Um Sistema Seguro de Comercialização de Dados Pessoais Sensíveis baseado em Reputação, Confiança e Corrente de Blocos

Resumo. *Todo cidadão tem direito a privacidade e, portanto, direito a decidir sobre suas informações pessoais sem a interferência de organizações. Ele pode informar e comercializar seus dados, definindo quem, quando e onde as suas informações estarão disponíveis. No entanto, sistemas atuais de comercialização de dados centralizam a confiança, restringindo o controle do usuário proprietário sobre os próprios dados. Este artigo propõe um sistema seguro, ágil e eficaz para a comercialização de dados automatizada, distribuída e transparente entre domínios utilizando corrente de blocos, contratos inteligentes, reputação e confiança. Um protótipo desenvolvido e implementado do sistema de reputação comprova a eficácia de um modelo de confiança baseado em interações da vida real. Os resultados de avaliação de desempenho mostram que o sistema proposto provê segurança e privacidade na comercialização de dados entre domínios de maneira distribuída, ágil e eficaz, atingindo taxas de centenas de transações por segundo e punindo comportamento malicioso.*

1. Introdução

O controle de acesso a dados pessoais está se tornando uma grande preocupação à medida que os dispositivos delegam cada vez mais o armazenamento de dados confidenciais às autoridades centralizadas baseadas na nuvem [Lodderstedt et al. 2013]. Essas autoridades, como governos e empresas, podem então acessar, controlar e compartilhar dados pessoais do usuário com terceiros. Para o proprietário dos dados, usar soluções centralizadas de armazenamento implica perder o controle dos dados pessoais, pagar altas taxas e assinar termos que comprometem a privacidade de dados. Além disso, mesmo que a autoridade centralizada se comporte de maneira legítima, usuários mal-intencionados podem comprometer os serviços baseados na nuvem por meio de ataques internos e ataques de negação de serviço (*Denial of Service* - DoS). Uma maneira mais eficiente de garantir segurança e privacidade preservando o controle do proprietário sobre os dados usa a tecnologia de corrente de blocos [Nakamoto 2008], que fornece um registro distribuído e auditável para registrar transferências de dados pessoais.

Em um artigo precedente [Anônimo 2020], os autores propuseram um sistema baseado em corrente de blocos para comercializar dados de forma segura, automática e distribuída. O sistema se serve das propriedades de integridade e auditabilidade da corrente de blocos para armazenar as permissões de acesso de cada usuário aos dados privados e sensíveis. O cenário considera grandes centros de dados interconectados através da tecnologia de redes definidas por software (*Software Defined Network* SDN), que é usada para implementar o controle de acesso aos dados. Apesar de manter a imutabilidade dos registros da transferência de ativos, a corrente de blocos não garante a entrega de

dados armazenados fora da corrente (*off-chain*) e tampouco a qualidade deles. Vendedores maliciosos podem aproveitar essa vulnerabilidade para anunciar dados falsos ou não entregar dados adquiridos por compradores honestos. Assim, o sistema não consegue impedir comportamentos maliciosos de vendedores causando prejuízos a compradores que executam ações honestas.

Um sistema de confiança e reputação (*Trust and Reputation System* - TRS) permite identificar comportamentos maliciosos, além de apresentar uma visão sobre a qualidade dos dados a partir de avaliações de compradores. Um sistema de reputação integrado com a corrente de blocos pode prover avaliações da reputação de vendedores de dados assim como da qualidade dos dados comercializados de maneira transparente e distribuída. Desta forma, compradores podem rapidamente verificar a reputação de um anúncio e decidir por adquirir ou não os dados.

Este artigo propõe um sistema seguro de comercialização de dados baseado em corrente de blocos, reputação e confiança. As principais contribuições são as seguintes:

- criação de um ambiente de comercialização de dados de maneira automática e distribuída entre domínios, em que proprietários e compradores podem anunciar e adquirir dados através de uma corrente de blocos. A corrente de blocos registra de maneira imutável as transferências, permitindo que o proprietário mantenha controle sobre quem possui acesso aos seus dados. O esquema de comercialização proposto é de forma automática através de contratos inteligentes;
- um sistema de reputação e confiança que considera o histórico de interações e a opinião dos participantes para construir a reputação de um vendedor. O sistema adapta o modelo ao cenário de comercialização de dados, introduzindo uma punição na reputação proporcional ao preço do produto anunciado;
- um modelo de comercialização eficiente para anúncio, compra e avaliação da qualidade dos dados baseado em transações na corrente de blocos. As transações incluem a avaliação do vendedor e da qualidade dos dados comercializados;
- concepção do modelo de transação e de confiança e reputação.
  Resultados de avaliação de desempenho de um protótipo desenvolvido mostram que o sistema proposto é eficaz na punição de comportamento malicioso dos vendedores.

O restante do artigo está organizado da seguinte forma. A Seção 2 discute os trabalhos relacionados. A Seção 3 detalha o modelo de atacante para as entidades envolvidas na comercialização dos dados. A Seção 4 apresenta os modelos de confiança e reputação propostos e a integração desses modelos com a corrente de blocos. A Seção5 descreve os procedimentos necessários para integrar a confiança e reputação na corrente de blocos. A Seção 6 avalia o desempenho do sistema proposto por simulações e discute os resultados. Por fim, a Seção 7 conclui o artigo e apresenta direções para trabalhos futuros.

2. Trabalhos Relacionados

O emprego de sistemas de reputação e confiança para prover segurança em redes de computadores é usual. Velloso *et al.* propõem um modelo de confiança baseada em interações humanas para estabelecer confiança entre nós de uma rede ad hoc [Velloso et al. 2010]. Oliveira *et al.* propõem o uso de confiança para mitigar ataques Sybil no controle de acesso a dispositivos de Internet das coisas [de Oliveira et al. 2019].

Kamvar *et al.* propõem o algoritmo Eigentrust, que delega um valor de confiança global a cada participante de um sistema de compartilhamento de arquivos par-a-par (*Peer-to-Peer* - P2P) [Kamvar et al. 2003]. Sun *et al.* apresentam mecanismos de defesa contra ataques a sistemas de reputação de confiança em uma rede móvel ad hooc (*Mobile Ad Hoc Network* - MANET) [Sun et al. 2008].

