Esteganografia usando Algoritmo Genético de Chaves Aleatórias Viciadas

Steganography using Biased Random Keys Genetic Algorithm

Getulio dos Santos Araujo[1] e Warley Gramacho da Silva[1]

[1] *Universidade Federal do Tocantins, Curso de Ciência da Computação, Palmas, Tocantins, Brasil*

Data de recebimento do manuscrito: 30/08/2023
Data de aceitação do manuscrito: 31/08/2023
Data de publicação: 16/10/2023

Resumo— A esteganografia é vista com bons olhos no mundo digital como uma alternativa de envio de dados via arquivos em formato digital, mas há questionamentos quanto à sua qualidade e segurança, visto que uma simples degradação a expõe à interceptação de dados. Para minimizar isso, o uso da técnica da matriz de substituição tende a "mascarar" as informações embutidas, que, se remodeladas, podem ser representadas por um vetor de tamanho n, com n valores diferentes entre 0 e $n-1$ que são ordenados de uma maneira específica pode determinar a sequência de bits que substituirá uma determinada sequência de bits. Encontrar uma configuração ótima para uma matriz de substituição é um problema combinatório e deterministicamente inviável, exigindo o uso de heurísticas para encontrar uma solução quase ótima. No presente projeto aplicamos os conceitos de Algoritmos Genéticos com Chaves Aleatórias Aleatórias (BRKGA) para encontrá-lo, ao mesmo tempo que é atestada sua eficiência em relação à abordagem da esteganografia, aplicações de seus conceitos em Algoritmos Genéticos convencionais. Ao final, concluímos que alguns aspectos podem ter contribuído para sua eficiência em relação à abordagem AG proposta por [1].

Palavras-chave—esteganografia; matriz de substituição; problema combinatorial; técnicas heurística; BRKGA.

Abstract— *Steganography is seen favorably in the digital world as an alternative for sending data via files in digital format, but there is some questioning regarding its quality and security, given that a simple degradation exposes it to data interception. In order to minimize this, the use of the substitution matrix technique tends to "mask" the embedded information, which, if remodeled, can be represented by a vector of size n, with n different values between 0 and n-1 that are ordered in a specific way can determine the sequence of bits that will replace a given sequence of bits. Finding an optimal configuration for a substitution matrix is a combinatorial problem, and deterministically infeasible, requiring the use of heuristics to find a quasi-optimal solution. In the present project we apply the concepts of Genetic Algorithms with Biased Random Keys (BRKGA) to find it, at the same time that its efficiency is attested in relation to the approach of steganography, applications of its concepts in conventional Genetic Algorithms. In the end, we concluded that some aspects may have contributed to its efficiency in relation to the GA approach proposed by [1].*

Keywords—*steganography; substitution matrix; combinatorial problem; heuristic techniques; BRKGA.*

I. Introdução

O termo esteganografia pode ser definido como a arte de ocultar informações, tornando-as ocultas para não serem percebidas por terceiros. Ao utilizar a técnica de esteganografia para ocultar informações secretamente, é comum utilizar a substituição do LSB (Least Significant Bit) como técnica para ocultar tais dados, permitindo que, além de uma fácil implementação, tal processo tenha alta capacidade. armazenar dados com segurança, podendo minimizar a dispersão entre o arquivo original e o arquivo gerado pelo processo de esteganografia [2]. Porém, a substituição LSB quando aplicada à esteganografia, tem suas restrições quanto ao tamanho do arquivo e/ou ao tipo de arquivo no qual tal informação será ocultada, pois dependendo do tamanho da informação a ser ocultada, o aparecimento de tal arquivo após processo pode ser viável [3].

Visando reduzir tal dispersão e uma melhor esquema-

Esta é uma versão revisada de um artigo de conferência apresentado em ARAUJO, Getúlio Dos Santos; SILVA, Warley Gramacho da. ESTEGANOGRAFIA USANDO ALGORITMO GENÉTICO DE CHAVES ALEATÓRIAS VICIADAS.. In: Anais do XVI Seminário de Iniciação Científica da Universidade Federal do Tocantins. Anais... Palmas (TO) DIGITAL (ONLINE), 2020. Disponível em: https://www.even3.com.br/anais/sicUFT/291845-ESTEGANOGRAFIA-USANDO-ALGORITMO-GENETICO-DE-CHAVES-ALEATORIAS-VICIADAS.
Dados de contato: Getulio dos Santos Araujo, getulio.santos@uft.edu.br

tização das substituições a serem realizadas, a utilização da Matriz de Substituição pode ser vista como uma forma eficiente, reversível e simples de mapear tais substituições, bastando estruturar quais e por qual determinada sequência LSB será substituído. Desta forma, a obtenção de uma matriz ótima permite, ao ocultar os dados em arquivos digitais, reduzir a sua degradação, bem como recuperar a informação oculta. No entanto, encontrar esta matriz ótima pode tornar este processo demorado, devido à abundância de matrizes possíveis, uma vez que para uma determinada sequência de LSBs de tamanho n, existem $n!$ matrizes diferentes, continuando assim os algoritmos heurísticos a encontrar matrizes de substituições ótimas ou quase ótimas, isso é visto com bons olhos.

