What is a Cryptosystem?

- $K = \{0,1\}^l$
- $P = \{0,1\}^m$
- $C' = \{0,1\}^n, C \subseteq C'$
- $E : P \times K \rightarrow C$
- $D : C \times K \rightarrow P$
- $\forall p \in P, k \in K : D(E(p,k),k) = p$
- It is infeasible to find $F : P \times C \rightarrow K$

Let's start again, in English. . .

What is a Cryptosystem?

A cryptosystem is pair of algorithms that take a *key* and convert *plaintext* to *ciphertext* and back.

Plaintext is what you want to protect; ciphertext should appear to be random gibberish.

The design and analysis of today's cryptographic algorithms is highly mathematical. Do *not* try to design your own algorithms.

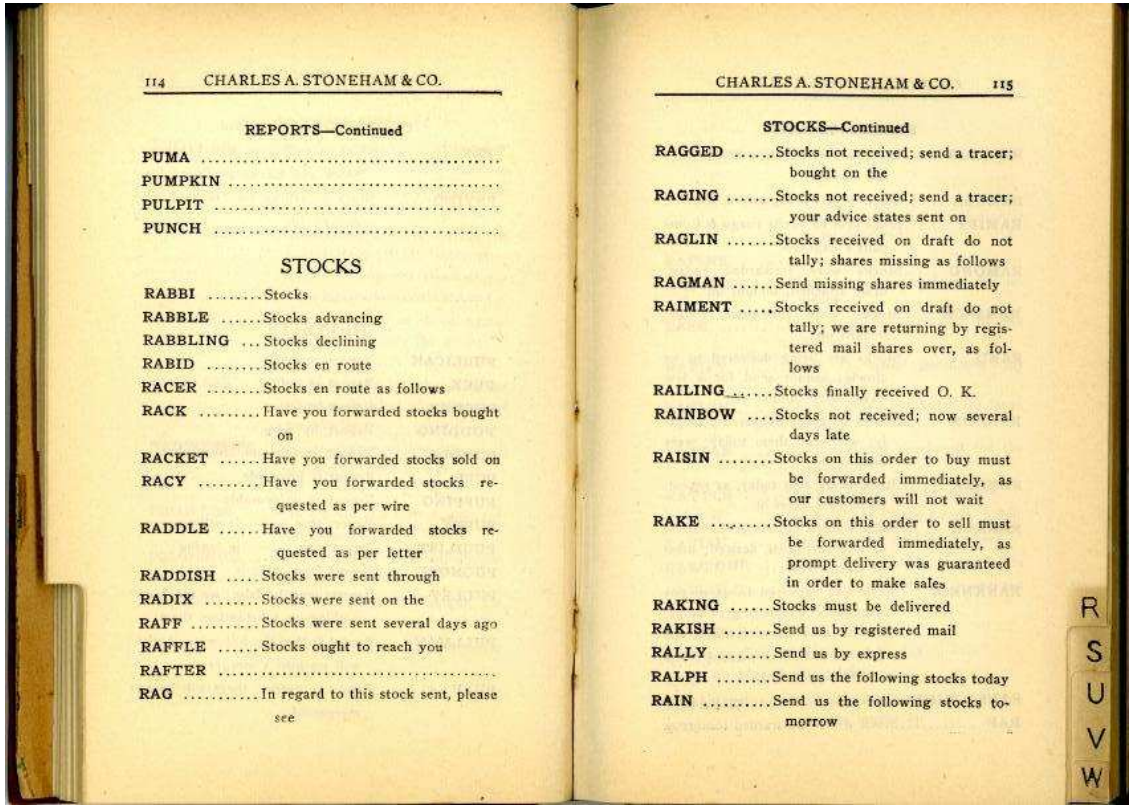
A Tiny Bit of History

- Encryption goes back thousands of years
- Classical ciphers encrypted letters (and perhaps digits), and yielded all sorts of bizarre outputs.
- The advent of military telegraphy led to ciphers that produced only letters.

