Computação da Quadratura Gaussiana em um Esquema Criptográfico Parcialmente Homomórfico

Paulo Ricardo Reis[1], **Pedro Lara**[2], **Fábio Borges**[1]

[1]Laboratório Nacional de Computação Científica (LNCC)
25651-075 – Petrópolis – RJ – Brazil

[2]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ)
25620-003 – Petrópolis – RJ – Brazil

{paulorbr, borges}@lncc.br, pedrocslara@gmail.com

Abstract. *There is a growing attention to the use of homomorphic encryption, that is, cryptographic systems able to perform mathematical operations with the data in the encrypted domain. Although such systems provide a huge gain in terms of data privacy, they prove to be significantly slower. This work evaluates the applicability of the Paillier system in the calculation of gaussian quadrature. A mean increase o 3234 times in the computational time was verified using the homomorphic system and nine decimal cases of precision.*

Resumo. *É crescente a atenção ao uso de criptografia homomórfica, isto é, sistemas criptográficos capazes de realizar operações matemáticas com os dados no domínio cifrado. Embora tais sistemas proporcionem um enorme ganho no que se refere à privacidade dos dados, estes demonstram ser expressivamente mais lentos. Este trabalho avalia a aplicabilidade do sistema Paillier no cálculo da quadratura gaussiana. Foi constatado um aumento médio de 3234 vezes no tempo computacional utilizando o sistema homomórfico e precisão de nove casas decimais.*

1. Introdução

Com a crescente demanda por processamento computacional ao longo das últimas décadas, foram pensadas alternativas descentralizadas para o processamento de aplicações de grande porte como, por exemplo, a computação em nuvem. Uma empresa ou instituição de pesquisa não mais necessita possuir supercomputadores em sua sede para realizar cálculos de grande porte, podendo contratar serviços computacionais de terceiros para tal. Enquanto esta descentralização possa proporcionar economia de custos físicos e operacionais, traz consigo uma nova preocupação com a segurança e privacidade dos dados processados em nuvem.

A criptografia homomórfica é uma alternativa para mitigar este problema, já que permite o processamento de dados no domínio cifrado, isto é, não é necessária a decriptação para efetuar cálculos com os dados. Estes sistemas criptográficos são a chave para o processamento seguro de dados sensíveis em sistemas de terceiros, viabilizando a segurança e privacidade dos dados em nuvem como, por exemplo, armazenamento seguro [Pallavi e Joshi 2020], processamento de dados médicos e farmacêuticos [Hasan e Shaw 2018] [Bhagadia et al. 2020], entre outros. Entretanto, o uso destas técnicas pode tornar o código expressivamente mais lento.

Este trabalho avalia a aplicabilidade do criptossistema parcialmente homomórfico Paillier no cálculo de aproximações de integrais definidas através da quadratura gaussiana. Na próxima seção é feita uma breve descrição da regra de quadratura gaussiana. Na terceira seção é descrito o criptossistema de Paillier, bem como suas propriedades homomórficas. Posteriormente é descrito o processamento homomórfico da quadratura gaussiana. Na quinta seção são apresentados resultados experimentais para o caso proposto. Na última seção são tecidas considerações finais e futuras direções de pesquisa.

2. A Regra de Quadratura Gaussiana

No escopo da análise numérica, uma regra de quadratura é uma aproximação para a integral de uma função, normalmente por um somatório ponderado dos valores assumidos pela função em pontos específicos do domínio de integração. A quadratura gaussiana [Epperson 2013] aproxima a integral definida de uma função no intervalo $[-1, 1]$, assumindo a forma

$$\int_{-1}^{1} f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i), \tag{1}$$

onde n é a ordem da aproximação, ou seja, o número de pontos x_i utilizados e w_i são os pesos da quadratura. Os pontos x_i podendo ser obtidos de diferentes maneiras. Tomando como exemplo a quadratura de Gauss-Legendre, os pontos x_i são as raízes do n-ésimo polinômio de Legendre, $P_n(x)$.

Para aproximar integrais definidas em um intervalo genérico $[a, b]$, pode-se efetuar uma mudança de variáveis, de maneira que

$$\int_a^b f(x)\,dx = \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{b-a}{2}x' + \frac{a+b}{2}\right)\,dx'. \tag{2}$$

Aplicando-se a regra de quadratura gaussiana de ordem n, obtém-se

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2}\sum_{i=1}^{n} w_i f\left(\frac{b-a}{2}x_i + \frac{a+b}{2}\right). \tag{3}$$

Na próxima seção será descrito brevemente o algoritmo criptográfico parcialmente homomórfico de Paillier.

3. Criptografia Parcialmente Homomórfica com o Algoritmo de Paillier

Em 1999, Pascal Paillier propôs um algoritmo probabilístico assimétrico para criptografia de chave-pública cuja primitiva criptográfica é baseada no problema de se computar resíduos quadráticos. Tal problema não demonstra maior dificuldade de se resolver do que o problema do RSA [Paillier 1999].

