UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Programa de Pós-Graduação em Ciência da Computação

Gabriel Gomes Gaspar

Um Protocolo de Votação com Privacidade Incondicional

Belo Horizonte
2021

Gabriel Gomes Gaspar

Um Protocolo de Votação com Privacidade Incondicional

Versão Final

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Ciência da Computação.

Orientador: Jeroen van de Graaf

Belo Horizonte
2021

Gaspar, Gabriel Gomes

G249p Um protocolo de votação com privacidade incondicional [recurso eletrônico] / Gabriel Gomes Gaspar — 2021.
1 recurso online (154 f. il, color.)

Orientador: Jeroen Antonius Maria van de Graaf.
Dissertação (mestrado) - Universidade Federal de Minas Gerais, Departamento de Ciência da Computação, Instituto de Ciências Exatas
Referências: f. 147-154.

1. Computação – Teses. 2. Criptografia – Teses. 3. Votação – Medidas de segurança – Teses. 4. Voto eletrônico – Teses. I. Graaf, Jeroen Antonius Maria van de. II. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Computação. III.Título.

CDU 519.6*43(043)

Ficha Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa CRB 6/1510 Universidade Federal de Minas Gerais - ICEx

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

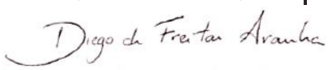
FOLHA DE APROVAÇÃO

Um Protocolo de Votação com Privacidade Incondicional

GABRIEL GOMES GASPAR

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. JEROEN ANTONIUS MARIA VAN DE GRAAF - Orientador
Departamento de Ciência da Computação - UFMG



PROF. DIEGO DE FREITAS ARANHA
Department of Computer Science - Aarhus University

PROF. MARIO SÉRGIO FERREIRA ALVIM JÚNIOR
Departamento de Ciência da Computação - UFMG

PROF. ALEJANDRO HEVIA
Departamento de Ciencias de la Computación - Universidad de Chile

elo Horizonte, 8 de Setembro de 2021.

Dedico este trabalho a todos aqueles que, mesmo em tempos de obscurantismo intelectual, ousam manter-se firmes na missão de erguer alto a vela da ciência a iluminar os caminhos trilhados pela humanidade.

Agradecimentos

Seria impossível e injusto iniciar uma sessão de agradecimentos sem primordialmente começar por minha mãe, a professora Maria do Rosário, que desde cedo me incentivou a trilhar o árduo caminho do aprender. Não posso também deixar de mencionar todo o suporte, carinho e apoio oferecidos pela cara Cristiane Holanda Costa, que me apoiou, auxiliou e acompanhou em grande parte dos desafios que tive de enfrentar no decorrer do mestrado. Agradeço também a meu pai, Viriato Gaspar, e a meu irmão, Danilo Gaspar, por suas presenças constantes, ainda que distantes. Agradeço também muito especialmente a meu orientador, Jeroen van de Graaf, pela acolhida na UFMG, por toda a necessária orientação e, primordialmente, pela confiança em meu trabalho e em minhas capacidades. Deixo também um saudoso agradecimento a meu padrasto, Gilberto Bezerra da Costa, o Gil, por seu apoio em minha mudança para Belo Horizonte, então uma cidade nova e desconhecida para mim, e por me possibilitar realizar meu curso de mestrado sem maiores preocupações com o que então aparentemente deixava para trás. Por fim, agradeço também saudosamente à Aika, minha Akita, que por noites a fio pacientemente me acompanhou enquanto desempenhava minhas atividades de mestrado nas solitárias noites de um mundo em pandemia.

Resumo

O desenvolvimento de protocolos eleitorais criptográficos ganhou especial tração nas primeiras décadas do século XXI, impulsionado principalmente pelos diversos avanços tecnológicos e científicos junto às novas demandas democráticas frente o contexto contemporâneo de crescente informatização. Central a tais protocolos encontra-se o conceito de independência de software, o qual especifica que um sistema eleitoral eletrônico deve produzir uma trilha de evidências que seja capaz de atestar a correção dos resultados eleitorais e que possa ser verificada de forma independente de maquinário eletrônico específico, sem contudo violar o sigilo individual dos eleitores. Em tais sistemas criptográficos, no entanto, faz-se necessário um compromisso com a integridade incondicional dos resultados ou o sigilo incondicional do voto dos eleitores, sendo comum nos protocolos mais proeminentes prezar pela incondicionalidade da integridade. Nesse sentido, o presente trabalho descreve uma proposta de protocolo eleitoral que também respeita o princípio da independência de software mas que, em contraste, preza pela incondicionalidade do sigilo individual dos votos, sob o credo de que este deve apresentar capacidades de manutenção eterna. Ademais, o protocolo proposto viabiliza verificação de ponta-a-ponta (E2E-V) dos votos por meio de um recibo impresso fornecido ao eleitor, além de prover uma metodologia de auditoria universal por meio de uma equação matemática simples que deve ser verificada como verdadeira para toda eleição íntegra, de tal forma permitindo a qualquer observador do processo, ante premissas criptográficas bem estabelecidas, atestar a integridade dos resultados eleitorais produzidos. O trabalho apresenta ainda uma instanciação do protocolo na forma de uma implementação, viabilizando ilustração e análise de alguns de seus aspectos mais centrais.

