Estudo de aplicações e limitações de técnicas de compartilhamento de segredos no armazenamento de dados pessoais em ambientes de computação em nuvem

Barreto, Luciano[1]; Daoud, Maher[2]; Nascimento, Laisa Karina do[3]

Financiamento Edital 10/2023/CAN

1 Introdução

A computação em nuvem consolidou-se como um paradigma essencial para a tecnologia da informação, permitindo escalabilidade, flexibilidade e redução de custos com infraestrutura. Essa modalidade de processamento e armazenamento trouxe benefícios significativos, mas também ampliou os riscos relacionados à privacidade e segurança da informação. Quando dados pessoais são delegados a provedores de nuvem, surgem preocupações com acessos não autorizados, vazamentos e conformidade regulatória, especialmente no contexto da Lei Geral de Proteção de Dados (LGPD) no Brasil (BRASIL, 2018).

Nesse cenário, técnicas que vão além da criptografia tradicional tornam-se indispensáveis. Uma abordagem é o compartilhamento de segredos (Secret Sharing), inicialmente proposto por Shamir (1979), que consiste em dividir um dado sensível em múltiplas partes independentes, sendo necessária a combinação de um número mínimo para a reconstrução. Avanços posteriores, como o *PVSS - Publicly Verifiable Secret Sharing* (SCHOENMAKERS, 1999), introduziram verificações públicas para aumentar a confiabilidade, sobretudo em ambientes distribuídos. Essas técnicas ganham relevância diante do desafio de conciliar desempenho, segurança e privacidade na nuvem.

2 Objetivos

O trabalho teve como objetivo principal avaliar a aplicação de algoritmos de compartilhamento de segredos no armazenamento de dados pessoais em nuvem, analisando suas potencialidades e restrições. Buscou-se revisar a literatura sobre diferentes algoritmos (SHAMIR, 1979; SCHOENMAKERS, 1999; KRAWCZYK, 1993), implementar soluções adaptadas a ambientes simulados e realizar testes comparativos de desempenho. Além disso, pretendeu-se compreender de que forma esses mecanismos podem contribuir para a conformidade com legislações de proteção de dados, como a LGPD.

3 Metodologia

A pesquisa foi desenvolvida em duas etapas complementares. A primeira envolveu uma revisão teórica e normativa, englobando o estudo da LGPD (BRASIL, 2018) e a categorização de diferentes tipos de dados pessoais, além do levantamento das principais técnicas de secret sharing, com destaque para os trabalhos de Shamir (1979) e Schoenmakers (1999). Também foram considerados estudos mais recentes, como as propostas de Basu et al. (2019) e Vassantlal et al. (2022), que discutem a aplicação dessas técnicas em protocolos tolerantes a falhas bizantinas e serviços *multi-cloud*.

A segunda etapa concentrou-se na implementação prática. Bibliotecas em Java foram refatoradas para Python, sobre as quais foi construído um web service REST em Django,

[1]Servidor Docente IFSC - PHB - luciano.barreto@ifsc.edu.br
[2]Egresso do curso de ADS IFSC - CAN - maher.d@aluno.ifsc.edu.br
[3]Egressa do curso de ADS IFSC - CAN - laisa.kn@aluno.ifsc.edu.br

adequado para execução em ambiente de nuvem. Os *datasets* utilizados foram gerados com a ferramenta *Mockaroo*, contendo informações fictícias como nome, CPF e dados clínicos, armazenados no formato JSON. As operações de *split* (divisão) e *reconstruct* (recomposição) foram testadas em ambiente controlado utilizando um servidor Dell PowerEdge R250. Os tempos de resposta foram medidos em milissegundos, e o desempenho foi avaliado considerando CPU, memória e *I/O*, assegurando que os resultados refletissem os algoritmos e não limitações da infraestrutura.

4 Discussão e Resultados

Os testes realizados possibilitaram observar diferenças significativas entre os algoritmos de Shamir (1979) e o PVSS proposto por Schoenmakers (1999), tanto nas operações de divisão quanto nas de reconstrução.

