**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**UNIDADE DE CURITIBA**

**DEPARTAMENTO ACADÊMICO DE INFORMÁTICA**

**CURSO DE ENGENHARIA DE COMPUTAÇÃO**

**Trabalho 01: Implementação de Cifradores Simétricos**

**Leonardo Winter Pereira**

Relatório apresentado como requisito parcial da disciplina de ‘Segurança e Auditoria de Sistemas’ do Curso de Engenharia de Computação da UTFPR.

**Curitiba**

**2017**

**Sumário**

[Introdução 3](#_Toc491472471)

[Desenvolvimento 4](#_Toc491472472)

[**Cifra de César** 4](#_Toc491472473)

[**Analisador de frequência** 4](#_Toc491472474)

[**Cifra de Vernam** 6](#_Toc491472475)

[**Cifrador RC4** 7](#_Toc491472476)

[Considerações finais 8](#_Toc491472477)

[Referências Bibliográficas 8](#_Toc491472478)

# Introdução

Cifradores simétricos, também conhecidos como cifradores convencionais ou cifradores de chave única, representam uma classe de algoritmos criptográficos cuja principal característica é a utilização de uma única chave para cifrar / decifrar uma informações. Tal característica, no entanto, também é a maior desvantagem destes algoritmos quando comparados com cifradores mais avançados (cifradores assimétricos), devido à facilidade de se conseguir encontrar a chave em questão.

Neste relatório, serão explicitados a implementação de cifradores simétricos bastante conhecidos: o cifrador de César e o cifrador de Vernan, bem como um analisador de frequência. Também é relatado o uso do cifrador RC4.

# Desenvolvimento

## **Cifra de César**

Dentre todos os algoritmos criptográficos, a cifra de César é provavelmente a mais conhecida e a mais simples de todas.

Trata-se de um tipo de cifra de substituição, na qual cada letra de um texto criptografado é substituído por outra.[1] Esta é definida através de um simples deslocamento da letra original por um certo número de posições. Este número é chamado de chave.

Para o desenvolvimento e implementação do algoritmo, foi utilizado a linguagem de programação JAVA, através da IDE Eclipse. No projeto, o alfabeto foi composto não apenas por letras, mas também por números ([A-Z,a-z,0-9]). Aumentando as possibilidades de chaves de 26 para 62 caracteres.

A operação matemática para criptografar e decriptografar cifras de César (considerando apenas as 26 letras do alfabeto) é ilustrada pelas Figuras 1 e 2.

E_n(x) = (x + n) \mod {26}.

Figura 1: Fórmula para codificar uma cifra de César.

D_n(x) = (x - n) \mod {26}.

Figura 2: Fórmula para decodificar uma cifra de César.

A implementação do algoritmo de César é relativamente simples e seu pseudo código é descrito a seguir:

1. Verificar se o caractere é um número, letra minúscula ou maiúscula.
2. Verifica se, com a chave N, a rotação do caractere ultrapassa os limites superiores para codificação e inferiores para decodificação.
3. Caso positivo, o caractere é alternado para o limite inferior do próximo grupo (número, minúscula ou maiúscula) ou superior, caso seja decodificação. Assim, a chave é diminuída da diferença entre o caractere e o respectivo limite para compensação. Assim, a tabela de caracteres ignora os caracteres entre números, letras minúsculas e maiúsculas, de forma a não se perceber estes caracteres na rotação.
4. Repete o passo 2 com a nova chave, e se necessário, o passo 3. Caso não extrapole, realiza a rotação normalmente.
5. Repete os passos para todos os caracteres da string, excluindo espaço e ponto.

## **Analisador de frequência**

A análise de frequência é o principal método empregado para decifrar mensagens criprografadas por um algoritmo de substituição (como é o caso da cifra de César). O método onsiste em verificar a ocorrência de uma letra e comparar este valor com uma base confiável para tentar descrobrir a chave que foi utilizada para realizar a codificação do texto em questão.

A mensagem a ser analisada através da análise de frequência, fornecida pelo professor, consiste em:

g5Bt5 t54yvtz3v4A5 wrG t53 7Bv r9 6v995r9 9v 9z4Ar3

58xB2y59r9. dBzA5 t54yvtz3v4A5, 7Bv 9v 9z4Ar3

yB3z2uv9. Vy r99z3 7Bv r9 v96zxr9 9v3 x8r59 v8xBv3

uv9uv4y59r3v4Av r trsvtr 6r8r 5 tvB, v47Br4A5 r9

tyvzr9 r9 srzEr3 6r8r r Av88r, 9Br 3rv.

cv54r8u5 Ur mz4tz.

Realizando o processo de análise de frequência de todas as letras do texto acima e comparando os resultados obtivos com a frequência do uso das letras na língua portuguesa (ilustrado na figura 3), obteve-se o resultado demonstrado na figura 4.