A principal vantagem deste algoritmo reside no fato de este ser um criptossistema homomórfico aditivo. Isto significa dizer que é possível efetuar a soma de dois textos planos m_1 e m_2 no domínio cifrado. Isto é, de posse apenas da chave pública e das mensagens cifradas $E(m_1)$ e $E(m_2)$, pode-se calcular $E(m_1 + m_2)$. O criptossistema pode ser divido em três funções: geração de chaves, encriptação e decriptação, descritas no Algoritmo 1

Algoritmo 1: Criptossistema Paillier.

função `Geração de Chaves()`**:**
- $p, q \leftarrow$ números primos aleatórios de k bits
- $n \leftarrow p \cdot q$
- $\lambda \leftarrow$MMC$((p-1),(q-1))$
- **repita**
  - $g \leftarrow g' \in \mathbb{Z}^*_{n^2}$
  - $\mu \leftarrow L(g^\lambda \mod n^2)$ onde $L(x) = \frac{x-1}{n}$
- **até** $\exists \mu^{-1} \in \mathbb{Z}_n$;
- $\mu \leftarrow \mu^{-1} \in \mathbb{Z}_n$
- **retorna** *Chave pública: (n, g); chave privada: (λ, μ)*

fim

função `Encriptação(`$m \in \mathbb{Z}_n$, g, n`)`**:**
- $r \leftarrow r' \in \mathbb{Z}^*_n$ aleatório
- $c \leftarrow g^m \cdot r^n \mod n^2$
- **retorna** $c \in \mathbb{Z}^*_{n^2}$

fim

função `Decriptação(`$c \in \mathbb{Z}^*_{n^2}$, λ, μ`)`**:**
- $m \leftarrow L(c^\lambda \mod n^2) \cdot \mu \mod n$ onde $L(x) = \frac{x-1}{n}$
- **retorna** $m \in \mathbb{Z}_n$

fim

Como anteriormente mencionado, o criptossistema Paillier é homomórfico aditivo, resultando em duas propriedades. A primeira constata que o produto de dois textos cifrados será decifrado como a soma dos textos planos correspondentes, isto é,

$$D(E(m_1, r_1) \cdot E(m_2, r_2) \mod n^2) = m_1 + m_2 \mod n. \tag{4}$$

Já a segunda diz que um texto cifrado elevado a uma constante k será decifrado como o produto do texto plano pela constante, ou seja,

$$D(E(m_1, r_1)^k \mod n^2) = km_1 \mod n. \tag{5}$$

Na próxima seção será demonstrada uma maneira de se usar as propriedades (4) e (5) para processar homomorficamente o cálculo da quadratura gaussiana.

4. Processamento Homomórfico da Quadratura Gaussiana

O processamento homomórfico é feito na parte direita da equação (1). Os valores $f(x_i)$ são cifrados, enquanto os pesos w_i são mantidos planos. Sendo assim, a propriedade (5) é utilizada para efetuar a multiplicação $w_i \cdot E(f(x_i))$ e a propriedade (4) é utilizada para o cálculo do somatório, como segue:

$$D(E(f(x_i), r_1)^{w_i} \mod n^2) = w_i \cdot f(x_i) \mod n; \tag{6}$$

$$D(E(w_i f(x_i), r_1) \cdot E(w_{i+1} f(x_{i+1}), r_2) \mod n^2) = w_i f(x_i) + w_{i+1} f(x_{i+1}) \mod n.$$

Os cálculos com Paillier só podem ser realizados com inteiros positivos, devido à aritmética modular, entretanto, a quadratura gaussiana utiliza números decimais. Para compatibilizar estes sistemas foi utilizada uma representação de ponto fixo. Escolhendo-se uma precisão fixa de casas decimais j, um número decimal α pode ser representado pelo inteiro a, cujos primeiros dígitos serão a parte inteira de α e os j últimos dígitos serão os j dígitos fracionários de α. A Tabela 4 traz alguns exemplos de representação de ponto fixo ao variar o número de casas decimais.

j	3.7564	12.45	0.0013	-1.5677
2	375	1245	0	-156
3	3756	12450	1	-1567
4	37564	124500	13	-15677

Tabela 1. Representação de ponto fixo para dado número de casas decimais.

Devido à aritmética modular, números negativos serão também representados por números positivos. No domínio cifrado, todo número será representado como um valor positivo menor do que n^2. Valores negativos são representados por seus inversos multiplicativos em $\mathbb{Z}^*_{n^2}$. Assim, ao decifrar o resultado obtido para o somatório (1), dois valores serão possíveis, um positivo e um negativo. Esta é uma das desvantagens encontradas neste experimento. Entretanto, os valores obtidos divergem muito, sendo razoável a decisão a partir da modelagem do problema. Na próxima seção serão descritos alguns resultados experimentais obtidos com relação ao tempo de processamento.