A tecnologia de corrente de blocos provê auditabilidade e rastreabilidade em ambientes sem confiança mútua [Rebello et al. 2019, Michelin et al. 2018].

Outros trabalhos investigam o uso de reputação aplicados a corrente de blocos. Oliveira *et al.* propõem um mecanismo de consenso para corrente de blocos baseado em reputação (*Blockchain Reputation-Based Consensus* - BRBC) [de Oliveira et al. 2020]. Dennis e Owen propõem um sistema de reputação para transferência de arquivos em que as recomendações dos usuários são armazenadas na corrente de blocos de forma pública [Dennis and Owen 2015]. Os autores, no entanto, não oferecem uma solução ao ataque de dissimulação e não implementam a proposta. Malik *et al.* propõem um arcabouço para gerenciar a confiança de participantes de uma corrente de blocos que registra transferências em uma cadeia de suprimentos [Malik et al. 2019]. Compradores, autoridades ficais e sensores atestam a qualidade da mercadoria emitindo avaliações através de transações na corrente de blocos. Um contrato inteligente calcula a reputação de um comerciante (*trader*) através de uma soma ponderada das avaliações das entidades. Malik *et al.* limitam a proposta ao caso de uma cadeia de suprimentos e não consideram o preço da mercadoria no cálculo da reputação. Ainda, a proposta é suscetível a ataques de dissimulação (*on-off attack*), em que um usuário altera o seu comportamento entre bom e mau, danificando a rede sem ser detectado.

Putra *et al.* propõem um sistema de gerenciamento de confiança e reputação para sistemas baseados em corrente de blocos para controle de acesso a dispositivos IoT [Putra et al. 2020]. O sistema utiliza contratos inteligentes para avaliar a confiança e reputação dos nós, detectando e eliminando nós maliciosos da rede. Os autores utilizam a reputação como um atributo para controlar o acesso a um dispositivo, definindo uma reputação mínima que um participante deve ter para garantir acesso aos dados. Entretanto, a proposta é suscetível a ataques de dissimulação e não recompensa os proprietários dados. A comercialização de dados pelos proprietários é uma propriedade desejável.

Ao contrário dos artigos citados, este artigo propõe um sistema de reputação e confiança baseado em corrente de blocos que é eficaz ao punir comportamento malicioso e adaptado à comercialização de dados baseada em corrente de blocos. Este artigo considera e adapta conceitos e mecanismos de defesa dos trabalhos citados para criar um sistema de reputação e confiança que considera as propriedades da corrente de blocos e de ambientes distribuídos. O sistema de comercialização como um todo é seguro, ágil e automático, remunerando os vendedores que comercializam os seus dados pessoais e punindo o vendedor malicioso.

3. Modelo de Atacante

Ataques ao sistema de confiança e reputação representam a tentativa de um nó malicioso aumentar a própria reputação ou prejudicar a reputação de nós honestos. O

Uma versão em inglês baseada neste artigo intitulada ”xxxxxxxxxxxxx”deve ser submetida a congressos ou revistas internacionais.

artigo considera cinco tipos de ataque: (i) de maldizer (*bad mouthing attack*); (ii) de dissimulação (*on-off attack*); (iii) Sybil; (iv) do recém-chegado (*newcommer attack*); e (v) do comportamento conflitante (*conflicting behaviour attack*). A Tabela 1 apresenta cada ataque e as contramedidas propostas.

Ataques	Descrição	Contramedidas
Ataque de maldizer (bad mouthing attack)	Usuários maliciosos publicam falsos relatos sobre um usuário alvo para afetar a reputação do alvo. Esse ataque é mais poderoso quando há o conluio entre nós maliciosos.	Utilização de uma sinalização de insatisfação por um vendedor insatisfeito com uma avaliação. O sistema registra os pedidos e detecta os nós maliciosos.
Ataque de dissimulação (on-off atack)	Um usuário alterna o seu comportamento entre bom e mau para dissimular ações maliciosas sem ser detectado.	Utilização de um fator de esquecimento adaptativo no cálculo da reputação.
Ataque Sybil	Um atacante gera múltiplas identidades falsas para aumentar sua influência no sistema. Um atacante pode usar as identidades para efetuar um ataque de maldizer sobre um usuário alvo.	Utilização de uma corrente de blocos permissionada e imposição de regras por organização sobre vendedores.
Ataque de recém-chegado (newcomer attack)	Um atacante gera novas identidades repetidamente e simula ser um novo usuário, para continuar agindo de maneira maliciosa sem ser punido.	Utilização de uma corrente de blocos permissionada e imposição de regras por organização sobre vendedores.
Ataque de comportamento conflitante (conflicting behaviour attack)	Um atacante se comporta de maneira diferente com cada vizinho, gerando recomendações conflitantes em participantes honestos.	Todas as avaliações dos usuários são publicamente verificáveis como transações na corrente de blocos.

Tabela 1. Ataques e contramedidas propostas para o sistema de reputação.

Ataques à corrente de blocos objetivam impedir que uma transação ou bloco legítimo sejam adicionados à corrente de blocos. A propriedade de tolerância a falhas do consenso requer que o participante controle a maioria das organizações para afetar o protocolo de consenso, mitigando esse tipo de ataque. Além disso, o registro imutável e distribuído da corrente de blocos permite que os emissores de transações verifiquem se a proposta foi corretamente adicionada à estrutura.

Ataques a vendedores ou a compradores consistem em tentar obter dados privados e sensíveis anunciados ou personificar o alvo. O trabalho considera que todo dado anunciado na corrente de blocos é encriptado, i.e., se o atacante ganhar acesso, ele precisa da chave que decripta os dados para obter informações pessoais sobre o vendedor. Um atacante pode tentar personificar o alvo para enganar outros participantes. Este ataque, no entanto, não é possível, pois os emissores das mensagens assinam todas as transações. Além disso, a corrente de blocos registra todas as tentativas de modificação usando par roubados de chaves, permitindo que a vítima substitua o par de chaves e evite mais danos.

Ataques à rede tentam de isolar um único alvo, impedindo assim que vendedores e compradores emitam transações. Este artigo mitiga esse tipo de ataque estabelecendo caminhos redundantes entre os participantes da rede de corrente de blocos.

