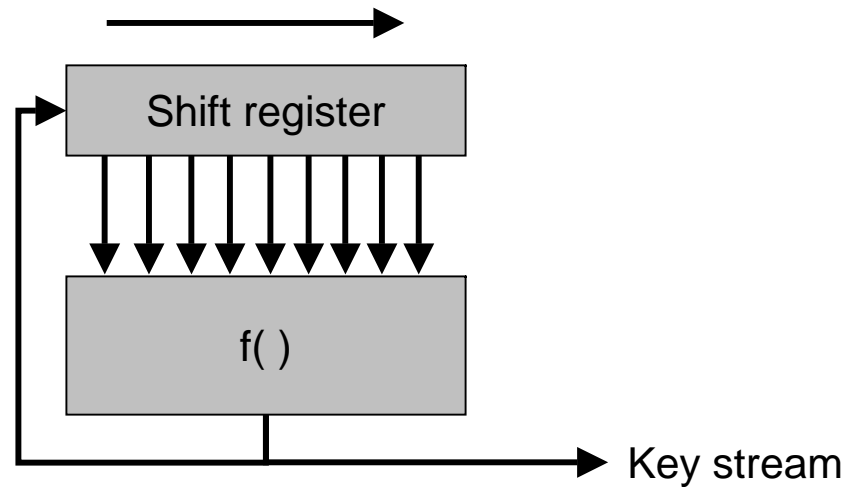
Generating Long Pseudo-Random Sequences

- For N bit register, if $f(x)$ is linear, $2N-1$ bits of key stream are sufficient to find $f()$
- So, $f()$ must be nonlinear



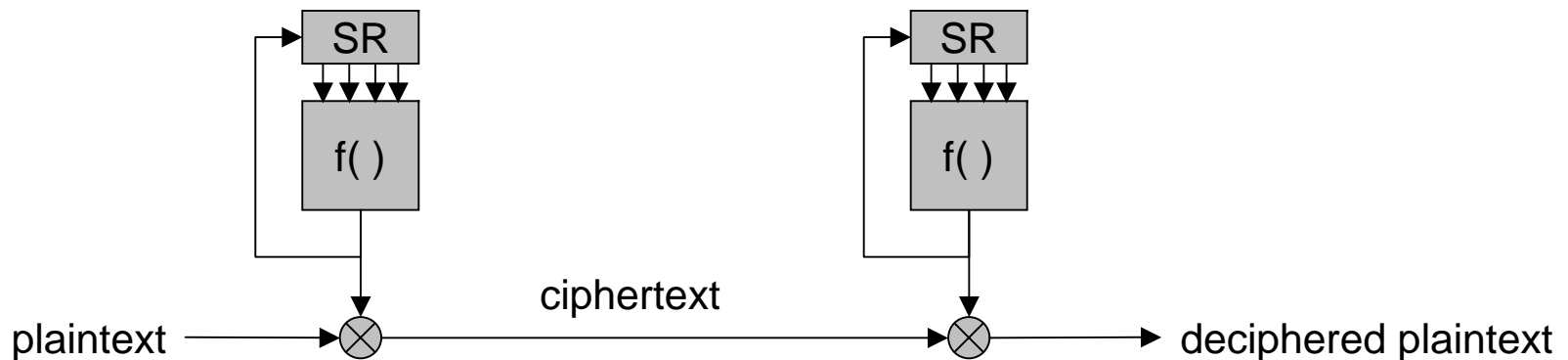
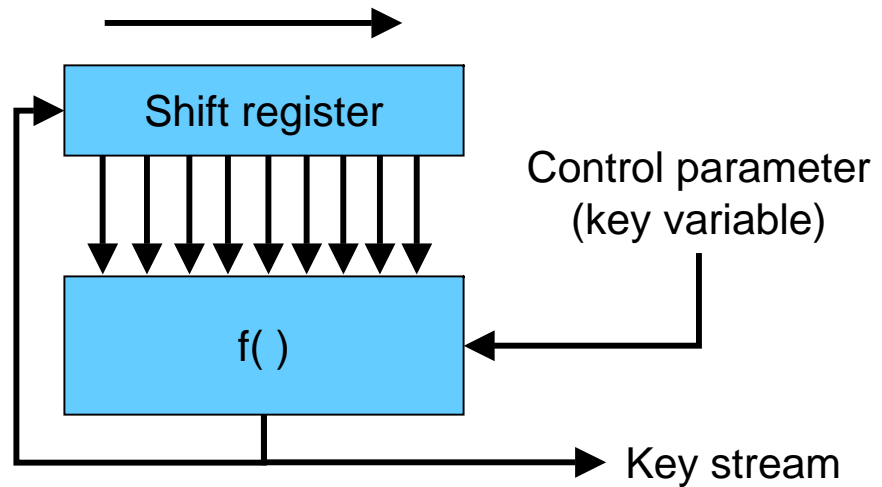
Generating Long Pseudo-Random Sequences

- How to make $f()$ easy to change?
- What about synchronization?



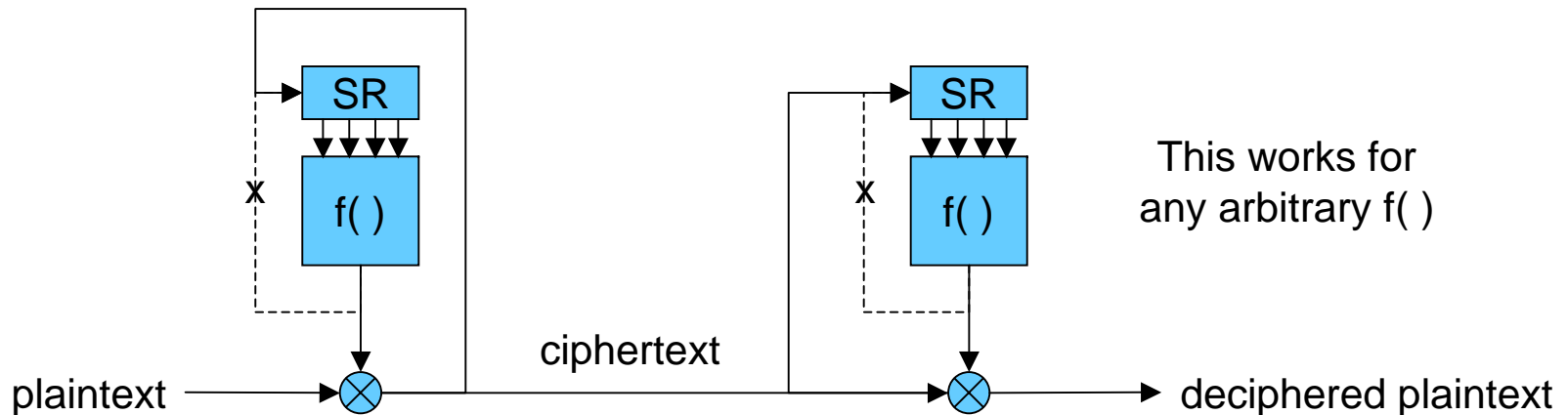
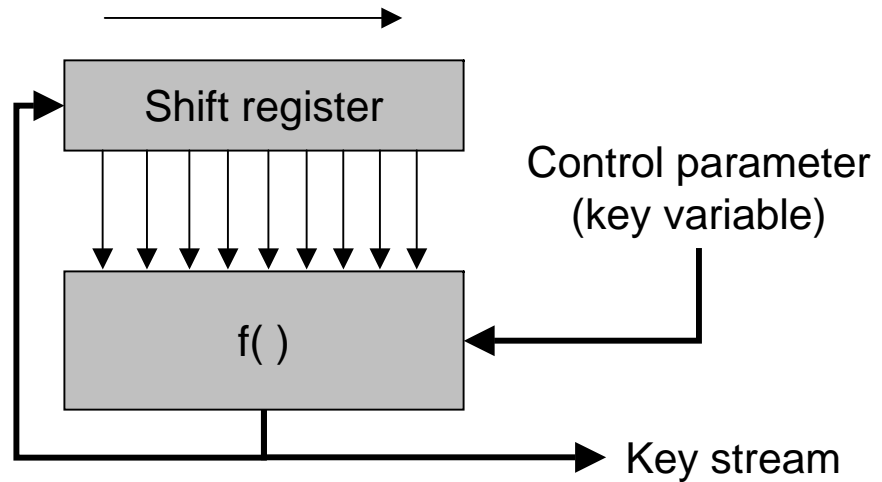
Generating Long Pseudo-Random Sequences

- How to make $f()$ easy to change?
- What about synchronization?