Codes vs. Ciphers

- Ciphers operate *syntactically*, on letters or groups of letters: $A \rightarrow D$, $B \rightarrow E$, etc.
- Codes operate semantically, on words, phrases, or sentences, per this 1910 codebook

A 1910 Commercial Codebook



114 CHARLES A. STONEHAM & CO.

REPORTS—Continued

PUMA .....
PUMPKIN .....
PULPIT .....
PUNCH .....

STOCKS

RABBI .....Stocks
RABBLE .....Stocks advancing
RABBLING ...Stocks declining
RABID .....Stocks en route
RACER .....Stocks en route as follows
RACK .....Have you forwarded stocks bought on
RACKET .....Have you forwarded stocks sold on
RACY .....Have you forwarded stocks requested as per wire
RADDLE .....Have you forwarded stocks requested as per letter
RADDISH .....Stocks were sent through
RADIX .....Stocks were sent on the
RAFF .....Stocks were sent several days ago
RAFFLE .....Stocks ought to reach you
RAFTER .....
RAG .....In regard to this stock sent, please see

CHARLES A. STONEHAM & CO. 115

STOCKS—Continued

RAGGED .....Stocks not received; send a tracer; bought on the
RAGING .....Stocks not received; send a tracer; your advice states sent on
RAGLIN .....Stocks received on draft do not tally; shares missing as follows
RAGMAN .....Send missing shares immediately
RAIMENT .....Stocks received on draft do not tally; we are returning by registered mail shares over, as follows
RAILING .....Stocks finally received O. K.
RAINBOW ....Stocks not received; now several days late
RAISIN .....Stocks on this order to buy must be forwarded immediately, as our customers will not wait
RAKE .....Stocks on this order to sell must be forwarded immediately, as prompt delivery was guaranteed in order to make sales
RAKING .....Stocks must be delivered
RAKISH .....Send us by registered mail
RALLY .....Send us by express
RALPH .....Send us the following stocks today
RAIN .....Send us the following stocks tomorrow

R S U V W

Commercial Telegraph Codes

- Most were aimed at economy
- Secrecy from casual snoopers was a useful side-effect, but *not* the primary motivation
- That said, a few such codes were intended for secrecy; I have some in my collection, including one intended for union use

Properties of a Good Cryptosystem

- There should be no way short of enumerating all possible keys to find the key from any reasonable amount of ciphertext and plaintext, nor any way to produce plaintext from ciphertext without the key.
- Enumerating all possible keys must be infeasible.
- The ciphertext must be indistinguishable from true random values.

Milestones in Modern Cryptography

1883 Kerckhoffs' Principles

1917-1918 Vernam/Mauborgne cipher (one-time pad)

1920s-1940s Mathematicization and mechanization of cryptography and cryptanalysis

1973 U.S. National Bureau of Standards issues a public call for a standard cipher; this led to the adoption of the Data Encryption Standard (DES)

1976 Diffie and Hellman describe public key cryptography

Kerckhoffs' Law

> The system must not be required to be secret, and it must be able to fall into the hands of the enemy without inconvenience.

In other words, the security of the system must rest entirely on the secrecy of the key.

Vernam/Mauborgne Cipher

- Exclusive-OR a key stream tape with the plaintext
- Online encryption of teletype traffic, combined with transmission
- For a one-time pad — which is provably secure — use true-random keying tapes and *never* reuse the keying material.
- If keying material is reusable, it's called a *stream cipher*

☞ Snake oil alert! *If the key stream is algorithmically generated, it's not a one-time pad!*

The Fall of a Variant

- Really long key tapes are unwieldy, so Vernam tried XORing the output of two modestly-long looped tapes
- Example: key tapes of 999 and 1000 characters
- This repeats — and it was cracked easily, way back when

Mathematicization and Mechanization

- Mechanical encryptors (Vernam, Enigma, Hagelin, Scherbius)
- Mathematical cryptanalysis (Friedman, Rejewski et al, Bletchley Park)
- Machine-aided cryptanalysis (Friedman, Turing et al.)