4. O Sistema de Confiança e Reputação Proposto

Este artigo define confiança e reputação como conceitos distintos em um sistema distribuído. A confiança é uma visão subjetiva de um comprador em relação a um vendedor com base em suas interações anteriores. Portanto, cada comprador calcula e atualiza de forma independente seus valores de confiança em um vendedor. A reputação representa uma visão global do sistema em relação a um vendedor específico

Da mesma forma que em Velloso *et al.* [Velloso et al. 2010], Putra *et al.* [Putra et al. 2020] e Malik *et al.* [Malik et al. 2019].

e deve ser construída agregando as confianças individuais de todos os compradores no sistema. O sistema, no entanto, estende a definição de confiança/reputação subdividindo-a em confiança/reputação de um vendedor e confiança/reputação dos dados anunciados. A subdivisão proposta objetiva regular as reputações no sistema de forma mais justa. O desacoplamento evita casos em que um bom vendedor comercializa dados de baixa qualidade sem ser punido, ou em que um vendedor ruim torna-se impossibilitado de recuperar sua reputação apesar de comercializar dados de alta qualidade. Assim, a reputação total de um anúncio é uma soma ponderada entre a reputação do vendedor e dos dados expressa por:

$$Rep_{an} = \alpha Rep_v + (1 - \alpha) Rep_d, \tag{1}$$

onde α é um parâmetro de priorização da reputação do vendedor ou da qualidade dos dados, ajustado em cada cenário diferente e acordado pelos participantes na inicialização da rede, Rep_v é a reputação do vendedor e Rep_d é a reputação dos dados.

4.1. Confiança do Vendedor

A ideia-chave do modelo de confiança proposto é simular as interações sociais da vida real, nas quais a confiança aumenta gradualmente com experiências positivas e diminui significativamente sempre que ocorre uma experiência negativa. Na confiança do vendedor, o modelo ainda considera que as interações recentes são mais relevantes que as interações do passado para permitir possíveis alterações no comportamento de um vendedor. O modelo utiliza uma função de envelhecimento adaptativa para implementar o esquecimento gradual de interações passadas, expressa por:

$$I_n = \sum_{i=1}^{n} \beta^{(n-i)} \delta_i, \tag{2}$$

onde β é o fator de esquecimento, n é o total de interações ocorridas e δ_i é o valor associado a cada interação. Se a interação i for positiva, $\delta_i = \delta_+ > 0$. Caso contrário, $\delta_i = \delta_- < 0$. Utilizando $|\delta_-| \gg |\delta_+|$, as interações negativas pesam mais do que as interações positivas e o modelo é capaz de simular a confiança igual a da vida real.

Diferente de trabalhos anteriores que utilizam fatores de esquecimento fixo [Malik et al. 2019, Putra et al. 2020], este trabalho adota um fator de esquecimento que se adapta de acordo com a probabilidade de o vendedor agir honestamente [Sun et al. 2008]. A probabilidade de que um vendedor seja honesto pode ser modelada por uma distribuição β com probabilidade *a priori* de $\frac{1}{2}$, que corresponde à probabilidade assumida pelo sistema de que um vendedor desconhecido ser honesto. A cada nova interação, o modelo atualiza a distribuição beta através de inferência bayesiana e utiliza o valor esperado $E[p] = \frac{\delta_+ + 1}{\delta_+ + \delta_- + 2} = \beta$ da nova distribuição para estimar a probabilidade de o vendedor seja honesto naquele momento. A principal vantagem do uso de um fator de esquecimento adaptável é a prevenção contra ataques de dissimulação, nos quais um mau vendedor comporta-se bem apenas o suficiente para recuperar sua reputação e voltar a agir maliciosamente. Com o fator de esquecimento adaptativo, se a probabilidade do vendedor ser honesto é alta, i.e. $E[p] \to 1$, o sistema leva mais tempo para esquecer o

passado do vendedor, recompensando-o por suas boas ações. Se o vendedor age maliciosamente, i.e. $E[p] \to 0$, suas ações passadas são esquecidas rapidamente e o peso de suas ações maliciosas recentes sobre a reputação é maior. Após calcular I_n, o modelo calcula o crescimento da confiança de um comprador i em um vendedor j através de uma função de Gompertz, expressa por:

$$s_{ij} = a \exp(-b \exp(-cI_n)), \tag{3}$$

onde a, b e c são constantes que representam a assíntota, o parâmetro de deslocamento ao longo do eixo x e a taxa de crescimento da confiança, respectivamente. A escolha da função de Gompertz se deve a característica do seu aumento gradual, simulando a confiança na vida real, e ao fácil ajuste do seu formato através das constantes a, b e c.

4.2. Reputação do Vendedor

O modelo proposto para reputação do vendedor agrega os valores locais de confiança inspirando-se no algoritmo Eigentrust [Kamvar et al. 2003]. O algoritmo Eigentrust é um dos algoritmos mais citados e utilizados em sistemas par-a-par (*peer-to-peer* - P2P) para prover uma forma eficiente de construir reputação em ambientes distribuídos. A proposta original do Eigentrust baseia-se na confiança local normalizada de um par *i* em um par *j* conhecido, expressa por:

$$c_{ij} = \frac{max(s_{ij}, 0)}{\sum_j max(s_{ij}, 0)}, \tag{4}$$

onde $s_{ij} = sat(i, j) - unsat(i, j)$ é a diferença entre interações satisfatórias e interações insatisfatórias entre os pares envolvidos. A normalização transforma os valores de confiança em uma distribuição de probabilidade que impede que um par malicioso emita valores arbitrariamente altos de confiança para outros pares maliciosos. O algoritmo então propõe que a maneira natural de um par i descobrir a reputação de um par k desconhecido é perguntar a seus conhecidos as suas opiniões sobre k. As opiniões dos conhecidos possuem pesos proporcionais à confiança que o par *i* possui em cada conhecido:

$$t_{ik} = \sum_j c_{ij} c_{jk} = c_{i1} c_{1k} + c_{i2} c_{2k} + ... + c_{in} c_{nk}, \tag{5}$$

onde t_{ik} representa a confiança que o par *i* possuirá sobre o par *k* com base nas opiniões de seus conhecidos.