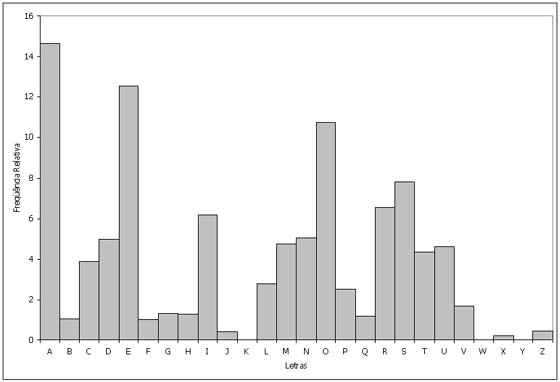


Figura 3: Frequência de uso das letras na [língua portuguesa](https://pt.wikipedia.org/wiki/L%C3%ADngua_portuguesa) [2]

Sabendo que o caractere de maior frequência na língua portuguesa é o 'A' e que a criptografia utilizada foi a cifra de César, a chave pode ser facilmente encontrada realizando uma simples operação matemática entre e a letra de maior frequência no texto (‘R’) e a letra ‘A’. Desta forma, encontramos N = 17.

O texto decriptografado segue:

“Pouco conhecimento faz com que as pessoas se sintam

orgulhosas. Muito conhecimento, que se sintam

humildes. Eh assim que as espigas sem graos erguem

desdenhosamente a cabeca para o ceu, enquanto as

cheias as baixam para a terra, sua mae.”

Leonardo Da Vinci.

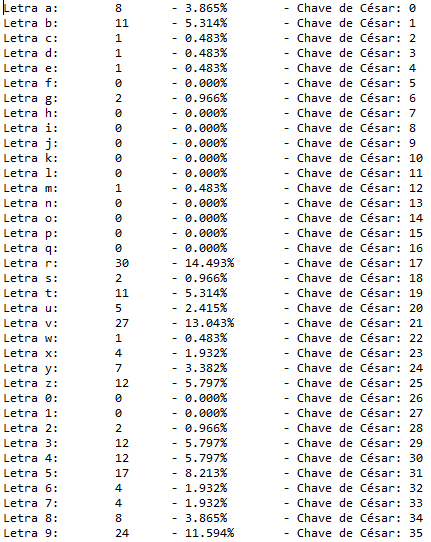


Figura 4: Frequência das letras no texto a ser decriptografado

## **Cifra de Vernam**

A cifra de Vernam, bem como a cifra de César, é uma cifra de substituição e utiliza a mesma lógica de rotação de caracteres que o cifrador de César.

A grande diferença da cifra de Vernam é que, enquanto um texto codificado através da cifra de César é facilmente decifrado através do processo de análise de frequência, o mesmo não pode ser dito para textos codificados através da cifra de Vernam.

Utilizando-se de um segundo texto como chave, este algoritmo se baseia na realização da operação XOR (OU exclusivo) entre os caracteres dos dois textos. Desta forma, não é possível detectar o caractere com maior frequência (como é o caso na cifra de César), pois é altamente improvável que todas as letras iguais do texto original realizem a operação XOR com o mesmo caractere na chave.

Entretanto, a maior vantagem da cifra de Vernam é também sua maior desvantagem . Uma vez que, para não poder ser analisado através de um analisador de frequência, o texto utilizado como chave deve ter comprimento maior ou igual ao texto original, a utilização desta cifra para criptografar arquivos grandes necessita de uma chave igualmente grande.

## **Cifrador RC4**

O cifrador RC4 é um cifrador de fluxo, onde os bits originais se combinam, geralmente com XOR, com bits de cifragem originados de um gerador de números “pseudo” aleatório.

Este cifrador é muito utilizado em ambientes onde a performance é um fator crucial[3], por ser extremamente rápida para processadores e por trabalhar com 1 byte de cada vez.

Este cifrador, assim como o de Vernam, não é vulnerável à análise de frequência. Com a chave possuindo um tamanho grande o suficiente, um mesmo caractere no texto original realizará a operação XOR com diferentes caracteres da chave e, portanto, resultará em diferentes caracteres no texto cifrado.

# Considerações finais

Algoritmos de criptografia com chaves simétricas são bastante simples de se implementar. Entretanto, por serem altamente dependentes de uma única chave, tornam-se vulneráveis a interceptações.

É possível, no entanto, criptografar textos de forma altamente complexa utilizando cifradores de Vernam e RC4, mesmo que a custo de chaves com tamanho elevado.

# Referências Bibliográficas

[1] <http://www.bosontreinamentos.com.br/seguranca/criptografia-cifra-de-cesar/>

[2] <https://pt.wikipedia.org/wiki/An%C3%A1lise_de_frequ%C3%AAncia>.

[3 ] http://www.cifraextrema.com/single-post/2016/09/21/RC4---adeus-ao-favorito-do-TLS-e-outros