Na operação de divisão do segredo (*split*), o algoritmo de Shamir apresentou desempenho superior, sobretudo em *datasets* menores. Para um conjunto de dados de *1kB*, o tempo médio de divisão foi de aproximadamente *83,56 ms*, enquanto o PVSS obteve valores próximos de *210 ms* nas mesmas condições. Esse padrão se manteve mesmo com o aumento do tamanho dos *datasets*: quando a carga foi elevada para *10kB*, Shamir atingiu cerca de *210 ms*, enquanto o PVSS manteve-se em patamares superiores, revelando maior custo computacional. Essa diferença confirma a eficiência do esquema de Shamir em ambientes de baixa latência, mas também evidencia sua variabilidade em cenários de maior volume de dados.

Já na operação de reconstrução do segredo (*reconstruct*), os dois algoritmos apresentaram resultados mais equilibrados, mas com características distintas. O Shamir manteve tempos médios próximos a *79-85 ms* para datasets pequenos, crescendo proporcionalmente com o aumento do tamanho dos arquivos. O PVSS, embora mais lento em comparação, apresentou maior consistência e previsibilidade, com tempos relativamente estáveis mesmo diante de variações de rede. Essa estabilidade sugere que, em aplicações críticas que demandam previsibilidade na recuperação de dados, o PVSS pode ser mais adequado, ainda que sacrifique velocidade.

Além dos tempos médios, a análise mostrou que os parâmetros *n* (número total de partes) e *k* (número mínimo necessário para reconstrução) influenciam diretamente no desempenho. Em cenários com *n* elevado, a divisão exigiu maior esforço computacional e ampliou os tempos de resposta. Essa relação já havia sido destacada em trabalhos anteriores (KRAWCZYK, 1993), e foi confirmada experimentalmente neste estudo.

Outro aspecto relevante foi o impacto da latência de rede, que em alguns casos elevou os tempos de resposta para patamares próximos a *230 ms*, independentemente do algoritmo utilizado. Esse resultado reforça que, além da escolha do método de compartilhamento, a qualidade da infraestrutura de rede da nuvem desempenha papel crucial no desempenho percebido pelo usuário.

5 Conclusão

Os resultados obtidos permitem concluir que técnicas de compartilhamento de segredos representam uma alternativa para o armazenamento seguro de dados pessoais em nuvem, especialmente diante das exigências regulatórias da LGPD (BRASIL, 2018). A operação de reconstrução mostrou-se mais eficiente que a de divisão, o algoritmo de Shamir apresentou vantagem em cenários de baixa latência, enquanto o PVSS se destacou pela consistência em ambientes distribuídos.

Assim, a escolha do algoritmo deve ser contextualizada, equilibrando rapidez, previsibilidade e robustez. A disponibilização do código em repositório público amplia a relevância prática da pesquisa, favorecendo a replicação dos testes por outras instituições.

Para trabalhos futuros, recomenda-se a exploração de algoritmos mais recentes e a realização de testes em cenários multi-cloud, conforme sugerido por Vassantlal et al. (2022).

Referências

BASU, S. et al. **Efficient verifiable secret sharing with share recovery in BFT protocols**. Proceedings of the ACM CCS, 2019.

BRASIL. **Lei nº 13.709**, de 14 de agosto de 2018. Lei Geral de Proteção de Dados. Disponível em: http://www.planalto.gov.br. Acesso em 30 julho 2024.

KRAWCZYK, H. **Secret sharing made short**. Annual International Cryptology Conference, 1993.

SCHOENMAKERS, B. **A simple publicly verifiable secret sharing scheme and its application to electronic voting**. Advances in Cryptology – CRYPTO99, 1999.

SHAMIR, A. **How to share a secret**. Communications of the ACM, 22(11), 1979.

VASSANTLAL, R. et al. **Cobra: Dynamic proactive secret sharing for confidential BFT services**. IEEE Symposium on Security and Privacy, 2022.