O modelo proposto neste artigo substitui, sem perda de generalidade, a equação de s_{ij} original pela função de Gompertz proposta em (3). Os pares i e j que perguntam opiniões equivalem aos compradores que desejam interagir com um vendedor k. Generalizando a Equação 5 na notação matricial para todo comprador i, tem-se:

$$C = \begin{bmatrix} c_{11} & c_{12} & ... & c_{1n} \\ c_{21} & c_{22} & ... & c_{2n} \\ ... & ... & ... & ... \\ c_{n1} & c_{n2} & ... & c_{nn} \end{bmatrix}, \vec{c_{ik}} = \begin{bmatrix} c_{1k} \\ c_{2k} \\ ... \\ c_{nk} \end{bmatrix}, \vec{t_{ik}} = \begin{bmatrix} t_{1k} \\ t_{2k} \\ ... \\ t_{nk} \end{bmatrix} \tag{6}$$

e $\vec{t_{ik}} = C^T.\vec{c_{ik}}$. No entanto, a confiança do comprador *i* no vendedor k ainda reflete apenas sua própria experiência e a de seus conhecidos. Para obter uma visão mais ampla, o comprador *i* pode perguntar aos conhecidos dos seus conhecidos ($\vec{t_{ik}} = (C^T)^2.\vec{c_{ik}}$). Ao repetir o processo repetidas vezes, o comprador adquire uma visão completa da rede após uma quantidade n suficiente de iterações:

$$\vec{t_{ik}} = (C^T)^n.\vec{c_{ik}} \tag{7}$$

Destaca-se que $\vec{c_{ik}}$, C e $\vec{t_{ik}}$ correspondem, respectivamente, ao estado inicial, à matriz de probabilidades e ao estado atual de uma cadeia de Markov estacionária. Portanto, se n for grande o suficiente e se a matriz C é irredutível e aperiódica, o estado atual $\vec{t_{ik}}$ de todos os compradores i converge para o estado estacionário $\vec{r_{ik}} = \begin{bmatrix} r_{1k} & r_{2k} & ... & r_{nk} \end{bmatrix}^T$ independente da confiança inicial $\vec{c_{ik}}$ de cada par. Como $r_{1k} = r_{2k} = ... = r_{nk}$, pode-se concluir que qualquer elemento do vetor $\vec{r_{ik}}$ representa a reputação Rep_k do vendedor k vista pelo sistema como um todo. Assim, define-se um vetor global de reputação $\vec{r} = \begin{bmatrix} Rep_1 & Rep_2 & ... & Rep_n \end{bmatrix}^T$ que contém a reputação de todos os vendedores do sistema e que deve ser armazenado no estado global da corrente de blocos.

4.3. Reputação dos Dados

A reputação dos dados no sistema proposto é simples por duas razões. Primeiro, considera-se que os dados não são modificados após sua publicação porque a transação de anúncio contém um resumo (*hash*) dos dados. Assim, um comprador detecta facilmente no caso de um vendedor malicioso modificar os dados anunciados e pode puni-lo por meio de uma avaliação negativa. Segundo, o artigo considera que um usuário nunca compra os mesmos dados duas vezes. Assim, em vez de usar valores de confiança que se modificam constantemente e um fator de esquecimento, é possível definir uma função baseada no número de compradores que já compraram os dados:

$$E_n = \sum_{i=1}^{n} \delta_i \ln(N_c) \frac{Pr}{\sum_{d=m-j}^{m} Pr_d}, \tag{8}$$

onde m é o total de anúncios, Pr é o preço dos dados normalizado em uma janela dos preços mais recentes do sistema, j é o tamanho da janela e N_c é o número de usuários que já compraram o dados. O fator $\ln(N_c)$ evita que um vendedor ganhe reputação interagindo somente com um comprador. O modelo normaliza o preço para determinar seu valor real em comparação com o preço médio no sistema. Dados de valor alto possuem maior risco associado e, portanto, sua reputação deve aumentar e diminuir proporcionalmente ao risco. O uso de uma janela evita que a reputação dos dados diminua naturalmente à medida que mais dados são anunciados e que a soma dos preços no sistema aumenta. A etapa final para construir a reputação dos dados é introduzir E_n na função Gompertz:

$$Rep_d = a \exp(-b \exp(-cE_n)) \tag{9}$$

5. Integrando Confiança e Reputação na Corrente de Blocos

A corrente de blocos tem a mesma estrutura da que os autores usaram em XYZ *et al.* [Anônimo 2020] com a definição de três tipos de transação para a comercialização dos

dados: i) transação de anúncio; ii) transação de compra e iii) transação de resposta. Para introduzir o sistema de reputação e confiança, o artigo propõe uma transação de avaliação além das três anteriores. Transações de anúncio são emitidas a partir de um proprietário interessado em disponibilizar e comercializar seus dados. O proprietário submete os dados a serem comercializados em um servidor de armazenamento capaz de suportar e processar grande quantidade de dados [Truong et al. 2019] e emite uma transação de anúncio assinada. As transações são assinadas pelo emissor usando criptografia assimétrica para garantir a autenticidade e a integridade da transação. A transação de anúncio deve conter uma breve descrição dos tipos de dados sendo ofertados, por exemplo dados de sensores médicos, e o preço para os dados serem adquiridos. Ainda, a transação registra o resumo (*hash*) do dado anunciado para manter a integridade dele.

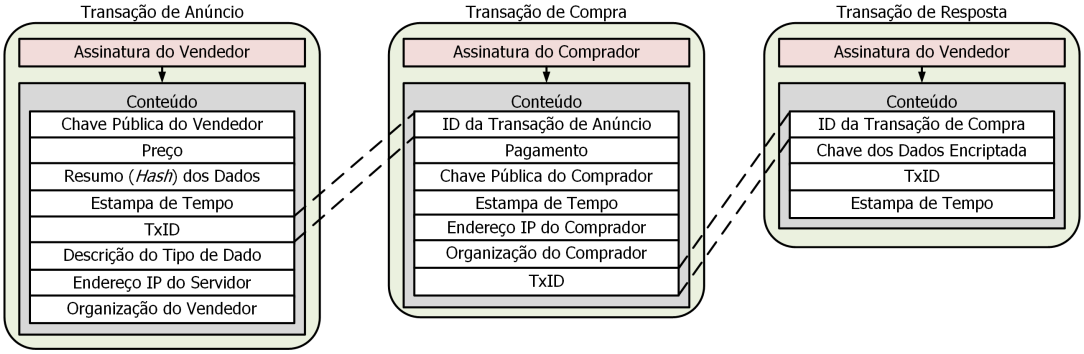


Figura 1. Modelos propostos de transações para a comercialização segura de dados. Compradores referenciam à transação de anúncio usando o identificador da transação. Após efetuar uma compra, o comprador deve ler a transação de reposta para obter a chave que decripta os dados.