Uma das heurísticas analisadas neste projeto de pesquisa é a utilização do Algoritmo Genético de Chave Aleatória Viciadas (BRKGA) proposto por [4], sendo uma variação dos Algoritmos Genéticos de Chave Aleatória (RKGA) de [5], ambos aplicados à esteganografia.

II. Algoritmo de Chaves Aleatória Viciadas (BRKGA)

Em geral, as etapas de um algoritmo genético podem consistir em etapas bem definidas. Essas etapas incluem gerar uma população inicial, iniciar o ciclo de geração realizando cruzamento entre indivíduos, simulando o processo evolutivo através da mutação dos genes dos indivíduos e submetendo-os à seleção para a próxima geração.

O Algoritmo Genético de Chave Aleatória Viciadas (BRKGA) é uma metaheurística baseada no Algoritmo Genético de Chaves Aleatórias (RKGA), que, por sua vez, é baseado no Algoritmo Genético. O sucesso do desempenho do RKGA na resolução de problemas de sequenciamento levou a comunidade científica a testá-lo em uma ampla gama de problemas combinatórios, onde as soluções podem ser representadas como vetores de permutação. Os genes desses vetores são chamados de chaves aleatórias, sendo essencialmente números reais gerados aleatoriamente no intervalo (0, 1].

O BRKGA, embora semelhante ao AG, possui uma peculiaridade. Enquanto o AG favorece indivíduos bem avaliados, no BRKGA, de indivíduos P, apenas a elite PE e os mutantes PM da geração atual avançam. Além disso, estão incluídos novos indivíduos $P - PE - PM$ provenientes de cruzamentos entre elites e não-elite.

III. Métodos

A substituição dos bits menos significativos nas imagens baseia-se na insignificância desses bits, tratados como ruído. Por exemplo, substituir o bit menos significativo (LSB) 1 em um byte de 00000000 (decimal 0) para 00000001 (decimal 1) faria pouca diferença em comparação com a substituição do bit mais significativo (MSB) para 10000000 (decimal 128).

Embora a substituição do LSB degrade ligeiramente a imagem e limite a capacidade, a substituição de vários bits por byte pode aumentar a carga, mas afeta gradualmente a imagem original. Um compromisso nos bits LSB é eficaz para carga e camuflagem.

Contudo, estas substituições devem permitir ao destinatário recuperar a mensagem original. O uso de uma matriz de substituição facilita a reversão das substituições, permitindo ao destinatário reconstruir a mensagem incorporada.

As substituições devem ser reversíveis para o destinatário recuperar a mensagem. O uso da matriz de substituição facilita esse processo, permitindo ao destinatário recuperar facilmente a mensagem incorporada. Matrizes de substituição por [1] utilizam um vetor de posições de tamanho N para LSBs a serem trocados, cada valor representando bits convertidos, únicos e entre 0 a $N-1$. As imagens coloridas podem ter vários vetores, melhorando a camuflagem e a qualidade. A diversidade de vetores representa características de cores, dificultando a quebra e melhorando a qualidade da esteganografia.

Nos testes foi adotado um cruzamento uniforme parametrizado de [6] com ajustes para o problema. Cabe à mutação inserir PM em indivíduos mutantes aleatórios para evitar convergência prematura.

O vetor de chaves aleatórias foi decodificado para viabilizar a solução para problemas combinatórios discretos. O processo de decodificação envolveu o mapeamento de chaves para elementos classificados, mantendo a associação classificando o vetor de chaves e a nova ordem de bits representando a solução. Outro decodificador utilizado consiste basicamente na discretização de valores vetoriais, o processo decodificador atribuirá a cada alelo do vetor resultante, um valor discreto de 0 a $N-1$, conforme a posição que tal alelo se encontra no vetor de chaves ordenadas.

Foi implementado um algoritmo genético para comparações, inspirado no trabalho de [1], com parâmetros ajustados para se assemelhar ao BRKGA, a fim de haver comparações justas das abordagens. Outros algoritmos foram implementados com base nas conclusões do pré-teste, como busca local em indivíduos mutantes, utilizando Hill Climbing e Simulated Annealing [7]. Essas abordagens foram utilizadas para compreender os resultados observados. Outra abordagem introduzida foi a combinação de conceitos de decodificadores e chaves aleatórias do BRKGA no algoritmo genético, alterando os processos de avaliação e a estrutura dos indivíduos para considerar valores discretos. Esta versão foi chamada $AG+RK$.