Standardized Ciphers

- Until the 1970s, most strong ciphers were government secrets
- The spread of computers created a new threat
- Reportedly, the Soviets eavesdropped on U.S. grain negotiators' conversations
- NBS (now called NIST) issued a public call for a cipher; eventually, IBM responded
- The eventual result — via a secret process — was DES

Public Key Cryptography

- Merkle invents a public key distribution scheme
- Diffie and Hellman invent public key encryption and digital signatures, but do not devise a suitable algorithm with all of the desired properties.
  Rivest, Shamir, and Adelman invent their algorithm soon thereafter
- In fact, the British GCHQ had invented “non-secret encryption” a few years earlier.
- There have been claims, but no evidence, that the American NSA invented it even earlier

What We Have Today

- Encryption is completely computerized, and operates on bits
- The basic primitives of encryption are combined to produce very powerful results
- Encryption is by far the strongest weapon in the computer security arsenal; host and operating system software is by far the weakest link
- *Bad software trumps good crypto*

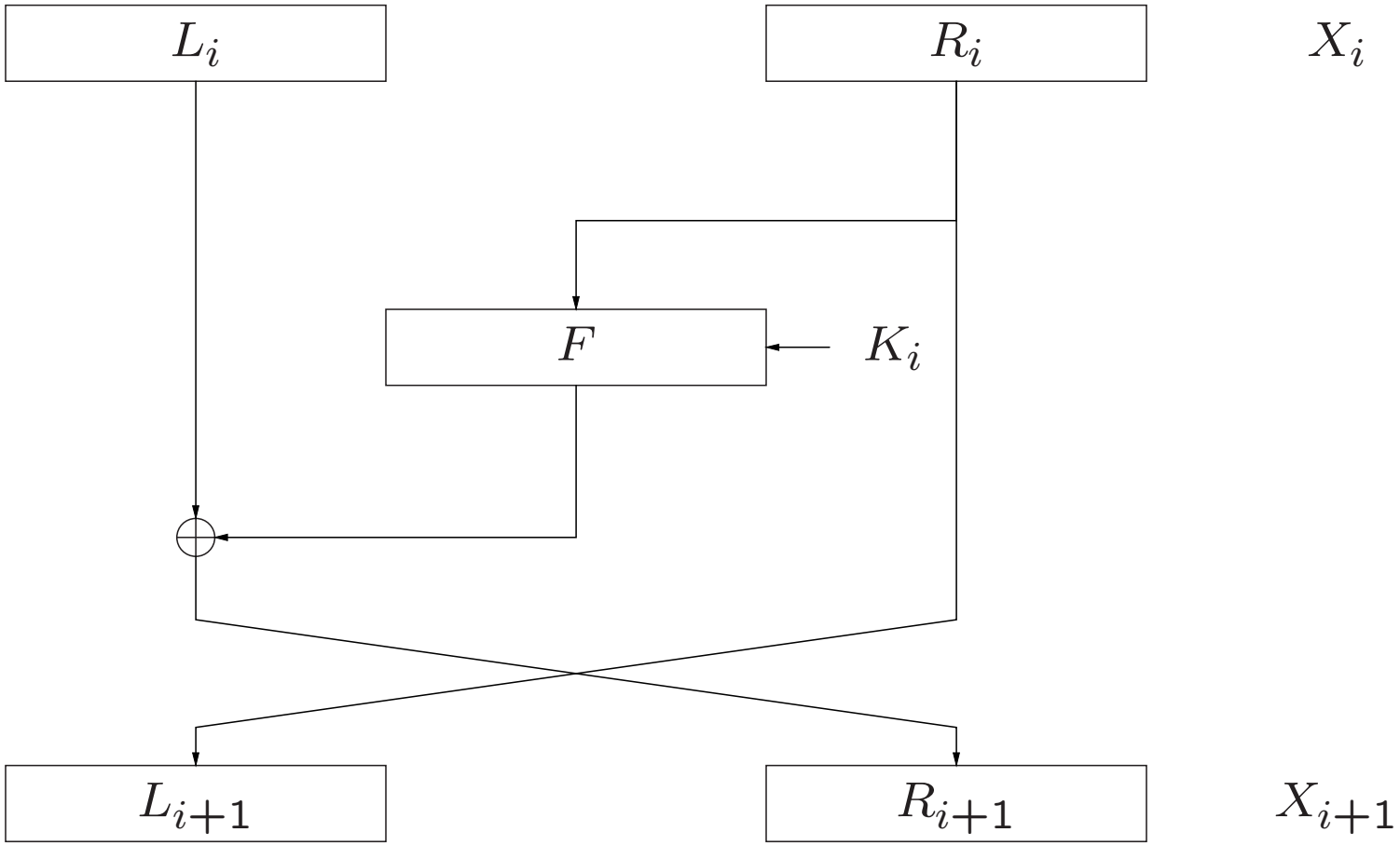