Transações de compra são emitidas a partir de interessados em adquirir dados anunciados na corrente de blocos. Os clientes compradores procuram por dados na corrente de blocos a partir de buscas por transações de anúncio. O sistema permite buscas específicas, como buscas pelo tipo de dado, ou buscas por todas as transações de anúncio. O interessado em comprar os dados anunciados deve emitir uma transação de compra referenciando o identificador da transação de anúncio correspondente e informando o endereço IP do comprador para que um controlador SDN possa liberar o acesso. A transação de compra também deve incluir o valor a ser pago pelos dados. Caso o valor oferecido seja menor que o constante na transação de anúncio correspondente ou o comprador não tenha saldo suficiente para a compra, a transação não é válida.

Transações de resposta são emitidas pelo proprietário automaticamente após um de seus dados receber uma transação de compra. A transação de resposta envia ao comprador a chave que decripta os dados. A chave é criptografada usando a chave pública do comprador para garantir que o servidor de armazenamento não tenha acesso aos dados decriptados e compartilhe com terceiros. A Figura 1 ilustra o modelo dos três tipos propostos de transação para a comercialização efetiva e segura de dados pessoais.

Um contrato inteligente executa as transações antes de serem adicionadas a um bloco. Esse processo consiste na verificação pelo contrato inteligente se o valor a ser pago é maior ou igual ao preço dos dados fixado pelo proprietário e se a organização possui *tokens* suficientes para efetuar a compra. Pares validam as transações que atendem

os requisitos anteriores e invalidam as que não atendem. Caso a transação seja válida, o valor pago de *tokens* é descontado da conta da organização do comprador e transferidos para a organização do vendedor.

O sistema proposto provê uma gerência da confiança de forma distribuída e autônoma através de contratos inteligentes executados na corrente de blocos. Propõe-se uma transação de avaliação TX_{fb} que invoca a função de contrato inteligente responsável pelo cálculo da confiança e da reputação de um vendedor. A transação de avaliação contém o identificador da transação de compra correspondente para garantir que a transação de compra exista e que o comprador precise assinar a transação de avaliação para garantir o não repúdio. Após adquirir os dados d de um vendedor v, um comprador c poderá emitir uma transação de avaliação TX_{ava} definida como:

$$TX_{ava} = [TX_{ID_{com}}|Sig_c|\beta_{c,d_r}|\lambda_{c,v_r}|T_j] \tag{10}$$

em que $TX_{ID_{com}}$ é a transação de compra correspondente, Sig_c é a assinatura do comprador, β_{c,d_r} é uma classificação do comprador c em os dados d, λ_{c,v_r} são uma classificação do comprador b no vendedor v e T_j é um campo de texto ao qual o comprador pode justificar sua classificação para outras partes da rede. O contrato inteligente define '0' como uma interação negativa e '1' como uma interação positiva nos campos β_{c,d_r} e λ_{c,v_r}.

A Figura 2 ilustra o processo completo para a comercialização de dados entre dois

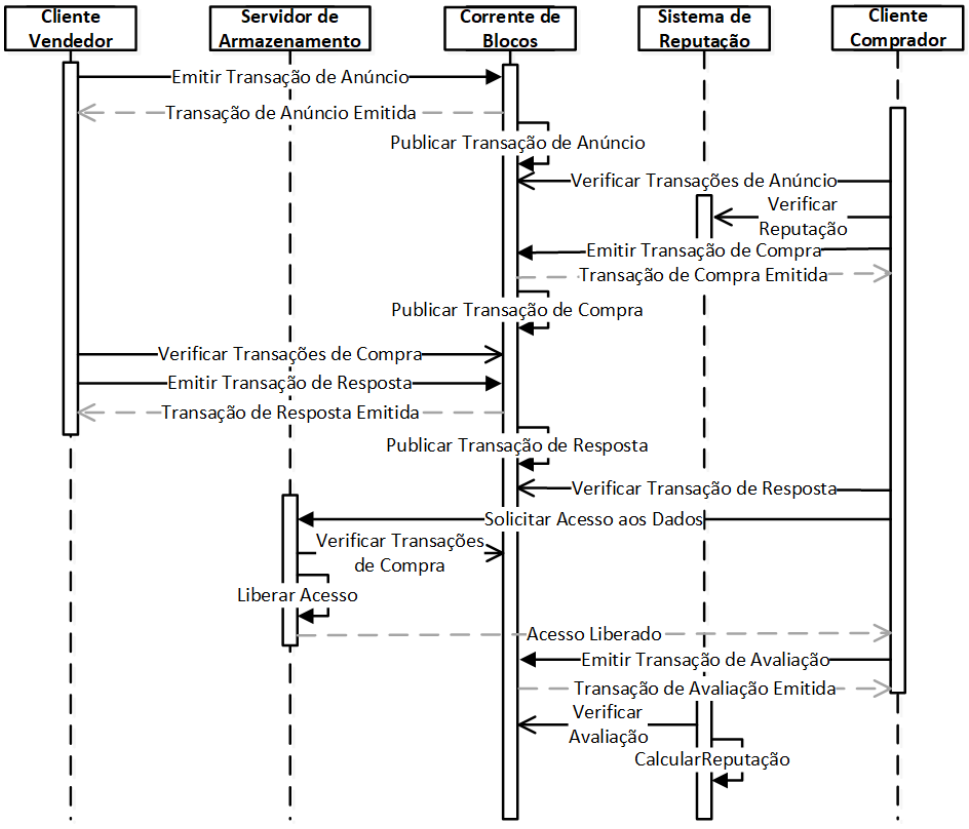


Figura 2. Diagrama de sequência do sistema proposto representando uma venda de dados entre duas organizações.

domínios. Se o vendedor considerar as classificações β_{c,d_r} e λ_{c,v_r} injustas, poderá exibir uma sinalização de insatisfação [Malik et al. 2019]. Os contratos verificam: (i) se o vendedor sinaliza todos os comentários negativos; (ii) se o comprador emite avaliações negativas para toda compra com o vendedor; (iii) se outros vendedores sinalizam o comprador. A verificação atenua ataques de maldizer, em que um comprador mal-intencionado fornece uma avaliação negativa para prejudicar um vendedor honesto.