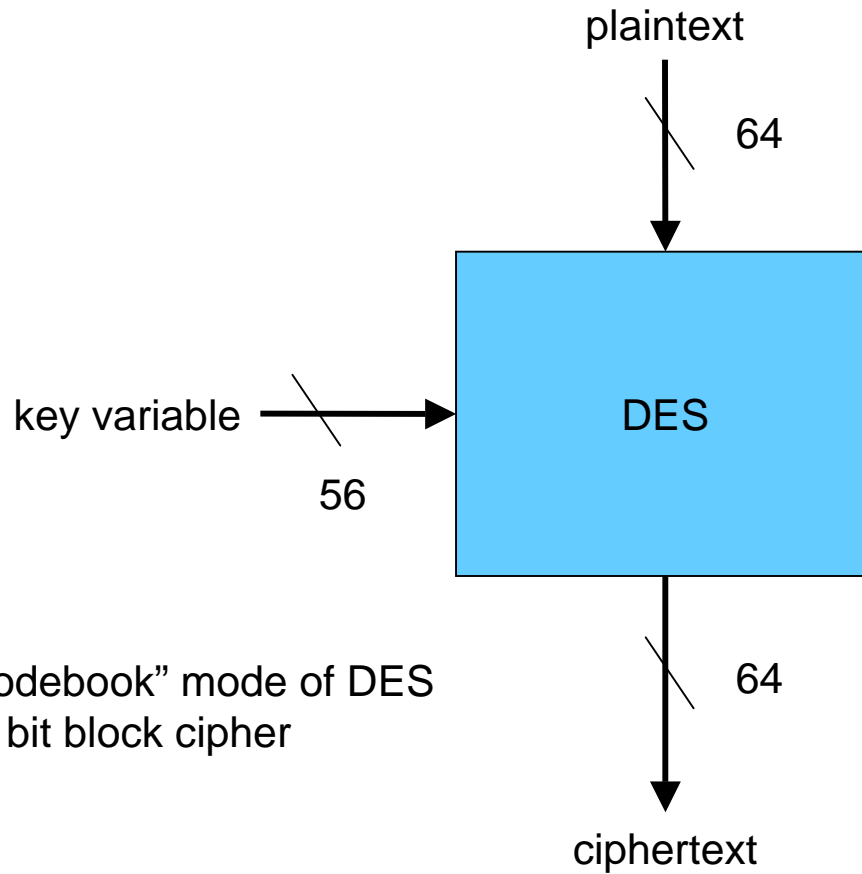


Generating Long Pseudo-Random Sequences

- How to make $f()$ easy to change?
- What about synchronization?