Block Ciphers

- Operate on a fixed-length set of bits
- Output blocksize generally the same as input blocksize
- Well-known examples: DES (56-bit keys; 64-bit blocksize); AES (128-, 192-, and 256-bit keys; 128-bit blocksize)

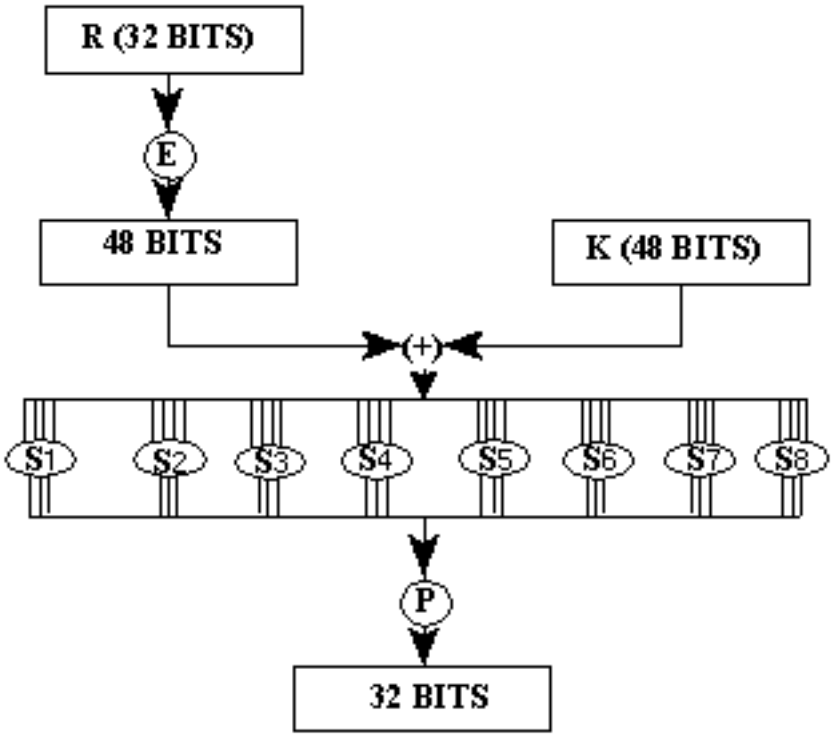
Basic Structure of (Most) Block Ciphers

- Optional key scheduling — convert supplied key to internal form
- Multiple *rounds* of combining the plaintext with the key.
- DES has 16 rounds; AES has 9-13 rounds, depending on key length

DES Round Structure



DES ”f” Funciton



How DES Works

For each round:

1. Divide the input block in half. The right half of each round becomes the left half of the next round's input.
2. Take the right half, pass it through a non-linear function of data and key, and exclusive-OR the result with the current input's left half.
3. The output of that function becomes the right half of the next round's input.
4. This is known as a *Fiestel network*

Decryption

- Run the rounds backwards
- In the example, L_{i+1} is passed unchanged to the previous round (as R_i)
- Accordingly, it can be fed into $F(K_i)$ to be XORed with R_{i+1} to produce L_i

What's Wrong with DES?