6. Simulação e Resultados

Um protótipo do sistema proposto foi desenvolvido utilizando a plataforma de código aberto Hyperledger Fabric v2.0 [Androulaki et al. 2018] para desenvolver a corrente de blocos permissionada. O aspecto organizacional do Hyperledger Fabric se ajusta ao cenário multi-domínio da proposta, em que empresas comercializam os dados. Um computador Intel i7-8700 CPU 3.20 GHz com 32 GB RAM e 12 núcleos de processamento hospeda os nós da rede de corrente de blocos como contêineres Docker. Estipulou-se um número de transações por bloco igual a 100, como usado em trabalho anterior sobre avaliação de desempenho da plataforma Hyperledger Fabric [Gorenflo et al. 2019]. A arquitetura do Hyperledger Fabric apresenta três tipos de nós: clientes, pares e ordenadores. Os clientes representam usuários e emitem transações que precisam ser executadas por pares endossadores (*endorsers*), que são responsáveis por verificar a validade da transação. Caso a transação seja válida, o cliente recebe as transações assinadas pelos pares endossadores e envia a transação com as assinaturas dos endossadores para nós ordenadores, que executam um protocolo de consenso e ordenam as transações em um bloco. A configuração do protótipo compreende cinco nós ordenadores e o protocolo de consenso Raft [Ongaro and Ousterhout 2014].

O contrato inteligente, escrito em Go, é executado em todos os pares, eliminando uma entidade centralizada de confiança e implementando as lógicas de transação descritas na seção anterior, além de um sistema de *tokens* e contas de organizações. Os *tokens* funcionam como moedas que as organizações podem usar para a comercialização de dados. O valor real desses *tokens* pode ser discutido fora da corrente (*off-chain*) entre as organizações. O contrato restringe o campo de justificativa T_j da transação de avaliação TX_{fb} em 280 caracteres para limitar o tamanho da transação.

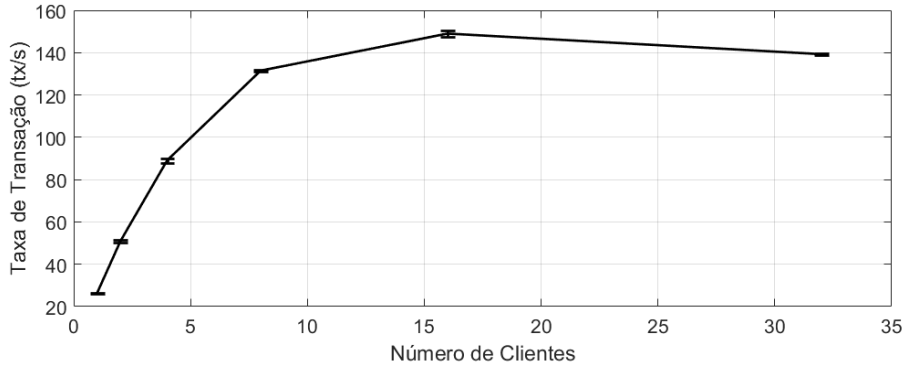
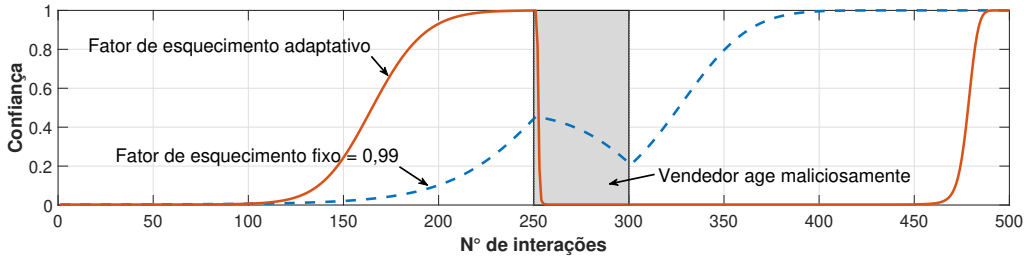


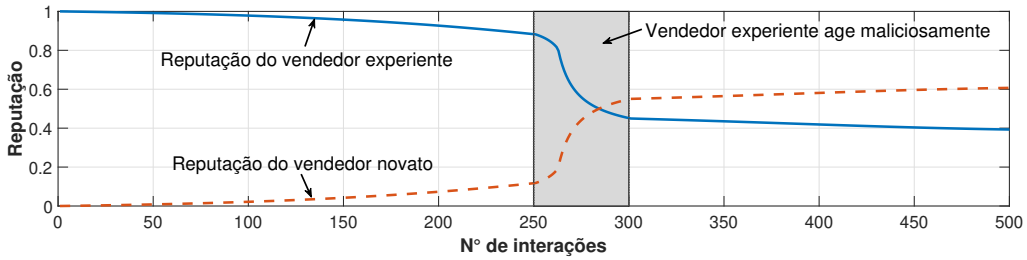
Figura 3. Taxa de transação por segundo em função do número de clientes.

A transação de resposta TX_{res} será implementada em trabalhos futuros.

O primeiro experimento mede vazão da transação de avaliação emitida por usuários que compram dados de vendedores na rede de corrente de blocos. A taxa de transação corresponde a razão entre o número total de transações emitidas e o tempo decorrido para todos os clientes emitirem todas as transações. A Figura 3 exibe os resultados experimentais da taxa de transação por segundo em função do número de clientes na rede, com intervalo de confiança de 95%. A taxa de transação cresce com o número de clientes emitindo transações paralelamente, pois um baixo número de clientes limita a vazão da rede na vazão em que as transações são emitidas. Pode-se observar que o sistema processa de maneira ágil a avaliações de centenas de usuários, atingindo o valor médio de 149 transações por segundo.