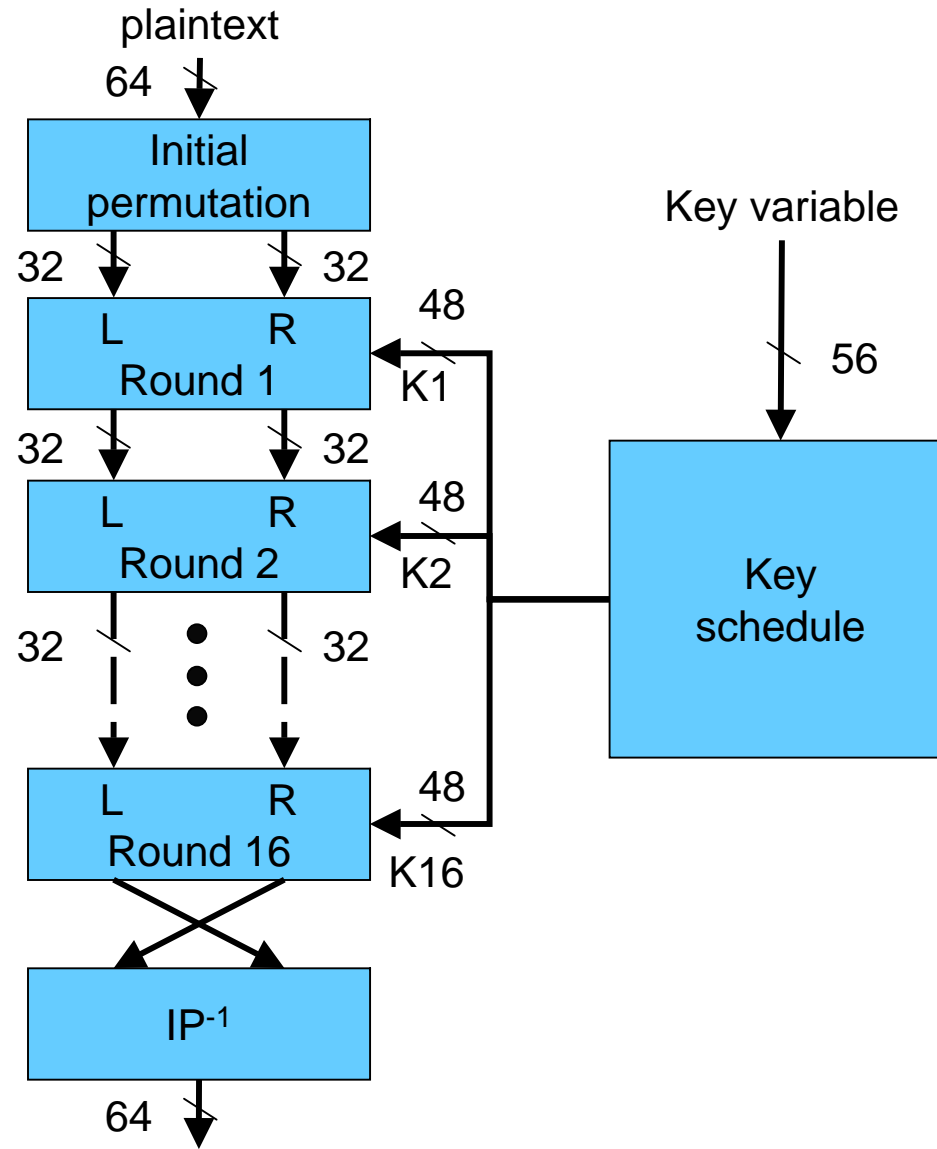


DES as one $f()$ option

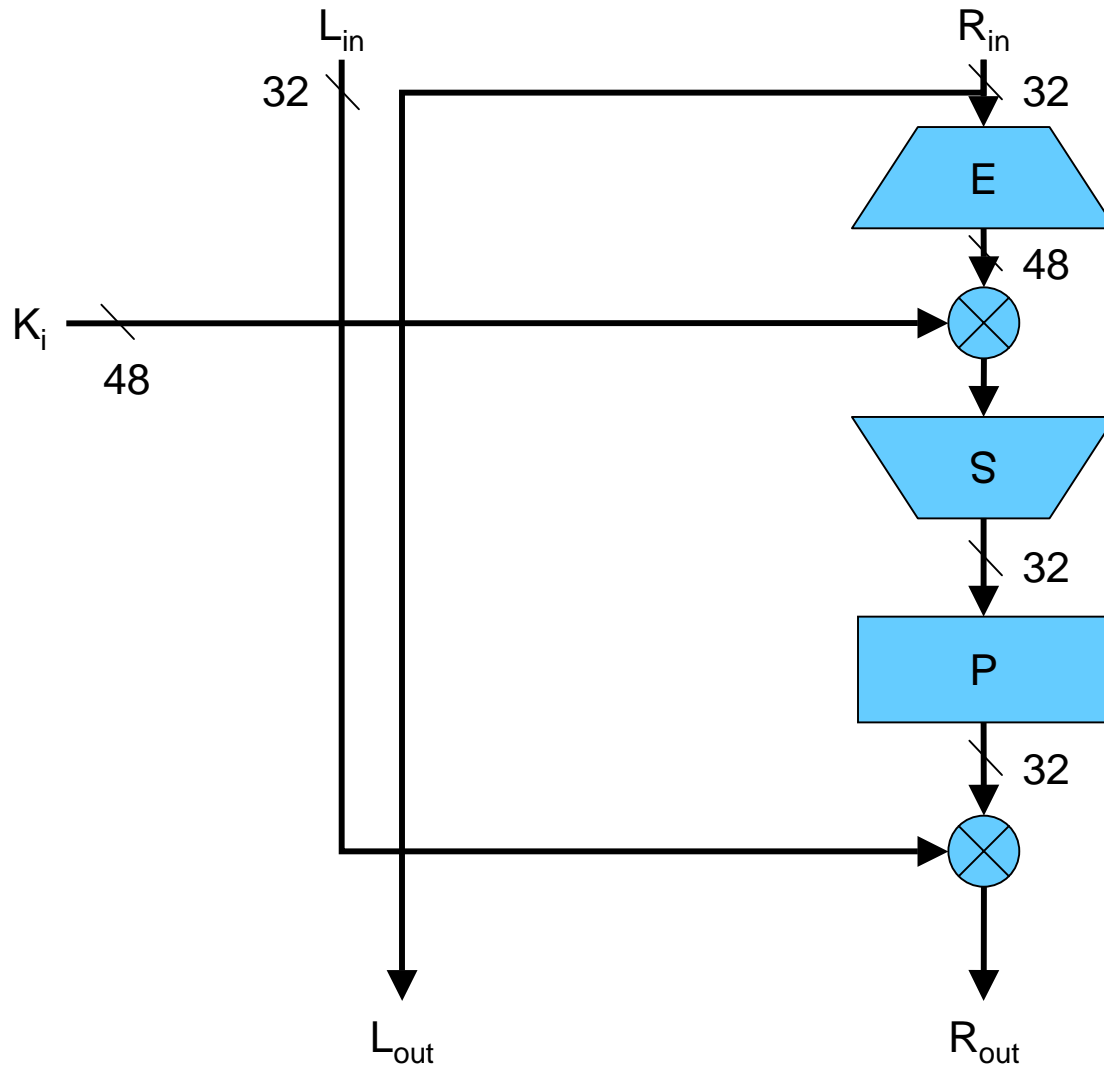


“Electronic codebook” mode of DES
– 64 bit block cipher

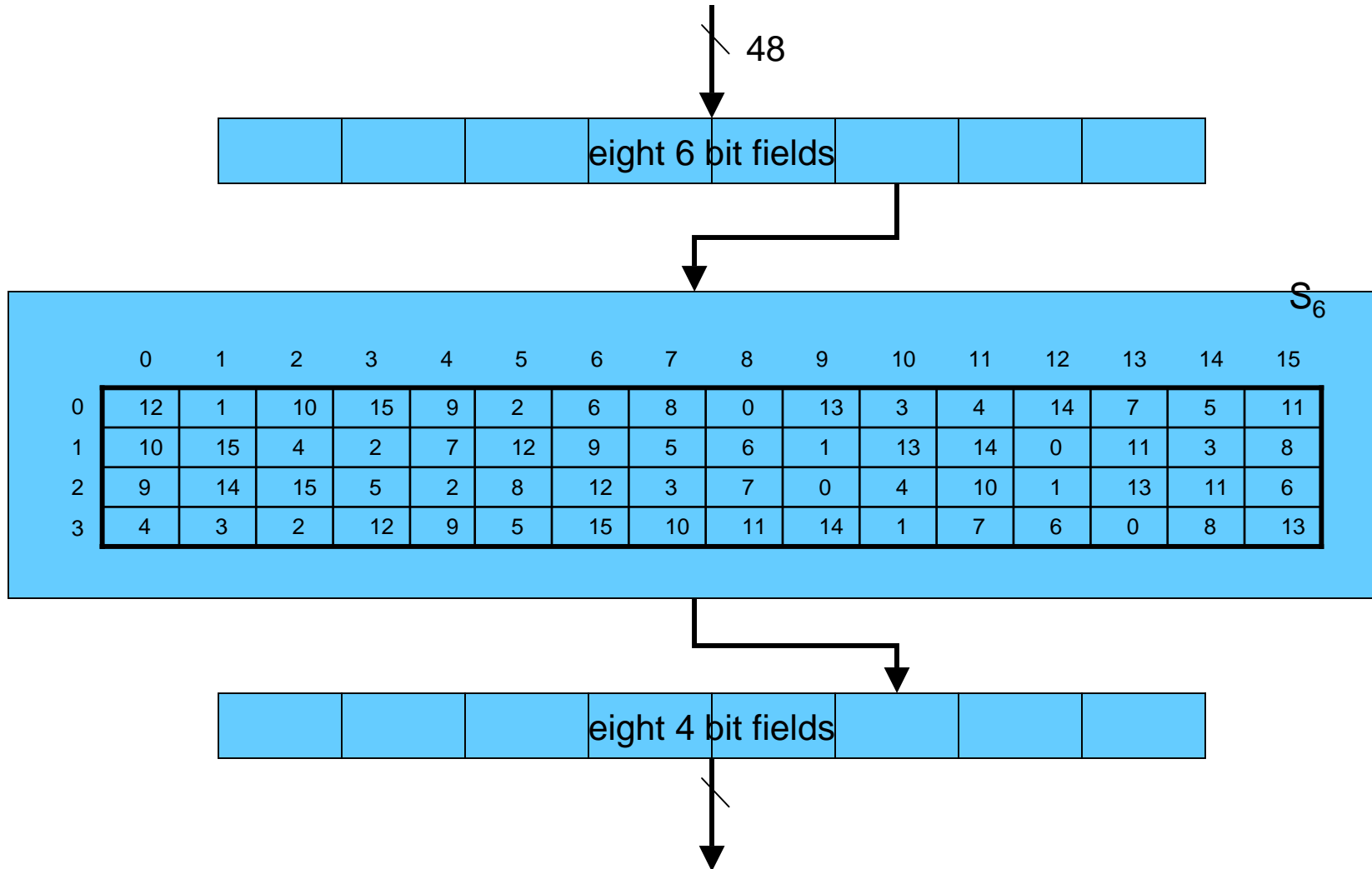
Internal operation of DES



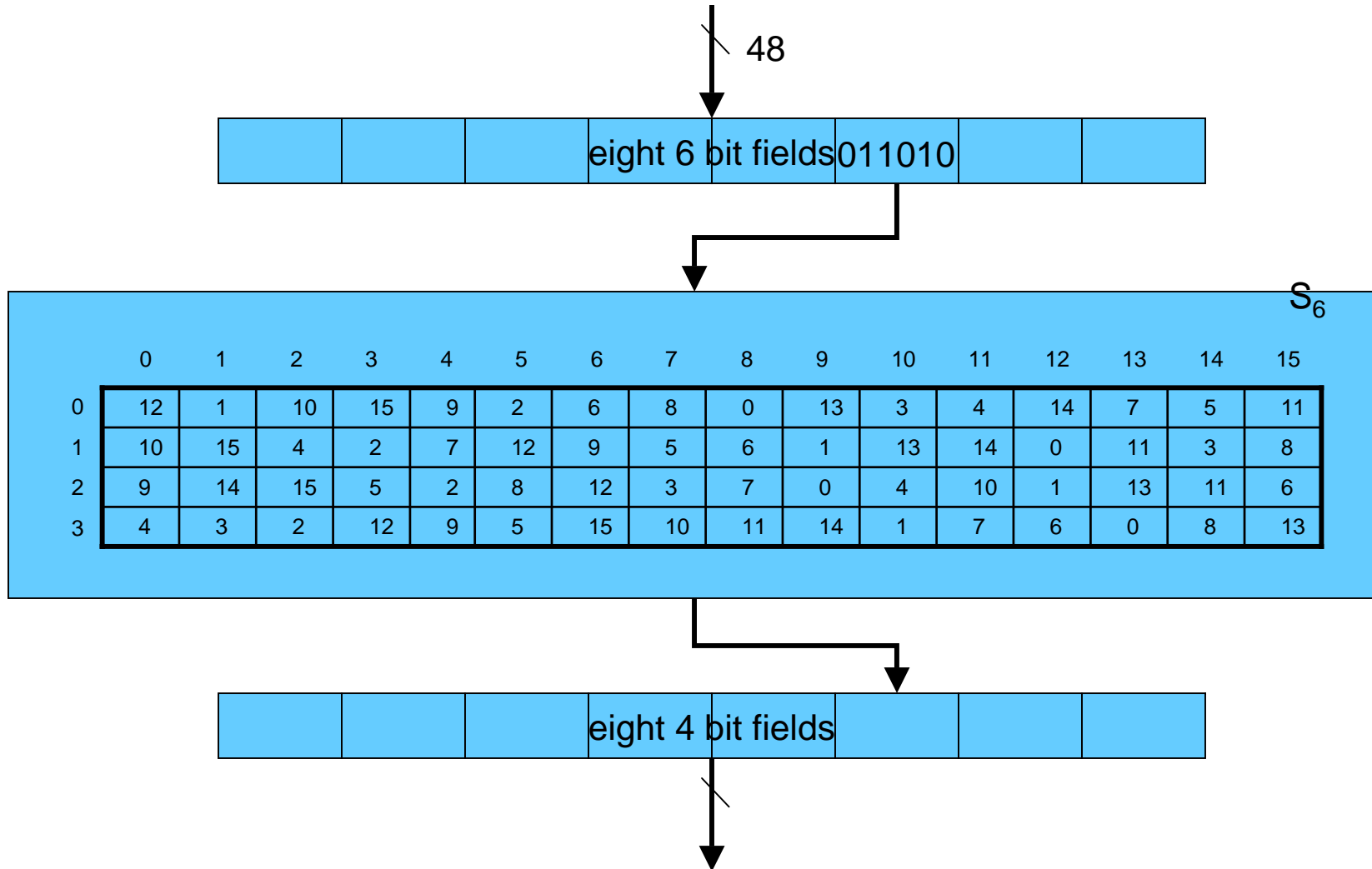