5. Resultados Experimentais

Todos os códigos utilizados para a montagem experimental descrita nesta seção foram escritos em linguagem C, utilizando da biblioteca GMP (GNU Multiple Precision Arithmetic Library), que fornece uma representação otimizada de inteiros de precisão arbitrária além de implementação nativa para o cálculo de inversa modular e testes de primalidade. Para medir o tempo foi utilizada a função `omp_get_wtime()` da API OpenMP, para a linguagem C.

A fim de comparar o tempo de processamento do cálculo da quadratura gaussiana em modo plano e homomorficamente com Paillier, foram executados testes experimentais utilizando um sistema com um processador Intel i7 $-$ 6500U de 2.50 GHz quad-core, 16GB RAM e sistema operacional Ubuntu 18.04 LTS de 64-bits. Para os testes, foi arbitrariamente escolhido usar a quadratura de Gauss-Legendre para aproximar a integral definida $\int_{-5}^{5} e^x \cdot \cos(x)dx$.

Foram executados testes para aproximações de ordem 5 até 100, e utilizando precisão de casas decimais de 3, 6 e 9 casas. Na Figura 1 encontram-se os tempos médios de processamento obtidos para o cálculo com os valores planos.

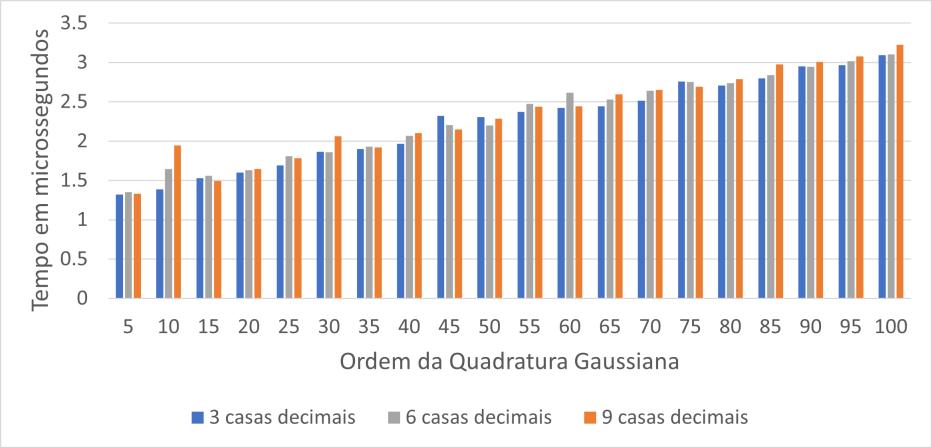

Figura 1. Tempo médio gasto para o cálculo em texto plano.

Na Figura 2 encontram-se os tempos médios de processamento obtidos para o cálculo homomórfico com os valores cifrados com Paillier.

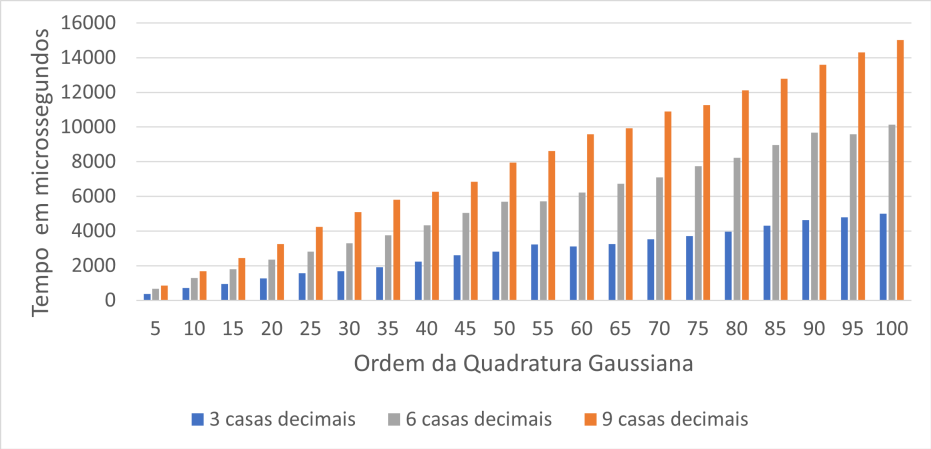

Figura 2. Tempo médio gasto para o cálculo em texto cifrado com Paillier.

Como esperado, o tempo de processamento têm uma tendência linear de crescimento conforme a ordem de aproximação utilizada. A fim de comparar as duas estratégias utilizadas, foi efetuado o cálculo da perda de eficiência do processamento homomórfico. Os valores foram obtidos fazendo-se a razão entre os tempos de processamento homomórfico e plano. A Figura 3 mostra a perda de eficiência ocasionada pelo uso do processamento homomórfico.

