Introdução à Criptografia:
Cifras de Fluxo e Cifras de Bloco

Roteiros e tópicos para estudo por

Vinícius da Silveira Serafim
professor@serafim.eti.br
http://professor.serafim.eti.br

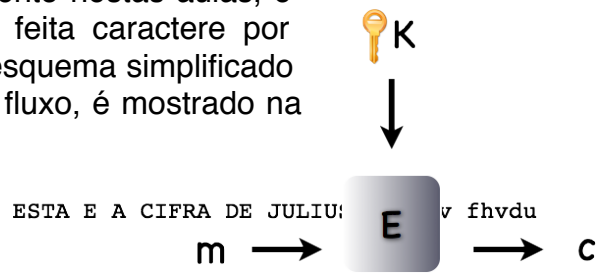
Palavras-chave: cifra de fluxo, cifra de bloco, ecb, cbc, modo de operação

Cifras de Fluxo

Definição: são cifras (ou algoritmos de cifragem) que convertem imediatamente um símbolo (ex.: bit, byte, caractere) do texto em claro em um símbolo do texto cifrado. (PFLEEGER, 2006)

O que Pfleeger quis dizer, ao utilizar a palavra imediatamente, é que a cifragem é feita símbolo por símbolo, de forma independente, isto é, a cifragem de um não interfere e nem depende da cifragem dos demais.

A cifra de César, que já é do nosso conhecimento nestas aulas, é um exemplo de cifra de fluxo. A cifragem é feita caractere por caractere, conforme vimos em exemplos. Um esquema simplificado da cifra de César e, portanto, de uma cifra de fluxo, é mostrado na figura ao lado.



A medida que os caracteres vão sendo fornecidos como entrada para o algoritmo de cifragem, imediatamente, vão sendo obtidos os resultados da cifragem de cada um deles.

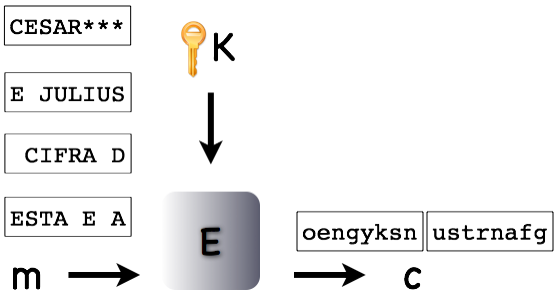
Cifras de Bloco

Definição: cifram um *grupo* de símbolos do texto em claro como *um* bloco. (PFLEEGER, 2006)

No caso das cifras de bloco, a mensagem é dividida em blocos de tamanho fixo (ex.: 64 bits, 128 bits). E cada bloco é cifrado como uma unidade. Ao invés de caractere por caractere, a mensagem é cifrada bloco por bloco.

Na figura a seguir é representado, de forma simples, o processo de cifragem por bloco da mensagem “ESTA E A CIFRA DE JULIUS CESAR”. Note que a mensagem foi dividida em blocos de 8 caracteres (tamanho escolhido por mim para o exemplo).

Você deve ter percebido que, no último bloco contendo a palavra CESAR, estão presentes três asteriscos que não fazem parte mensagem original. O que ocorre é que o tamanho da mensagem nem sempre será múltiplo do tamanho do bloco, assim é comum que no último bloco faltem alguns símbolos (caracteres no caso da cifra de César) para completar o bloco. Nesses casos é feito o ajuste (em inglês: Padding) adicionando o número de símbolos que faltam ao último bloco.



Cifras de Fluxo *versus* Cifras de Bloco

Abaixo, na tabela, um comparativo entre as vantages e desvantagens de ambas cifras. (PFLEEGER, 2006)

	Cifras de Fluxo	Cifras de Bloco
Vantagens	• *Velocidade da cifragem/ decifragem*: o símbolo pode ser transformado assim que recebido pelo algoritmo, sem ter que esperar pelos demais símbolos. • *Baixa propagação de erros*: se houver um erro na transformação de um símbolo, esse erro não afeta a transformação dos demais	• *Alta difusão*: o texto cifrado de um bloco depende de vários símbolos do texto em claro. • *Imunidade à inserção de símbolos*: se um símbolo for inserido em um bloco cifrado, o tamanho do bloco será alterado e a inserção será imediatamente detectada no momento da decifragem, pois o bloco terá um tamanho incorreto.
Desvantagens	• *Baixa difusão*: cada símbolo do texto cifrado corresponde à informação de apenas um símbolo do texto em claro. • *Suscetibilidade à inserção e modificação maliciosa*: novos símbolos podem ser adicionados, o atacante opera com um caractere por vez.	• *Mais lentos para cifrar/ decifrar*: o algoritmo tem que receber todo o bloco para então cifrá-lo. • *Propagação de erros*: um erro irá afetar a transformação de todos os caracteres do bloco.

As cifras de bloco acabam sendo mais seguras que as cifras de fluxo devido a sua característica de alta difusão (estudada no roteiro anterior). Por essa razão as cifras (ou algoritmos) de bloco são mais comuns que as cifras de fluxo. Quando tiver que escolher, escolha cifras de bloco.