DES Round_i



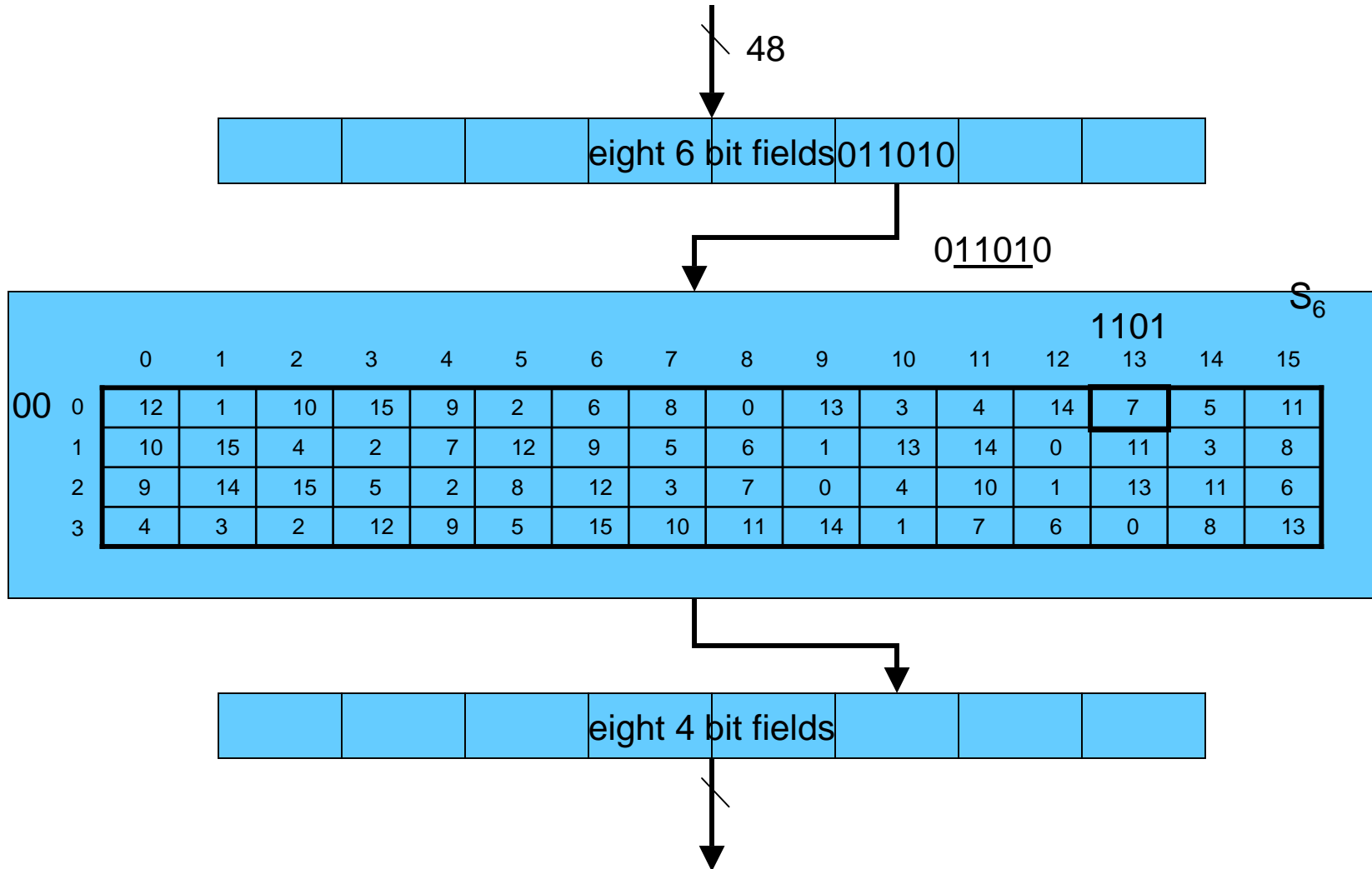
DES S-Boxes



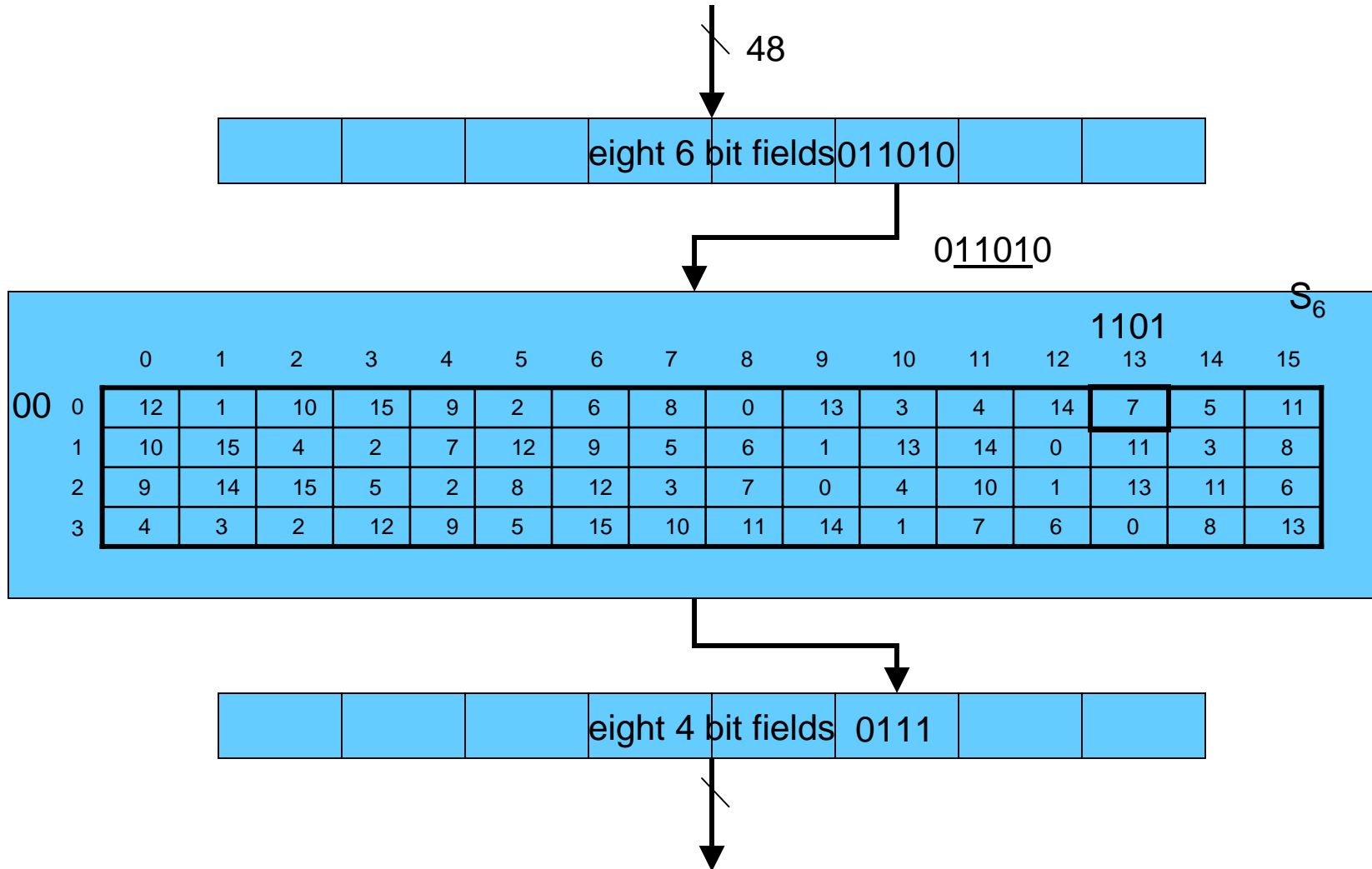
DES S-Boxes



DES S-Boxes



DES S-Boxes



DES S-Boxes

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
3	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13