- The key size is too short — a machine to crack DES was built in 1998.
- (Charges that NSA could crack DES were leveled in 1979. But the claim that NSA designed in a back door are false.)
- The blocksize is too short.
- It depends on bit-manipulation, and is too slow in software

Selecting the Advanced Encryption Standard

- NIST issued an open call for submissions
- 15 ciphers were submitted, from all over the world
- Several open conferences were held (and the NSA did its own private evaluations)
- 5 ciphers were eliminated as not secure enough
- 5 more were dropped for inefficiency or low security margin
- Of the 5 finalists, Rijndael — a Belgian submission — was chosen because of good security and very high efficiency across a wide range of platforms

How Does Rijndael Work?

- Input block viewed as a byte array; key viewed as a two-dimensional matrix
- Each round consists of a series of simple, byte-oriented operations: ByteSubstitution, ShiftRow, MixColumn, AddRoundKey.
- The key is mixed with the entire block in each round
- The basic operations are individually reasonably tractable mathematically, but are combined in a hard-to-invert fashion.

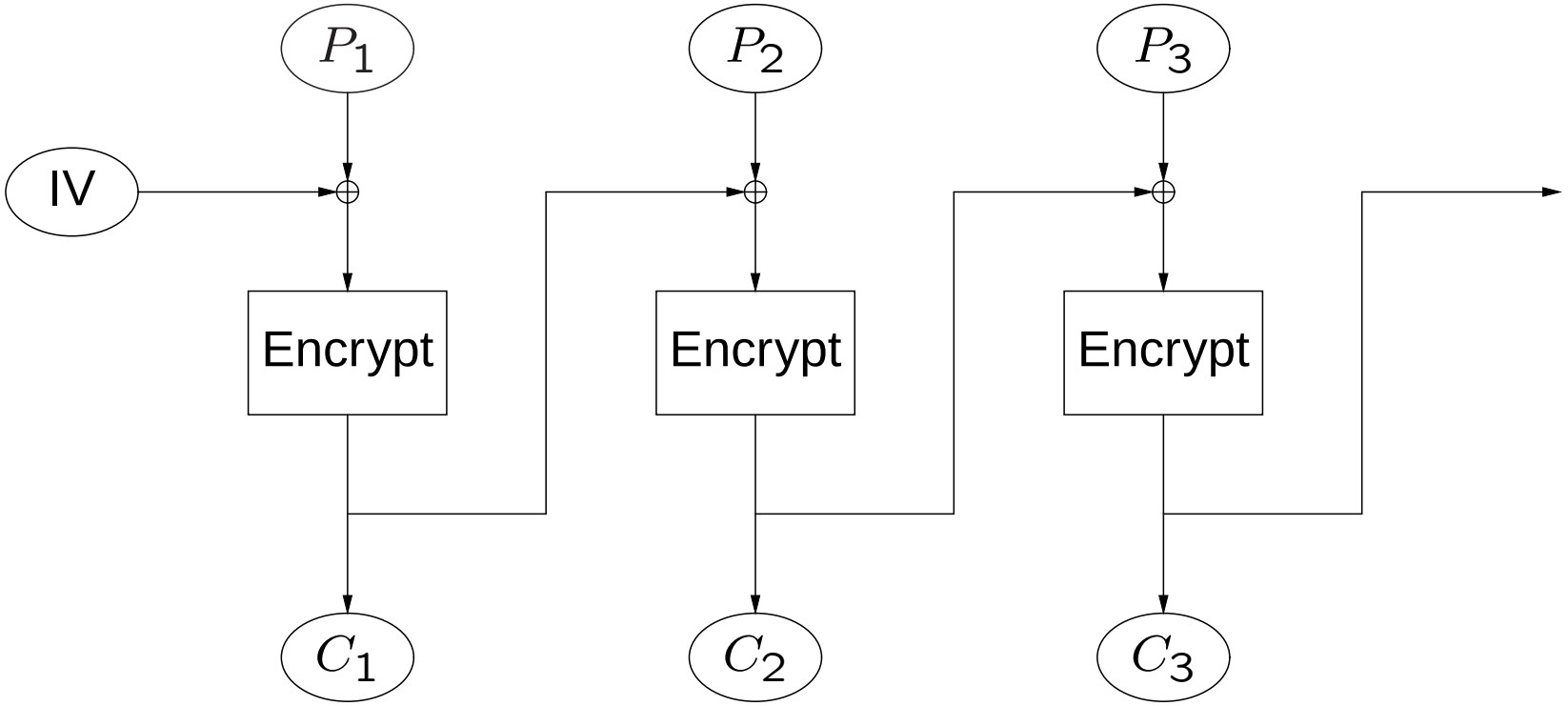
Modes of Operation

- Direct use of a block cipher is inadvisable
- Enemy can build up “code book” of plaintext/ciphertext equivalents
- Beyond that, direct use only works on messages that are a multiple of the cipher block size in length
- Solution: five standard *Modes of Operation*: Electronic Code Book (ECB), Cipher Block Chaining (CBC), Cipher Feedback (CFB), Output Feedback (OFB), and Counter (CTR).