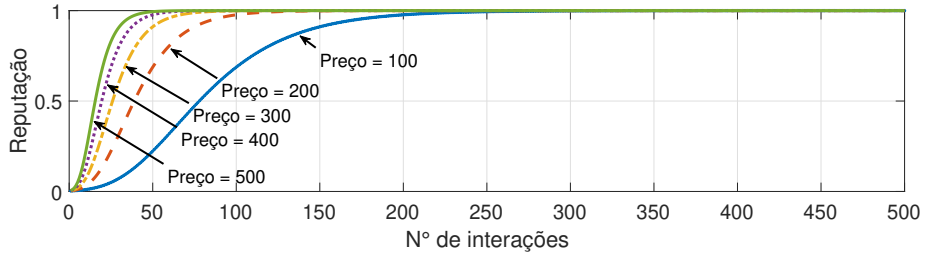
(a) Evolução da confiança de um vendedor para fatores de esquecimento fixo e adaptativo durante um ataque de dissimulação. O fator de esquecimento adaptativo recompensa vendedores que sempre foram honestos e pune fortemente vendedores maliciosos, entre 250 e 300.



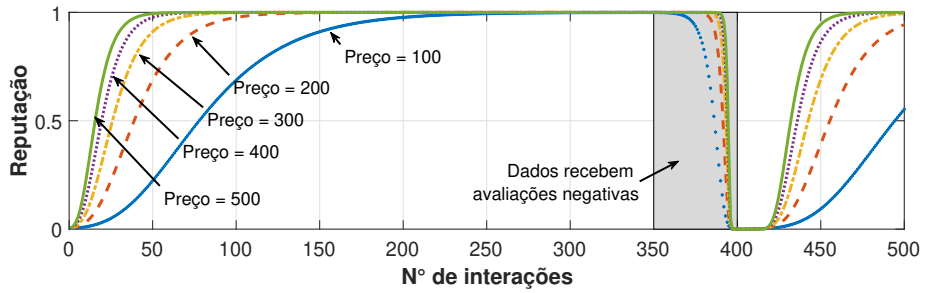
(b) Evolução da reputação de um vendedor novato que recebe o dobro de avaliações positivas que um vendedor experiente. Ao agir maliciosamente, no intervalo de 250 a 300, o vendedor experiente perde reputação rapidamente e é ultrapassado pelo vendedor novato.

Figura 4. Evolução da confiança e reputação de vendedores.

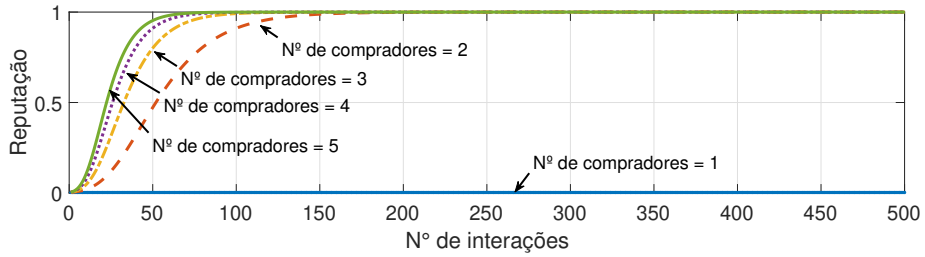
O segundo experimento avalia o impacto do fator de esquecimento adaptativo em comparação a um fator de esquecimento fixo quando ocorre um ataque de dissimulação. A Figura 4(a) ilustra as três fases do ataque. Entre as interações 0 e 250, o atacante age honestamente para ganhar confiança no sistema. Nessa fase, o fator de esquecimento adaptativo recompensa suas boas ações, pois, até então, não houve qualquer ação maliciosa. Entre as interações 251 e 300, o atacante se aproveita da confiança construída para agir maliciosamente. Com o fator de esquecimento fixo, sua confiança cai gradualmente. Com o fator de esquecimento adaptativo, o atacante é rapidamente punido nas primeiras ações maliciosas até chegar à confiança nula. Entre as interações 301 e 500, o atacante volta a se comportar honestamente para recuperar sua confiança. Com o fator de esquecimento fixo, o sistema esquece rapidamente o comportamento malicioso do atacante, que ganha confiança de forma imediata. No entanto, com o fator de esquecimento adaptativo, o atacante continua com confiança nula e precisa de mais interações para recuperá-la novamente. O experimento demonstra que, com o fator de esquecimento adaptativo, o sistema pune atacantes de forma mais eficiente e previne possíveis danos a



(a) Evolução da reputação de um dado com diferentes preços anunciados. O experimento considera somente avaliações positivas.



(b) Evolução da reputação de um dado com diferentes preços anunciados. Ao receber avaliações negativas, o dado perde reputação rapidamente e precisa de mais interações positivas para recuperar uma alta reputação.



(c) Evolução da reputação de um dado com diferentes números de compradores. Para ganhar reputação rapidamente, o dado precisa receber avaliações positivas de múltiplos compradores.

Figura 5. Evolução da reputação dos dados no sistema proposto.

diversos compradores.

O terceiro experimento avalia a evolução da reputação de um vendedor novato ao entrar em um sistema onde os demais vendedores já possuem alta reputação. A Figura 4(b) ilustra o caso em que um vendedor novato realiza duas vezes mais vendas positivas que um vendedor experiente. Na primeira fase, o vendedor novato começa com reputação nula e a aumenta gradualmente conforme recebe mais interações positivas que o vendedor experiente. Nesse caso, o vendedor experiente, apesar de ainda possuir reputação significativamente maior, deve aumentar suas vendas positivas para competir com o vendedor novato ou será lentamente ultrapassado. No entanto, na segunda fase o vendedor experiente decide agir maliciosamente anunciando dados de baixa qualidade apenas para alavancar seu número de vendas. Com as interações negativas resultantes, os compradores punem o vendedor experiente, que perde ainda mais reputação e é rapidamente ultrapassado pelo vendedor novato. Por fim, na terceira fase, o vendedor experiente retorna à situação da primeira fase. Porém, apesar de estar no sistema há mais tempo, sua reputação é menor que a do vendedor novato. O experimento demonstra que, assim como

na vida real, o sistema beneficia vendedores experientes e exige que os novatos possuam um diferencial. O vendedor experiente não pode se basear apenas na experiência para preservar sua reputação, pois será rapidamente ultrapassado pelo vendedor novato.