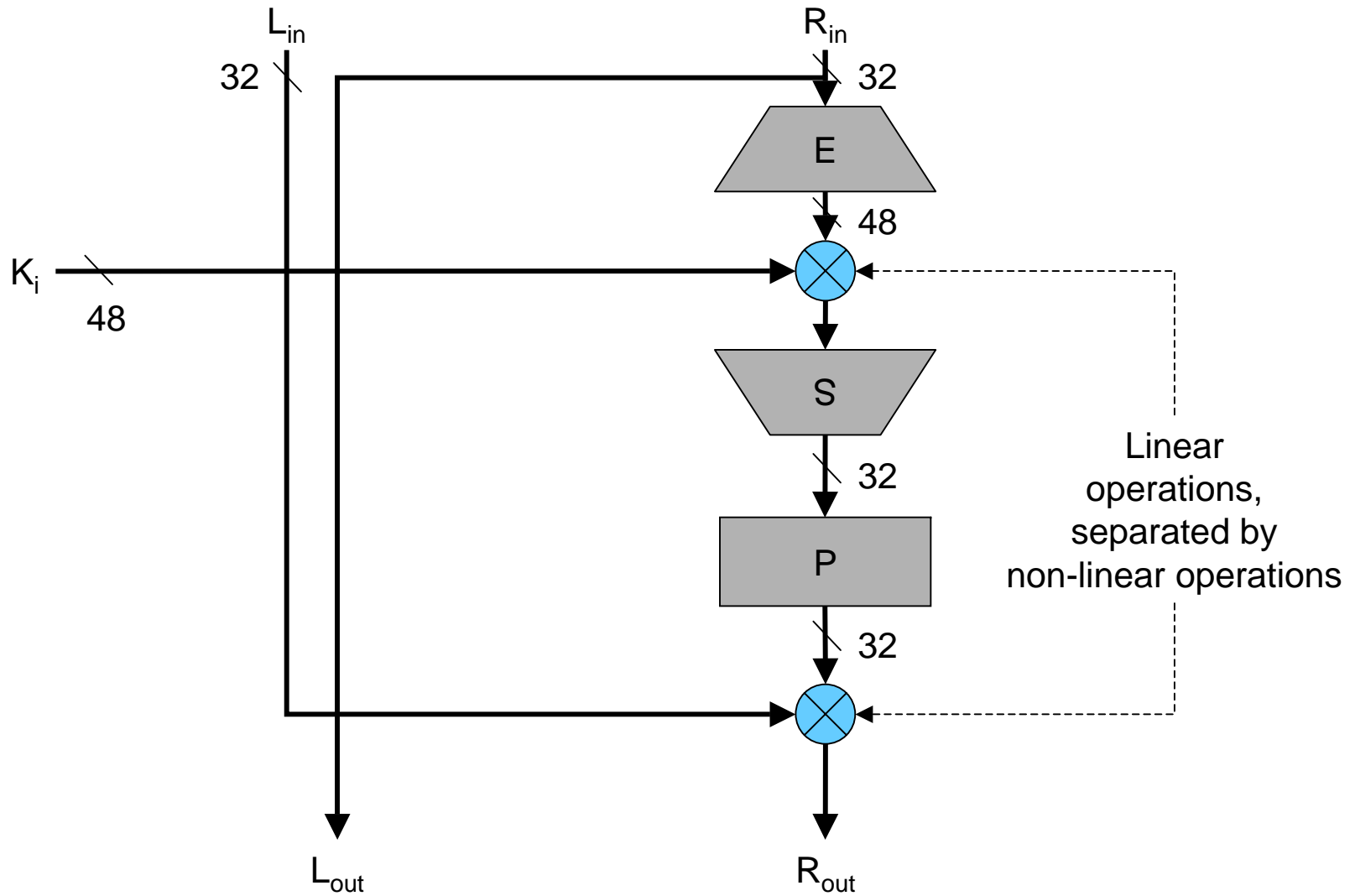
S_6

	0	1	3	2	6	7	5	4	12	13	15	14	10	11	9	8
0	1100	0001	1111	1010	0110	1000	0010	1001	1110	0111	1011	0101	0011	0100	1101	0000
1	1010	1111	0010	0100	1001	0101	1100	0111	0000	1011	1000	0011	1101	1110	0001	0110
3	0100	0011	1100	0010	1111	1010	0101	1001	0110	0000	1101	1000	0001	0111	1110	1011
2	1001	1110	0101	1111	1100	0011	1000	0010	0001	1101	0110	1011	0100	1010	0000	0111

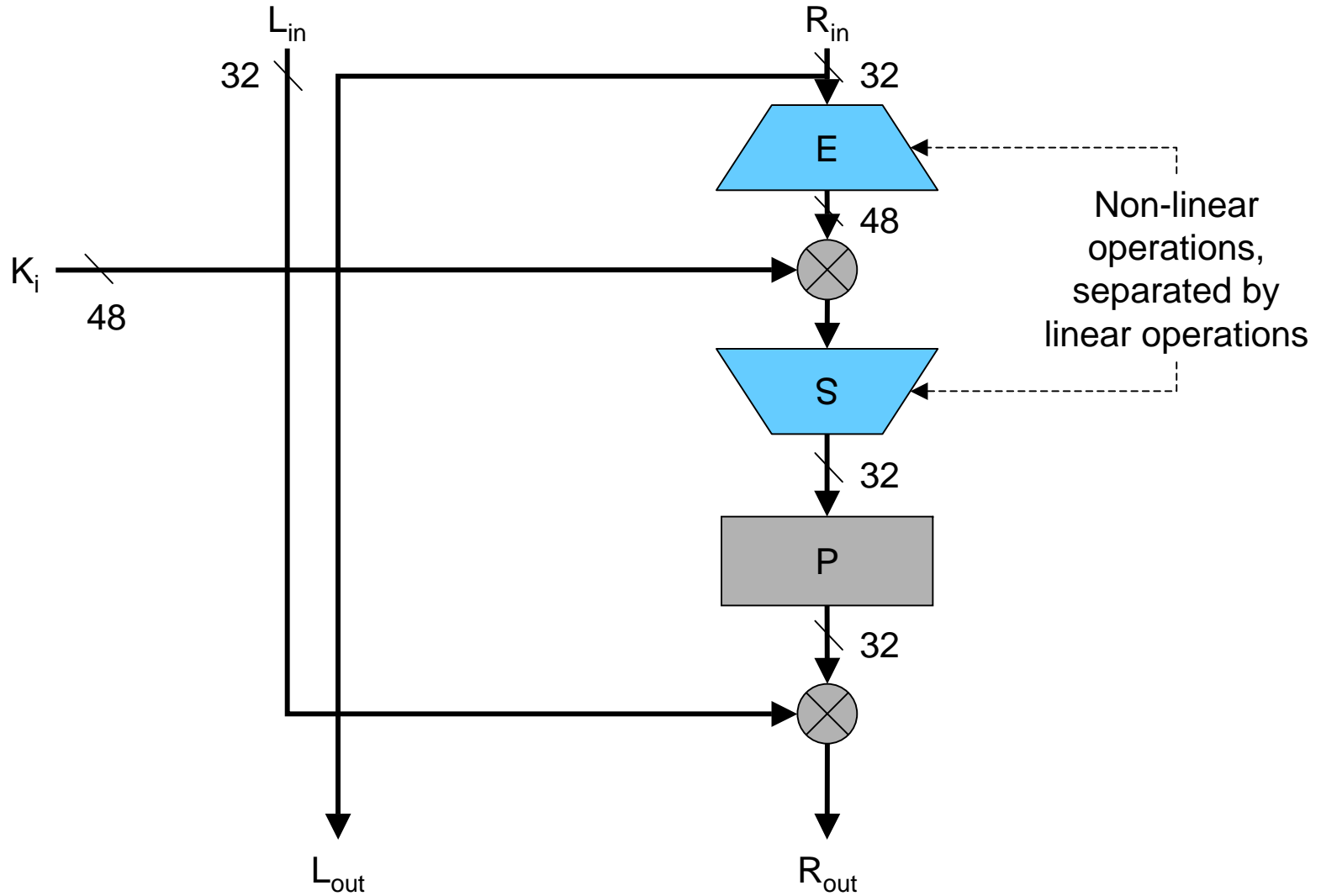
In the lower version, adjacent cells differ by one input bit