Electronic Code Book

- Direct use of the block cipher
- Used primarily to transmit encrypted keys
- Very weak if used for general-purpose encryption; never use it for a file or a message.
- We write $\{P\}_k \rightarrow C$ to denote "encryption of plaintext P with key k to produce ciphertext C"

Cipher Block Chaining



$$\{P_i \oplus C_{i-1}\}_k \rightarrow C_i$$
$$\{C_i\}_{k^{-1}} \oplus C_{i-1} \rightarrow P_i$$

Properties of CBC

- The ciphertext of each encrypted block depends on the plaintext of all preceeding blocks.
- There is a dummy initial ciphertext block C_0 known as the *Initialization Vector* (IV); the receiver must know this value.
- Consider a 4-block message:

$$\begin{aligned} C_1 &= \{P_1 \oplus IV\}_k \\ C_2 &= \{P_2 \oplus C_1\}_k \\ C_3 &= \{P_3 \oplus C_2\}_k \\ C_4 &= \{P_4 \oplus C_3\}_k \end{aligned}$$

If C_2 is damaged during transmission, what happens to the plaintext?

Error Propagation in CBC Mode

- Look at the decryption process, where C' is a garbled version of C:

$$\begin{aligned} P_1 &= \{C_1\}_{k^{-1}} \oplus IV \\ P_2 &= \{C_2'\}_{k^{-1}} \oplus C_1 \\ P_3 &= \{C_3\}_{k^{-1}} \oplus C_2' \\ P_4 &= \{C_4\}_{k^{-1}} \oplus C_3 \end{aligned}$$

- P_1 depends only on C_1 and IV, and is unaffected
- P_2 depends on C_2 and C_1, and hence is garbled
- P_3 depends on C_3 and C_2, and is also garbled. The enemy can control the change to P_3.
- P_4 depends on C_4 and C_3, and not C_2; it thus isn't affected.
- Conclusion: Two blocks change, one of them predicatably