O último experimento verifica o comportamento da reputação dos dados ao longo de diversas interações. A Figura 5(a) ilustra o crescimento da reputação dos dados de acordo com o preço anunciado, com somente interações positivas. O experimento considera uma janela em que a soma de preços de outros vendedores mais recentes é igual a 2000 e que os vendedores interagem com 4 compradores. A taxa de crescimento da reputação dos dados é proporcional ao preço anunciado dos dados, como esperado pela Equação 8. Portanto, o sistema recompensa de maneira efetiva dados de alto valor. No cenário da Figura 5(b), os dados passam a receber avaliações negativas entre as interações 300 e 400 e retornam a receber avaliações positivas na interação número 400. Percebe-se que ao receber 50 avaliações negativas, a reputação dos dados em todas as faixas de preço atinge o valor mínimo. Após receberem 50 avaliações positivas novamente, os dados não atingem a reputação máxima que possuíam. Dessa forma, o sistema pune de maneira eficaz o comportamento malicioso. Além disso, os dados com preço mais alto caem a uma taxa maior que a de preços mais baixos. Assim, o sistema pune de maneira proporcional ao dano financeiro causado pelo vendedor. A Figura 5(c) ilustra o crescimento da reputação de dados de acordo com o número de compradores. O experimento considera uma janela em que a soma de preços de outros vendedores mais recentes é igual a 2000 e que o dado anunciado custa 400. Neste cenário, um vendedor que interage com somente um comprador mantém a reputação em zero, mitigando conluio entre vendedor e comprador e recompensando os vendedores que interagem com mais compradores.

7. Conclusão

A tecnologia de corrente de blocos, aliada aos contratos inteligentes, provê a transparência necessária para a comercialização de dados, permitindo que os proprietários mantenham o controle sobre os próprios dados. Entretanto, a corrente de blocos não garante a entrega de dados adquiridos ou a qualidade dos dados anunciados. Portanto, é necessário identificar o comportamento malicioso de vendedores no sistema e torná-lo público aos compradores. Este artigo propõe um sistema seguro de comercialização de dados utilizando corrente de blocos, reputação e confiança. O sistema proposto permite a comercialização de dados de maneira automática, sem a necessidade de intervenção do vendedor. Ainda, o artigo define uma reputação para os dados anunciados que considera o preço deles e constrói a reputação de um vendedor a partir da visão global da rede sobre ele. Um protótipo do sistema proposto foi implementado e avaliado. Os resultados mostram que a proposta consegue mitigar ataques tradicionais de sistemas de reputação de maneira eficiente e registra avaliações de maneira rápida na corrente de blocos, atingindo a taxa de 150 avaliações por segundo. O sistema de reputação implementado pune fortemente vendedores maliciosos e dados com baixa qualidade. Como trabalhos futuros, pretende-se implementar o sistema de reputação proposto em contratos inteligentes para garantir o processamento de avaliações de forma distribuída e automática.

Referências

Androulaki, E., Barger, A., Bortnikov, V., Cachin, C., Christidis, K., Caro, D., et al. (2018). Hyperledger Fabric: A distributed operating system for permissioned block-

chains. In *Proceedings of the Thirteenth EuroSys Conference*, page 30. ACM.

Anônimo (2020). Oculto para revisão cega.

de Oliveira, G. H. C., Nogueira, M., and dos Santos, A. L. (2019). Controle de acesso à IoT baseado na percepção de comunidade e confiança social contra ataques Sybil. In *SBSeg'2019*. SBC.

de Oliveira, M. T., Reis, L. H., Medeiros, D. S., Carrano, R. C., Olabarriaga, S. D., and Mattos, D. M. (2020). Blockchain reputation-based consensus: A scalable and resilient mechanism for distributed mistrusting applications. *Computer Networks*, 179:107367.

Dennis, R. and Owen, G. (2015). Rep on the block: A next generation reputation system based on the blockchain. In *ICITST'2015*, pages 131–138, London. IEEE.

Gorenflo, C., Lee, S., Golab, L., and Keshav, S. (2019). FastFabric: Scaling hyperledger fabric to 20,000 transactions per second. In *IEEE ICBC'2019*, pages 455–463.

Kamvar, S. D., Schlosser, M. T., and Garcia-Molina, H. (2003). The Eigentrust algorithm for reputation management in P2P networks. In *WWW'03*, pages 640–651.

Lodderstedt, E., McGloin, M., and Hunt, P. (2013). OAuth 2.0 threat model and security considerations. IETF. RFC 6819. Disponível em http://www.rfc-editor.org/rfc/rfc6819.txt. Acessado em 8 de julho de 2020.

Malik, S., Dedeoglu, V., Kanhere, S. S., and Jurdak, R. (2019). TrustChain: Trust Management in Blockchain and IoT Supported Supply Chains. In *IEEE Blockchain'2019*, pages 184–193, Atlanta, GA, USA. IEEE.

Michelin, R. A., Dorri, A., Lunardi, R. C., Steger, M., Kanhere, S. S., Jurdak, R., and Zorzo, A. F. (2018). SpeedyChain: A framework for decoupling data from blockchain for smart cities. In *EAI MobiQuitous'18*, pages 145–154.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Disponível em https://bitcoin.org/bitcoin.pdf. Acessado em 8 de julho de 2020.

Ongaro, D. and Ousterhout, J. (2014). In search of an understandable consensus algorithm. In *USENIX ATC'2014*, pages 305–319, Philadelphia, PA. USENIX.

Putra, G. D., Dedeoglu, V., Kanhere, S. S., and Jurdak, R. (2020). Trust Management in Decentralized IoT Access Control System. In *IEEE ICBC'2020*, page 9. IEEE.

Rebello, G. A. F., Alvarenga, I. D., Sanz, I. J., and Duarte, O. C. M. B. (2019). BSec-NFVO: A blockchain-based security for network function virtualization orchestration. In *IEEE ICC*, pages 1–6.

Sun, Y., Han, Z., and Liu, K. R. (2008). Defense of trust management vulnerabilities in distributed networks. *IEEE Communications Magazine*, 46(2):112–119.

Truong, H. T. T., Almeida, M., Karame, G., and Soriente, C. (2019). Towards secure and decentralized sharing of IoT data. In *IEEE Blockchain'2019*, pages 176–183.

Velloso, P. B., Laufer, R. P., de Oliveira Cunha, D., Duarte, O. C. M. B., and Pujolle, G. (2010). Trust management in mobile ad hoc networks using a scalable maturity-based model. *IEEE Transactions on Network and Service Management*, 7(3):172–185.