Outputs differ by 2-3 bits

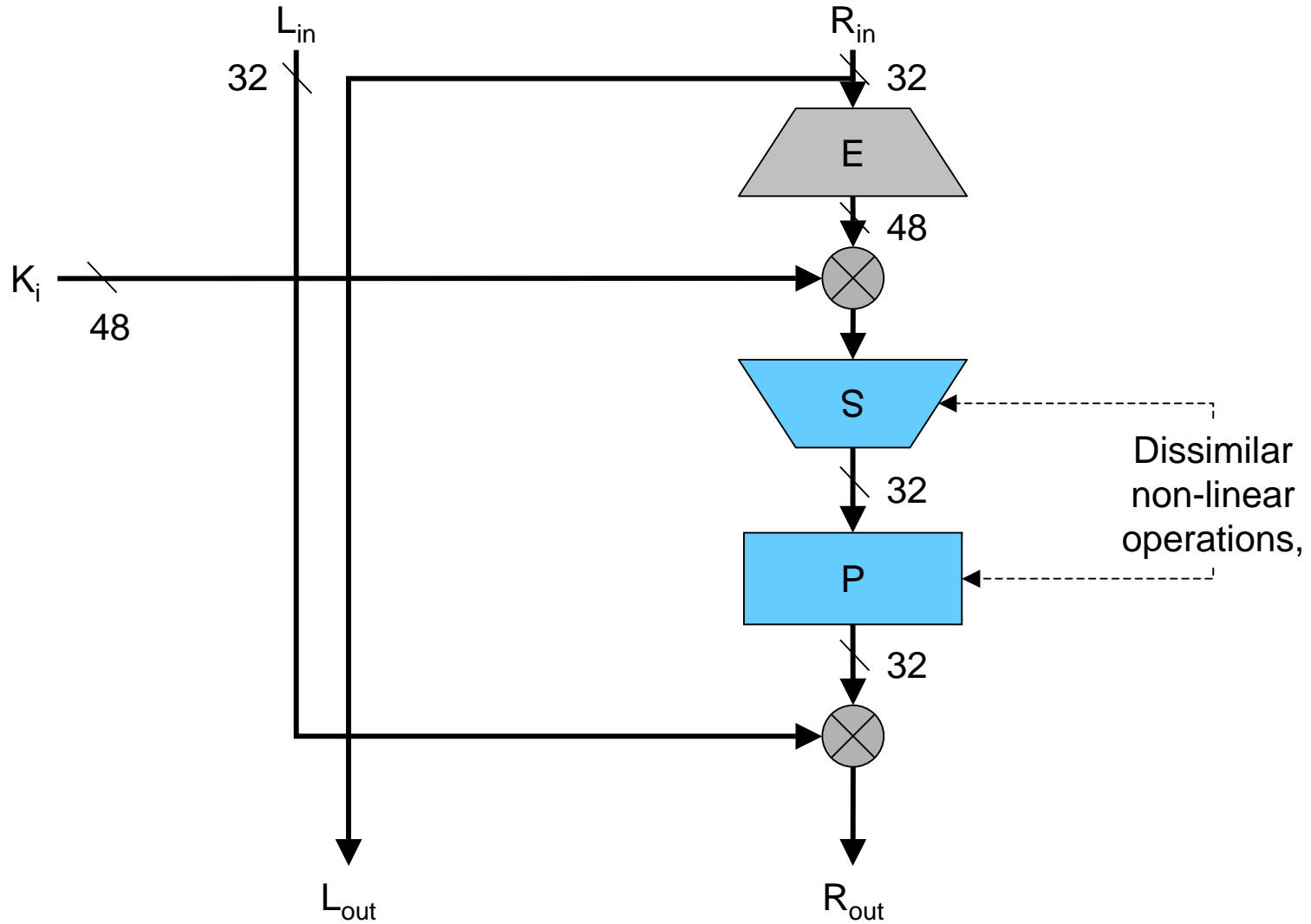
DES Round_i



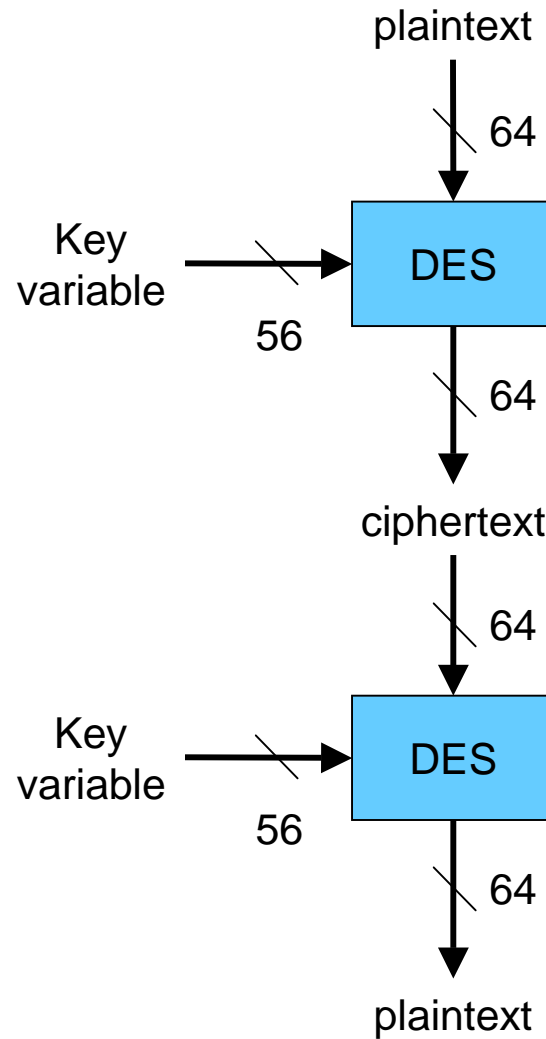
DES Round_i



DES Round_i

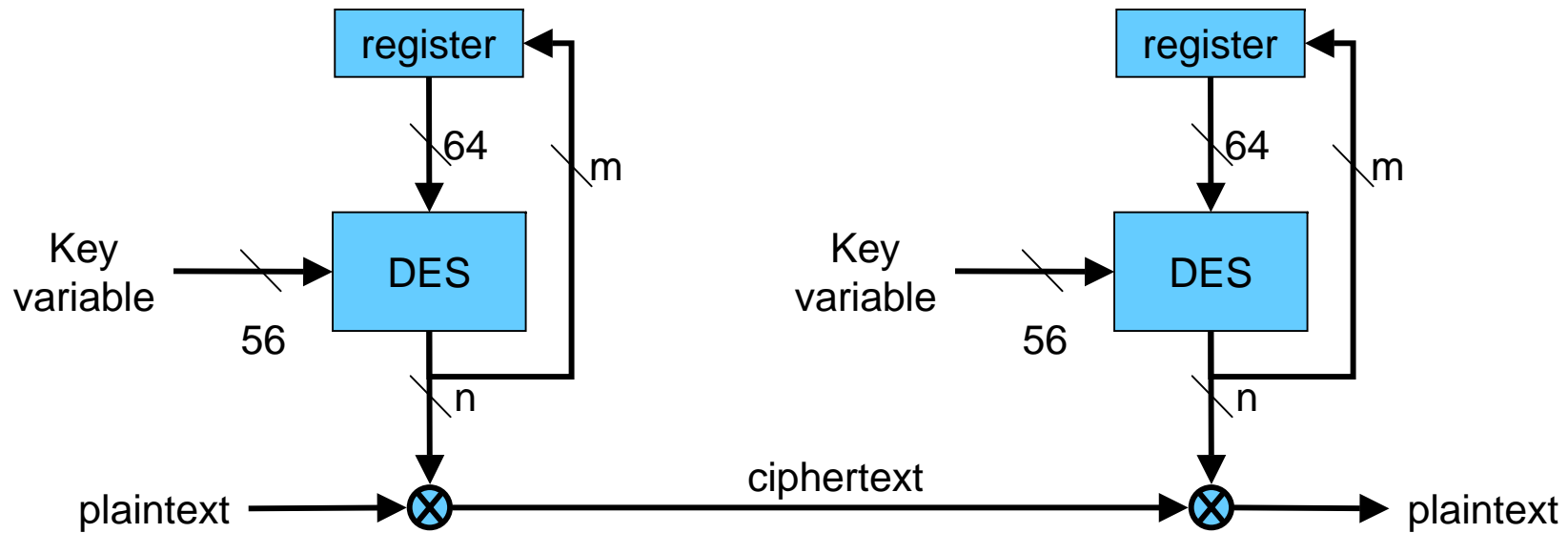


DES Modes



Electronic Codebook Mode

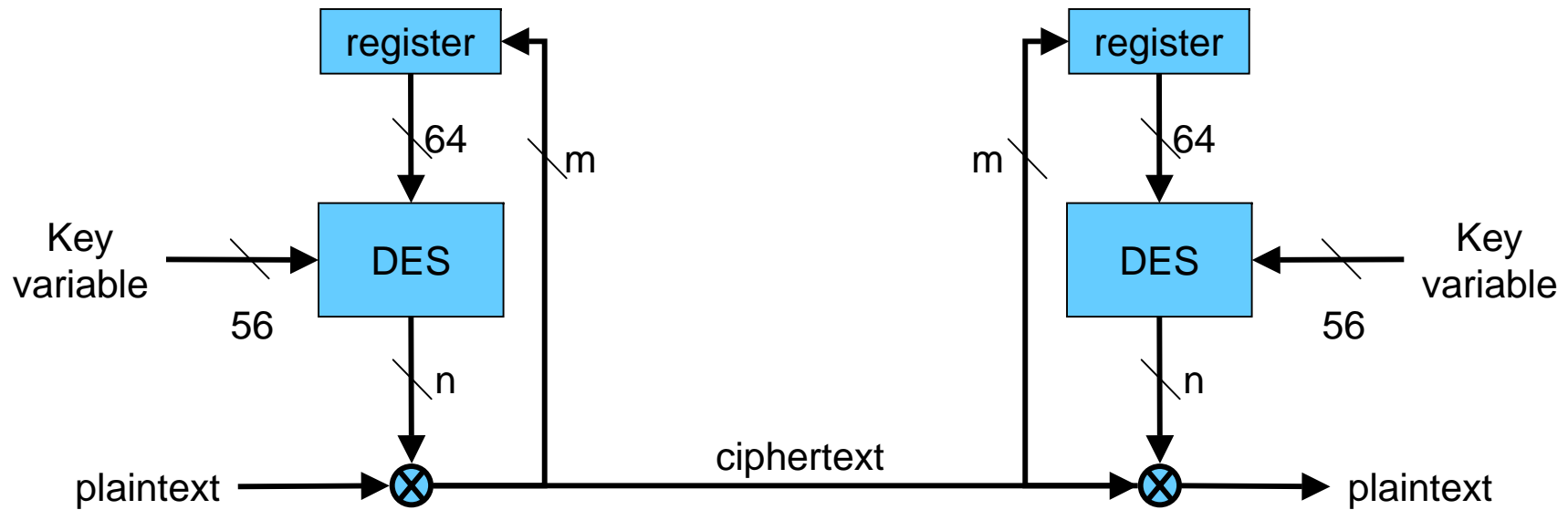
DES Modes



$m, n = 1, 8, 64$

Output Feedback Mode

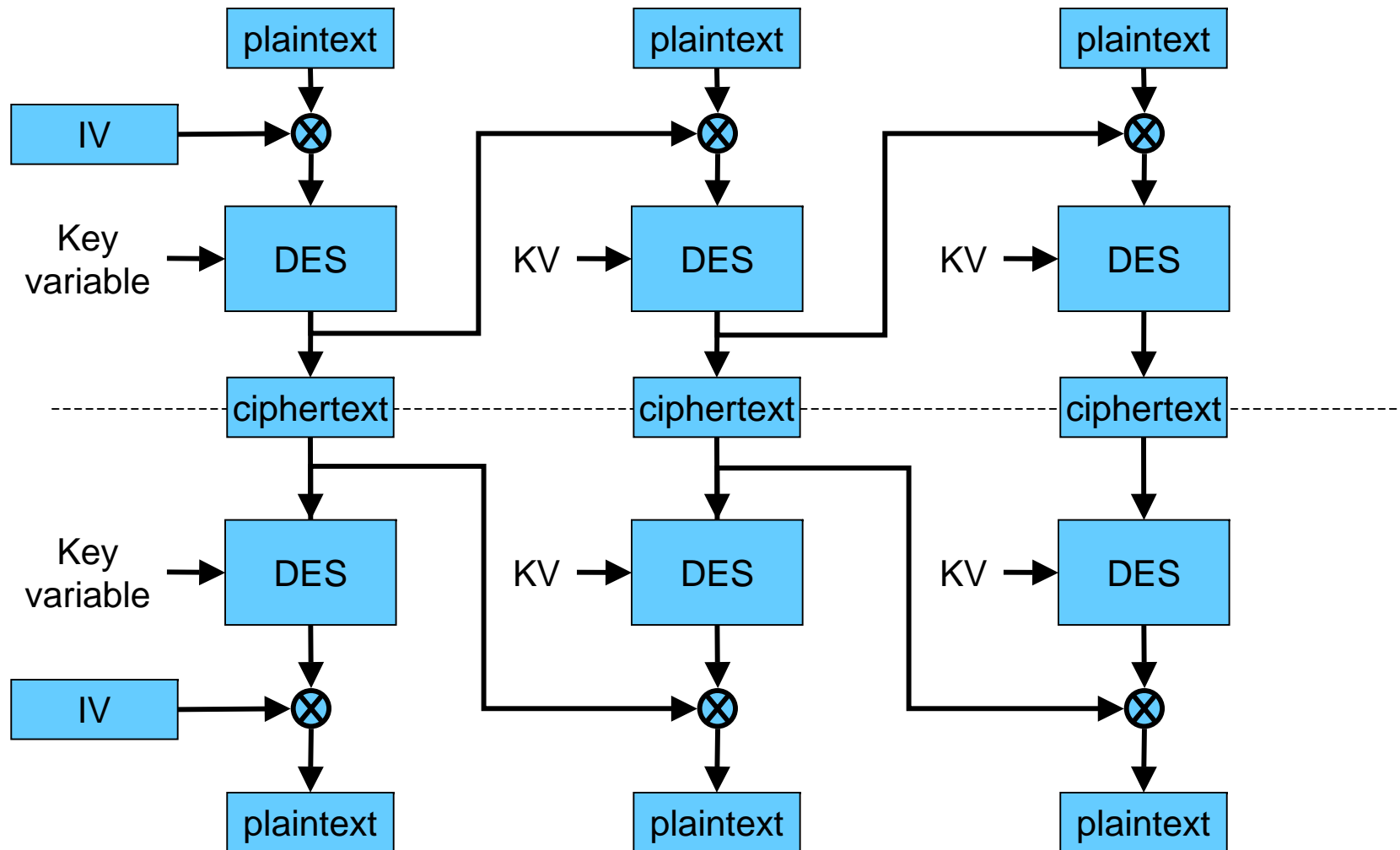
DES Modes



$m, n = 1, 8, 64$

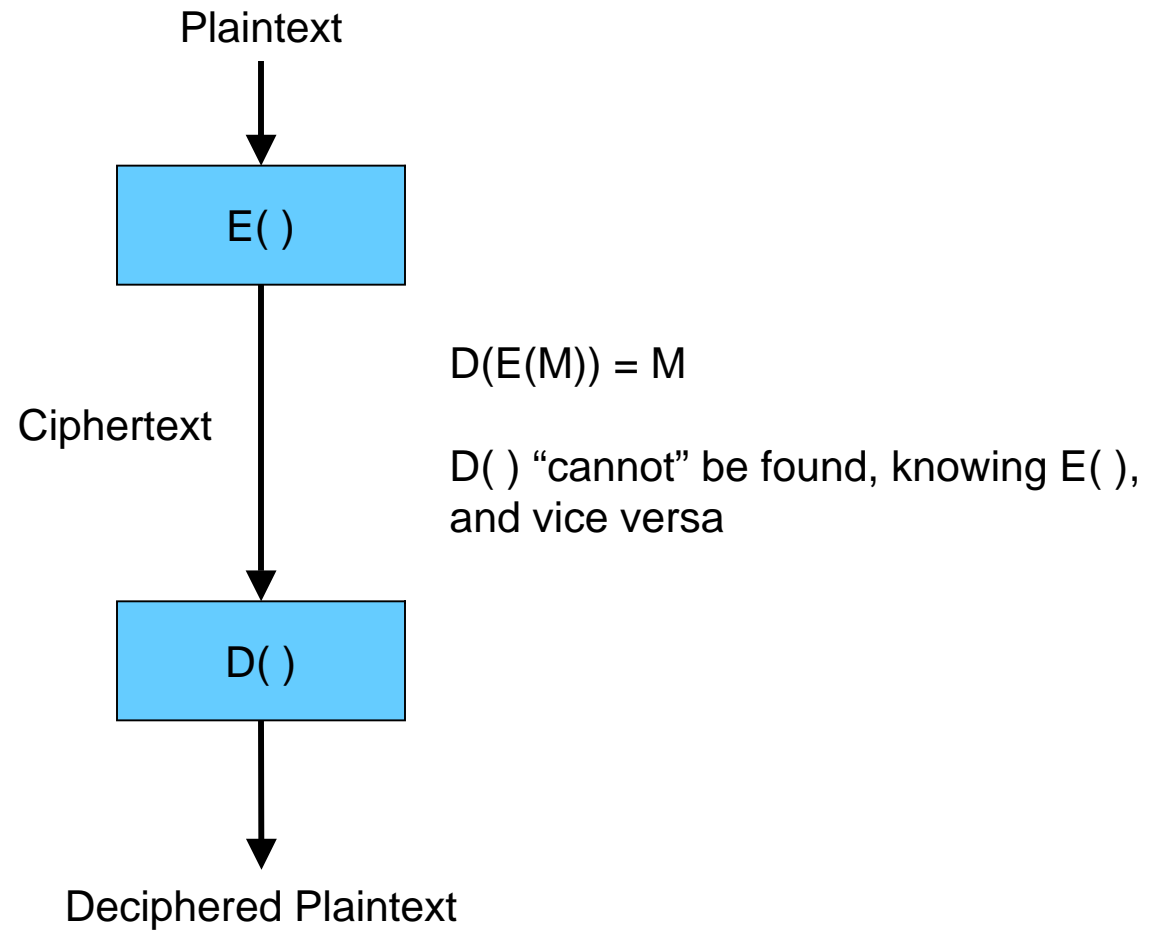
Cipher Feedback Mode

DES Modes

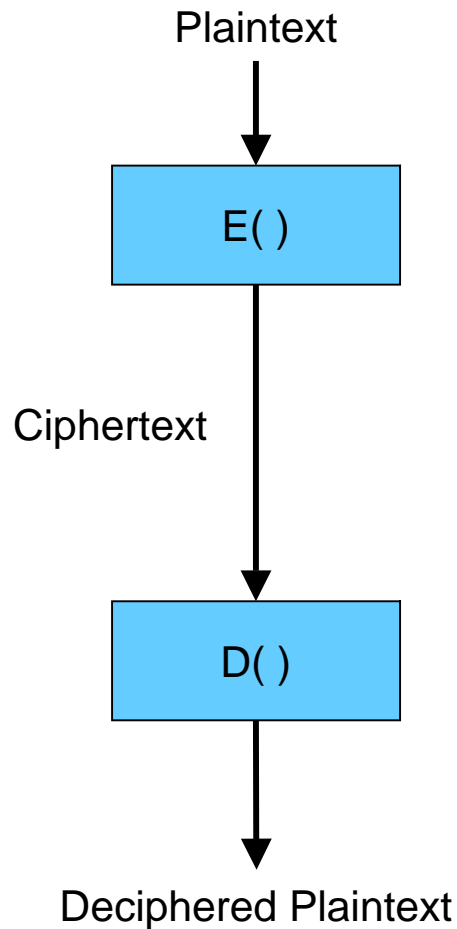


Cipher Block Chaining Mode

Public Key Cryptosystems



Public Key Cryptosystems

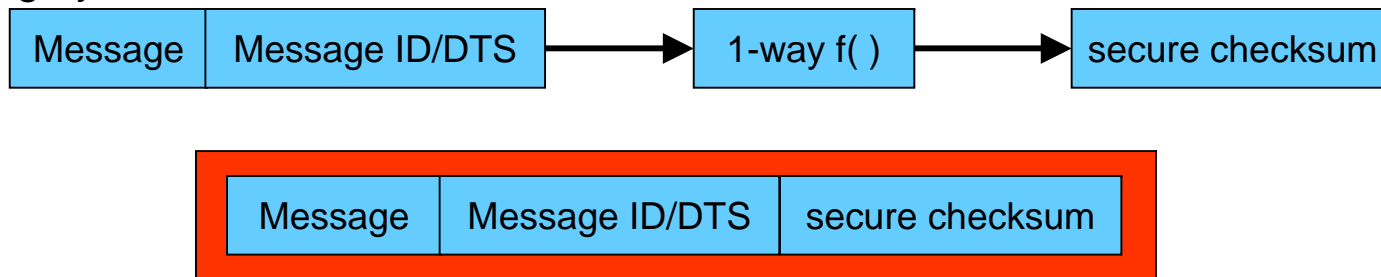


- $D()$ and $E()$ must be built on commutative functions:
 $f(g(x)) = g(f(x))$
- Multiplication and exponentiation work:
 - These form bases for Rivest-Shamir-Adleman (RSA) and Diffie-Hellman PKCs
 - Operations on an elliptic curve in a finite field are used for the El-Gamal and related algorithms
- The apparent security of PKCs come from difficulty of computing logarithms and factoring composite numbers in a finite field. **Thought** to be NP-Complete problems, Which **might** make them mathematically intractable
- E.g.,
 $E(M) = M^e$
 $D(C) = C^d$
 $D(E(M)) = (M^e)^d = M^{ed} = M^1$, if $d=e^{-1}$ in the field