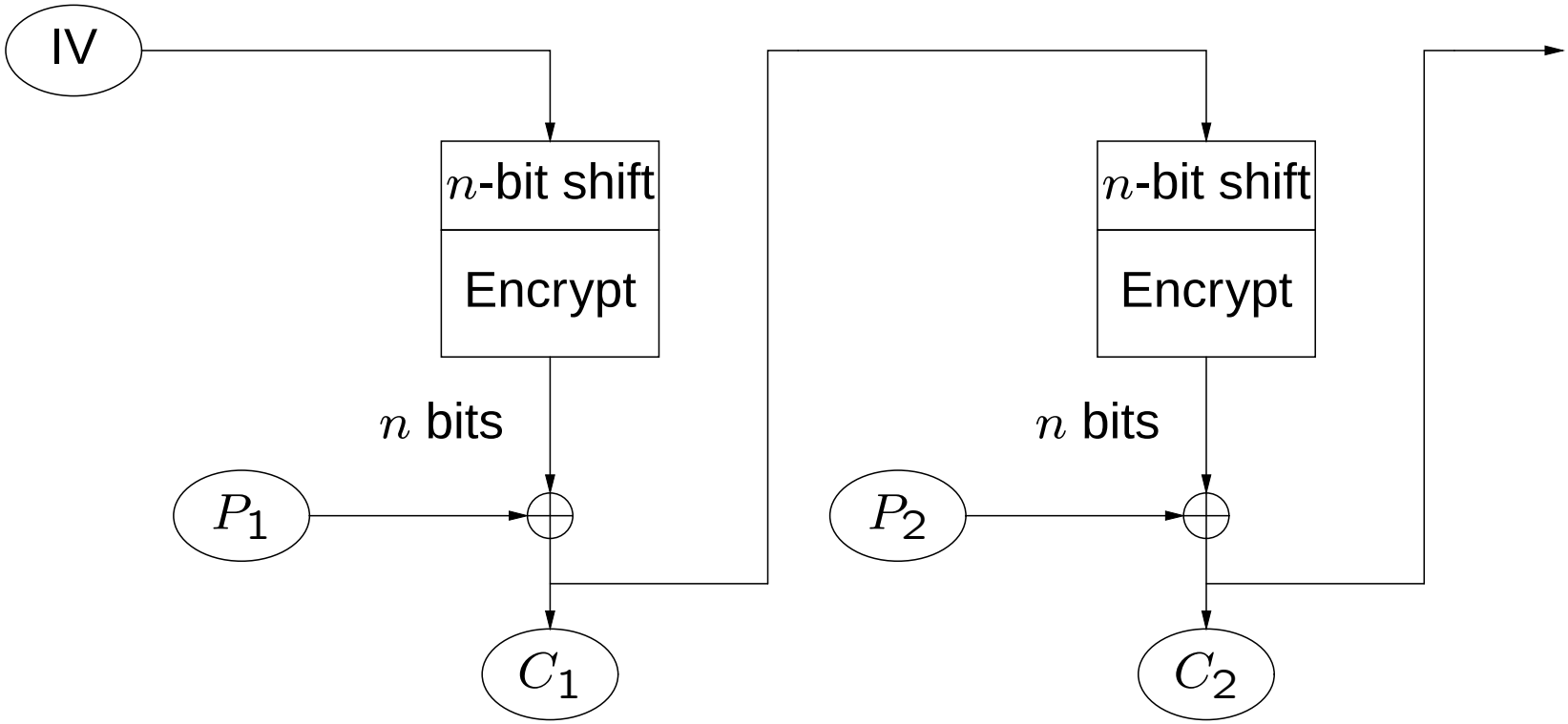
Cutting and Pasting CBC Messages

- Consider the encrypted message

$$IV, C_1, C_2, C_3, C_4, C_5$$

- The shortened message IV, C_1, C_2, C_3, C_4 appears valid
- The truncated message C_2, C_3, C_4, C_5 is valid: C_2 acts as the IV.
- Even C_2, C_3, C_4 is valid, and will decrypt properly.
- Any subset of a CBC message will decrypt cleanly.
- If we snip out blocks, leaving IV, C_1, C_4, C_5, we only garble one block of plaintext.
- Conclusion: if you want message integrity, you have to do it yourself.

n-bit Cipher Feedback



$$P_i \oplus \{C_{i-1}\}_k \rightarrow C_i$$
$$\{C_{i-1}\}_k \oplus C_i \rightarrow P_i$$

Properties of Cipher Feedback Mode

- Underlying block cipher used only in encryption mode
- Feedback path actually incorporates a shift register; some of the previous cycle's ciphertext can be retained.
- 8-bit CFB is good for asynchronous terminal traffic.
- Errors propagate while bad data is in the shift register — 17 bytes for CFB_8 when using AES.
- Copes gracefully with deletion of n-bit unit