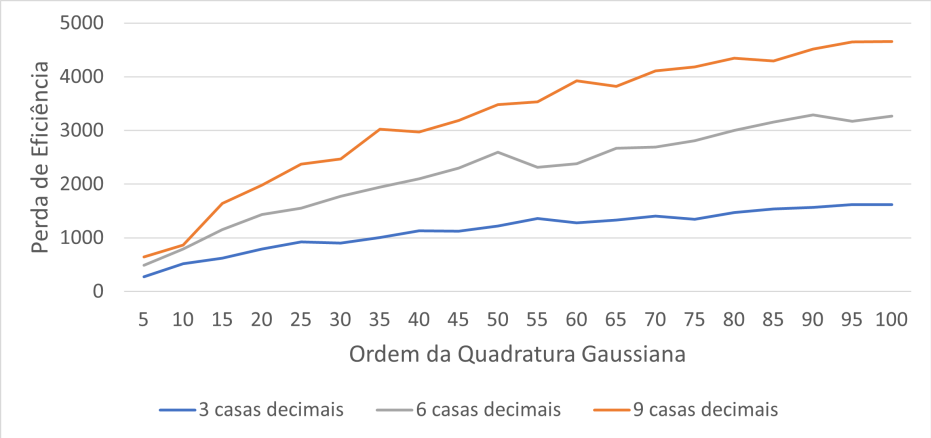

Figura 3. Perda de eficiência entre o processamento homomórfico e plano.

A análise dos dados e do gráfico permite inferir que a perda de eficiência ao se utilizar o processamento homomórfico cresce com o nível de precisão utilizado. Para três casas decimais o tempo de processamento homomórfico é, em média, aproximadamente 1153 vezes mais lento que o tempo de processamento do texto plano, subindo para o valor médio de 3234 usando precisão de nove casas decimais. Uma alternativa viável para melhorar o desempenho da criptografia com Paillier é o uso de técnicas de paralelização. Uma estratégia eficiente de paralelização para a exponenciação modular foi desenvolvida por [Borges et al. 2017]. Na seção seguinte serão feitas as considerações finais deste trabalho.

6. Considerações Finais

Esperava-se que o uso da criptografia homomórfica tornasse o tempo de processamento mais lento. Os resultados experimentais obtidos corroboram a grande perda de eficiência obtida ao se optar pelo processamento homomórfico dos dados, que chega a ser, em média, 3234 vezes mais lento que o processamento dos dados planos, em precisão de nove casas decimais. Isto forma um grande contraponto ao ganho em segurança e privacidade trazido pelo uso da criptografia homomórfica. Sempre cabe o princípio da razoabilidade para a tomada de decisão em se optar ou não pelo uso de uma ou outra técnica.

É importante ressaltar que os resultados foram obtidos sem o uso de qualquer otimização e para a mesma máquina. A grande vantagem da criptografia homomórfica é poder realizar o processamento de dados sensíveis em máquinas de terceiros. Para dispositivos com baixo poder computacional, como sensores e cartões inteligentes, esta pode ser uma estratégia viável, pois o uso de processamento em nuvem ou de máquinas com alto poder computacional pode viabilizar a computação homomórfica dos dados sem comprometer sua segurança e privacidade.

Objetiva-se futuramente otimizar o código utilizado, a fim de obter melhores resultados para o processamento homomórfico, além de reduzir a carga computacional inerente ao pré-processamento dos dados. Alternativas que utilizem o processamento paralelo podem se mostrar benéficas para o processamento com Paillier permitindo, também, a sua aplicação em métodos numéricos de maior carga computacional.

Agradecimentos

Agradecemos ao CNPq e à Petrobras pelo apoio na realização deste trabalho.

Referências

Bhagadia, D., Bhanpurawala, M., Dalal, D., and Kanani, P. (2020). Securing pharmaceutical data using homomorphic encryption. *International Journal of Future Generation Communication and Networking*, 13(1s):331–341.

Borges, F., Lara, P., and Portugal, R. (2017). Parallel algorithms for modular multi-exponentiation. *Applied Mathematics and Computation*, 292:406–416.

Epperson, J. F. (2013). *An introduction to numerical methods and analysis*. John Wiley & Sons.

Hasan, R. and Shaw, K. (2018). A survey paper on privacy of medical data storage on the cloud.

Paillier, P. (1999). Public-key cryptosystems based on composite degree residuosity classes. In *International conference on the theory and applications of cryptographic techniques*, pages 223–238. Springer.

Pallavi and Joshi, S. (2020). An efficient Paillier cryptographic technique for secure data storage on the cloud. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 145–149.