A função de avaliação foi o cálculo do $PSNR$ da imagem resultante utilizando a matriz.

Algumas das mesmas imagens de coberturas e secretas usadas por [1], e presentes e adaptadas de [8] foram utilizados nos teste realizados. Como imagens de coberturas elencamos as figuras 1 e 2, e como imagens secretas as figuras 3, 4 e 5.

IV. Resultados e Discussões

Independentemente das configurações dos parâmetros, o BRKGA produziu consistentemente melhores matrizes de substituição quando os parâmetros permitiram indiretamente que o processo de cruzamento, produza a maioria dos indivíduos, com uma população elite reduzida. Isto resultou em indivíduos progressivamente mais semelhantes às elites atuais, evitando a estagnação nos ótimos locais e promovendo a procura de melhores soluções. A discretização se destacou em relação aos demais decodificadores, embora algumas operações tenham piorado com o passar das

 ISSN: 2675-3588

Figura 1: Imagens de Cobertura Babuíno.

Figura 2: Imagens de Cobertura Lena.

Figura 3: Imagens de Secreta Jato.

gerações devido às características das imagens. O primeiro decodificador com cruzamento parametrizado assemelha-se ao cruzamento de pontos, enquanto o segundo com

Figura 4: Imagem secreta Veleiro no lago.

Figura 5: Imagem secreta Casa.

cruzamento proposto é mais parecido com o cruzamento uniforme, indicando nenhuma relação direta entre valores vizinhos no cromossomo.

A figura 6 compara o AG de [1] com a melhor configuração de parâmetros do BRKGA, usando as mesmas gerações e população.

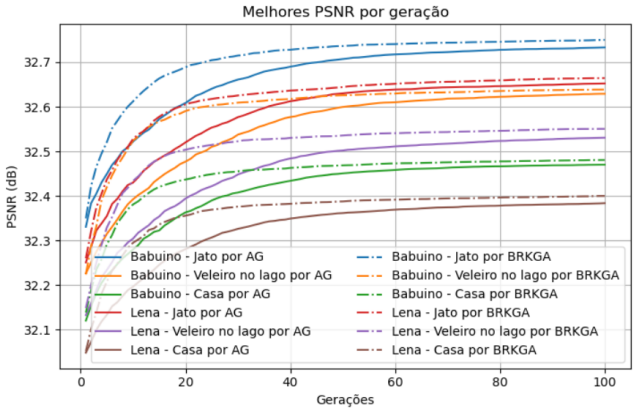

Figura 6: Comparação dos resultados obtidos entre o Algoritmo Genético abordado por [1] e BRKGA.

A abordagem com chaves aleatórias e o decodificador no BRKGA oferece a vantagem de permitir mudanças discretas nos indivíduos, como alterar um alelo para afetar sua representação no modelo discreto. Isso é notável nas operações de cruzamento, conforme mostrado na figura 7, onde a melhoria fica evidente ao aplicar esses conceitos ao algoritmo de [1]. Mas ainda assim não supera a abordagem do BRKGA, vide a figura 8.

Além do BRKGA, foi implementado um algoritmo genético baseado em configuração semelhante à do [1], assimilando ao máximo os parâmetros da abordagem proposta, em termos das operações realizadas, a fim de fazer uma avaliação justa comparação de ambas as abordagens.