Applications of cryptography to security

- Confidentiality – the most obvious application

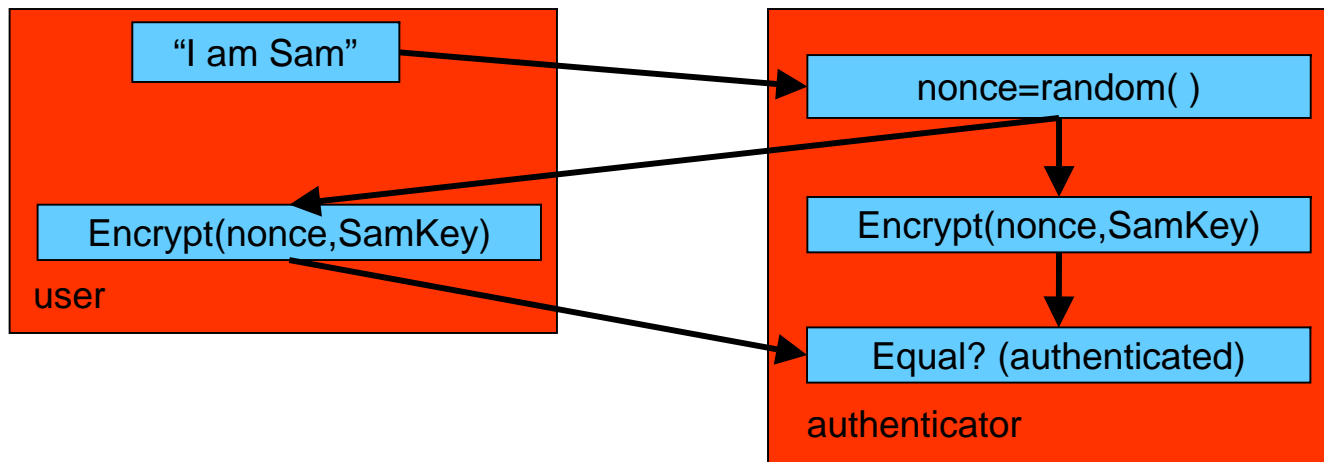
- Integrity



- Non-repudiation

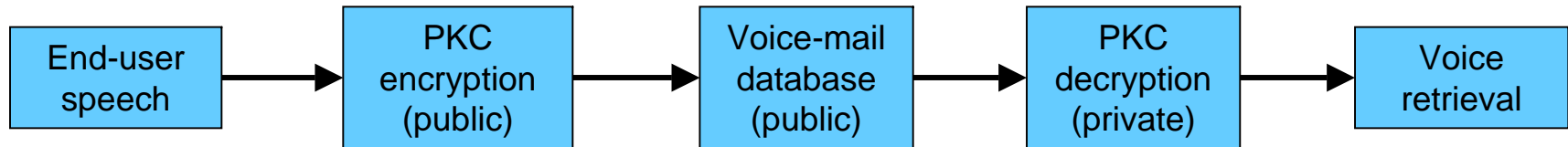
– Same as integrity, but seal the message: with user ID and user-specific key

- Authentication Challenge-response

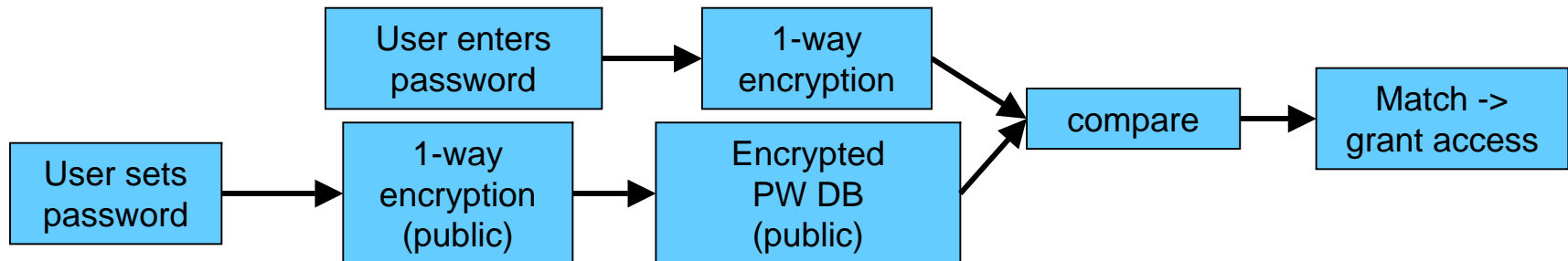


What can go wrong with cryptography?

- Gus Simmons' attack on voice mail system



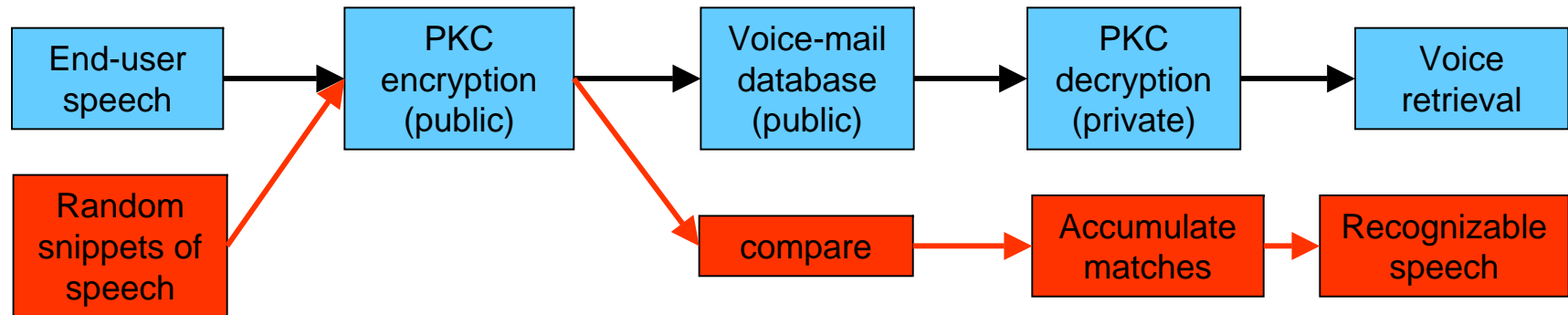
- Hacking UNIX passwords



- 802.11 WEP

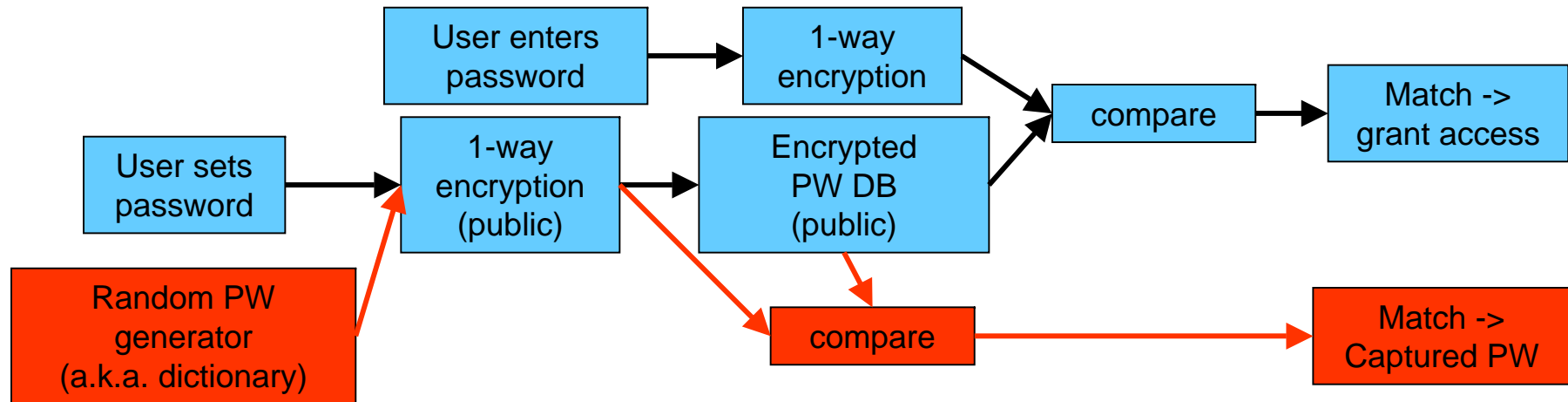
What can go wrong with cryptography?

- Gus Simmons' attack on voice mail system



Redundancy in Source Material

- Hacking UNIX passwords



- 802.11 WEP