Cifras de bloco: modos de operação

As cifras de bloco possuem ainda outra particularidade: os modos de operação. O modo de operação é uma escolha feita algumas vezes pelo projetista de um sistema, outras vezes pelos desenvolvedores e, em alguns casos, deixada a cargo do usuário. Independente de quem, é comum essa escolha ser feita sem o claro entendimento do que ela significa e qual o impacto dela na segurança. (Você lembra o que comentei no roteiro anterior sobre a segurança do ambiente?)

Entre os diversos modos disponíveis, vamos abordar apenas dois: Electronic CodeBook (ECB) e o Cipher Chaining Block (CBC). Me adiantando um pouco, o primeiro você nunca deve utilizar, o segundo é a escolha mais indicada. (FERGUSON, 2003) Use outro no lugar do CBC se você souber o que está fazendo e tiver fortes razões para isso.

Electronic CodeBook (ECB)

O problema do modo ECB é semelhante ao problema da baixa difusão das cifras de fluxo. Porém, no caso do ECB, ao invés de símbolos, são os blocos que são cifrados de forma independente uns dos outros.

Uma mensagem **m** a ser cifrada será dividida em blocos e seus blocos serão cifrados um por um. A fórmula abaixo representa a cifragem do bloco *n* (bn):

$$C_{bn} = E(K, M_{bn})$$

Se, na mensagem **m**, houverem blocos com o mesmo conteúdo o resultado da cifragem para esses blocos será o mesmo. O detalhe é que um atacante pode coletar diversas mensagens cifradas **c** com a mesma chave **K** e buscar por padrões de formação.

Numa aplicação real o ataque é ainda mais fácil, pois as mensagens trocadas entre duas partes (procurem a definição de parte no roteiro sobre os objetivos da segurança) obedecem a um padrão de formação (imagine um formulário eletrônico). Dessa forma fica ainda mais fácil para um atacante inferir informações a partir da análise dos padrões de repetição dos blocos cifrados sem nunca ter que descobrir a chave ou quebrar o algoritmo utilizado.

Cipher Chaining Block (CBC)

O CBC inclui na cifragem de cada bloco o resultado da cifragem do bloco anterior. Dessa forma, uma diferença em um bloco de texto claro irá alterar o resultado da cifragem de todos os blocos subsequentes. Compare a fórmula abaixo com a fórmula do modo ECB:

$$C_{bn} = E(K, M_{bn}, C_{bn-1})$$

Agora, na cifragem do bloco *n*, entram na equação não só a chave e o próprio bloco, mas o resultado da cifragem do bloco *n-1*. Isso reduz substancialmente a probabilidade de ocorrência do problema do modo ECB.

Ainda resta uma questão a ser resolvida: como fica a cifragem do primeiro bloco, já que não há bloco anterior? A resposta para isso é utilizar como primeiro bloco um *vetor de inicialização* (VI) que pode ser um número randômico (aleatório) pra cada mensagem ou um número sequencial incrementado a cada mensagem **m** cifrada.

Importante: um IV fixo (sempre o mesmo para todas as mensagens) não deve ser utilizado pois gera o problema do modo ECB para o primeiro bloco de todas as mensagens. Para entender melhor isso vamos ver a fórmula da cifragem do primeiro bloco:

$$C_{bn} = E(K, M_{bn}, VI)$$

A única diferença dessa formula para a anterior é que o valor da cifragem do bloco anterior C_{bn-1} foi substituído pelo vetor de inicialização *VI*. Se os *primeiros blocos* de mensagens distintas forem iguais, a chave de cifragem for a mesma e o VI for fixo, o resultado da cifragem desses dois blocos será igual. Altere somente o VI para cada uma delas e os resultados serão completamente diferentes.

Assim, usando uma cifra de bloco em modo CBC e com um VI diferente para cada mensagem, mesmo que seja utilizanda a mesma chave de cifragem e mesmo para duas mensagens com conteúdos iguais, os resultados das cifragens serão diferentes. Dificulta-se assim, substancialmente, o trabalho do atacante ao tentar identificar padrões de formação das mensagens através do texto cifrado das mesmas.

Em resumo

- Use sempre cifras de bloco ao invés de cifras de fluxo.
- Use o modo CBC e nunca o ECB.
- Utilize vetores de inicialização diferentes para cada mensagem
- E saiba o porque de cada um dos itens anteriores

Bibliografia

ANDERSON, R. J. Security Engineering: A Guide to Building Dependable Distributed Systems. 2a edição. Wiley, 14 de abril de 2008. 1080 pág.

FERGUSON, N. F.; SCHNEIER, B. Practical Criptography. 1a edição. Wiley, 17 de abril de 2003. 432 pág.

MENEZES, A.; OORSCHOT P. V.; VANSTONE, S. **Handbook of Applied Cryptography**. 1a edição. CRC Press, 16 de dezembro de 1996. 780 pág.

PFLEEGER, C. P. Security in Computing. 4a edição. Prentice Hall, 23 de outubro de 2006. 880 pág.

SCHNEIER, B. Applied Cryptography: Protocols, Algorithms, and Source Code in C. 2a edição. Wiley, 18 de outubro de 1996. 758 pág.