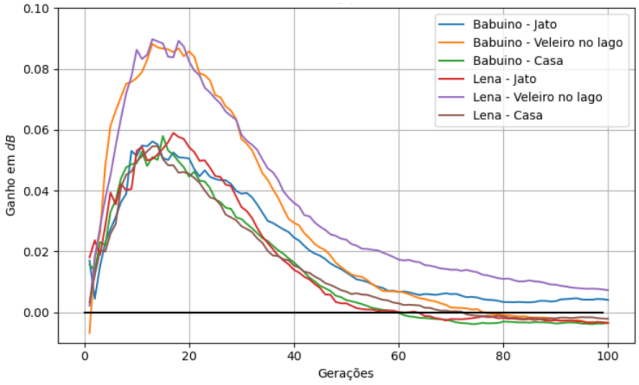

Figura 7: Ganho PSNR em dB de AG + RK comparado à abordagem AG simples

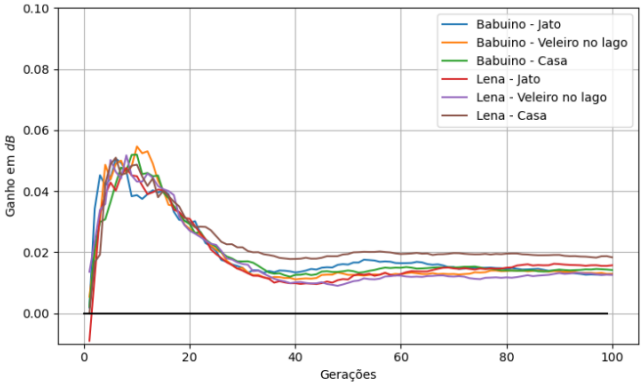

Figura 8: Ganho PSNR em dB do BRKGA comparado ao AG+RK

Algoritmos de busca local Hill Climbing e Simulated Annealing introduzidos como uma proposta de busca local no BRKGA. Tanto AG+RK quanto busca local em BRKGA e AG indicam ganhos relativos em dB em comparação com abordagens convencionais. A Figura 9 e 10 mostra que chaves aleatórias com decodificador melhoram a evolução populacional. Contudo, após a 50ª a 60ª geração, estes benefícios diminuem em ambas as abordagens.

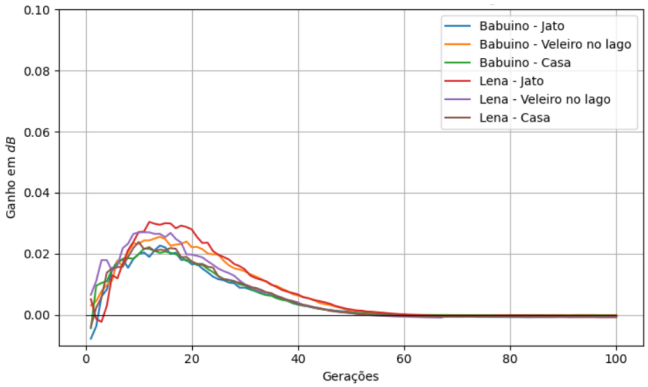

Figura 9: Ganho PSNR em dB aplicando busca local por Hill Climbing em BRKGA com relação ao AG

V. Conclusão

Em resumo, os conceitos BRKGA, como decodificador, vetor de chaves aleatórias, população elite e cruzamento parametrizado, oferecem melhorias nos resultados até um limite, condicionado pelos parâmetros escolhidos. Um

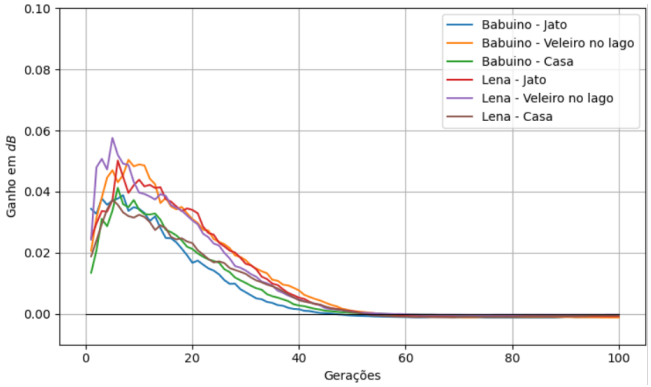

Figura 10: Ganho PSNR em dB aplicando busca local por Simulated Anneling em BRKGA com relação ao AG

estudo detalhado do problema em questão é crucial para a determinação desses parâmetros, conforme mostrado neste artigo.

VI. Agradecimentos

Este trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil.

Referências

[1] A. L. Brazil *et al.*, "Path relinking and aes cryptography in color image steganography," 2009.

[2] S. A. Laskar and K. Hemachandran, "High capacity data hiding using lsb steganography and encryption," *International Journal of Database Management Systems*, vol. 4, no. 6, p. 57, 2012.

[3] M. PETRI, "Esteganografia," *Sociedade Educacional De Santa Catarina–SOCIESC, Instituto Superior Tupy. Joinville, TCC de sistemas de informação*, 2004.

[4] J. F. Gonçalves and M. G. C. Resende, "Biased random-key genetic algorithms for combinatorial optimization," *Journal of Heuristics*, vol. 17, no. 5, pp. 487–525, Aug. 2010. [Online]. Available: https://doi.org/10.1007/s10732-010-9143-1

[5] J. C. Bean, "Genetic algorithms and random keys for sequencing and optimization," *ORSA Journal on Computing*, vol. 6, no. 2, pp. 154–160, May 1994. [Online]. Available: https://doi.org/10.1287/ijoc.6.2.154

[6] W. M. Spears, K. A. De Jong *et al.*, "On the virtues of parameterised uniform crossover." in *ICGA*, 1991, pp. 230–236.

[7] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, "Optimization by simulated annealing," *science*, vol. 220, no. 4598, pp. 671–680, 1983.

[8] A. Weber, "The usc-sipi image database volume 3: miscellaneous."

 ISSN: 2675-3588