Palavras-chave: Criptografia. Sistemas Eleitorais Eletrônicos. Votação.

Abstract

The development of cryptographic electoral protocols gained further traction in the first decades of the twenty-first century, largely due to various advances in science and technology and also motivated by the renewed demands of the various democracies in the context of growing informatization. Central to these protocols is the concept of software independence, which states that an electoral system must produce an evidence trail that allows for integrity checks on the election's results independently from election-specific electronic machinery, all the while preserving voters' privacy. In such cryptographic systems, however, it seems inevitable that a compromise needs to be made between unconditional integrity of the results and unconditional privacy of the votes, unconditional integrity being the common choice in many of the most prominent protocols. In this sense, this work proposes an electoral protocol that likewise respects the principle of software independence, but that contrasts to many other proposals in its choice of unconditional privacy of the votes under the belief that voters need to be assured eternal secrecy regarding their electoral choices. Furthermore, the protocol proposed enables end-to-end verifiability (E2E-V) of individual votes via a printed receipt issued to each voter, besides providing universal audit features by means of a simple mathematical equation that must hold for a correctly executed election. This work also presents an implementation of the protocol that allows for illustration of its workings and analysis of some of its core aspects.

Keywords: Cryptography. Electronic Electoral Systems. Voting.

Lista de Figuras

Lista de Códigos

Sumário

Capítulo 1

Introdução

Em seu influente escrito *On Democracy*, o cientista político Robert Dahl promove uma reflexão acerca daquilo que se vem a compreender como uma *democracia*, haja vista a imensa diversidade de aplicações do termo e suas mais variadas formas de existência, seja nos idos da história humana, seja em sua contemporaneidade [15]. Dahl ressalta que um processo que se pretenda democrático necessariamente requer participação efetiva dos membros de uma associação democrática nas tomadas de decisão e que a todos os membros seja dada a oportunidade de *votar*, sendo seus *votos* contados de maneira igualitária.

Dahl estende esta sua compreensão a democracias de larga escala, elencando como um de seus principais requisitos a realização de "eleições livres, justas e frequentes, nas quais os riscos de coerção sejam comparativamente raros" [15, p.85]. Em tal consideração, Dahl estabelece de maneira sutil ao menos três aspectos cruciais referentes a processos eleitorais em organizações democráticas: primeiramente, eleições são processos constantes e frequentes; em segundo, eleições devem ser justas com relação à sua condução, nisto inclusa a integridade de seus resultados; e em terceiro, os participantes da eleição, os *eleitores*, devem se sentir livres para expor suas opções eleitorais sem receios de coerção ou retaliações por conta delas. Assim sendo, tendo a realização de eleições como evento recorrente em uma democracia, faz-se de crucial importância resguardar as supracitadas garantias identificadas como necessárias ao devido processo eleitoral, quais sejam, a *integridade* de seus resultados e a liberdade dos eleitores para manifestação livre de suas escolhas, conferida atual e historicamente pelo aspecto do *sigilo* do voto individual.

Em termos de tais garantias, talvez a mais óbvia seja aquela referente à *integridade* dos resultados eleitorais. Afinal, de modo a que os resultados de uma eleição possam ser considerados válidos e, com isto, devidamente aceitos pelos membros de uma organização democrática, faz-se necessário que os votos individuais possam ser demonstrados como livres de adulterações semânticas quando agregados para formação de um resultado final. Do contrário, os resultados justificadamente abrem margem à contestação de sua validade e, por conseguinte, da legitimidade dos vencedores. Isto posto, é de limitada valia democrática um sistema que seja incapaz de comprovar os próprios resultados eleitorais que externaliza. Por isso é comum o dizer de que é importante a um sistema eleitoral não apenas revelar o vencedor de um pleito, mas também convencer o perdedor de que ele de

fato perdeu [68, p.9].

Intrínseca a tal necessidade de integridade dos resultados de uma eleição, e em conformidade com o estabelecido por Dahl, é perceptível, no sentido de viabilizar a livre manifestação dos eleitores, a necessidade crítica de que também seja garantido o *sigilo* de cada eleitor com relação a seu voto. Isto implica que ninguém mais, além do próprio eleitor, deve ter conhecimento acerca de suas opções de escolha manifestas em seu voto depositado. Tal percepção da necessidade do sigilo inclusive data de idos tempos, a exemplo da introdução do voto secreto na Roma antiga [69] no intuito de evitar influências indevidas sobre os processos decisórios, principalmente por parte das castas mais altas da sociedade, e oferecer ao eleitor maior liberdade em suas escolhas.

Ambas estas garantias, portanto, encontram-se intimamente interligadas. Afinal, se um coator é capaz de se utilizar de seu poder de influência a fim de induzir eleitores a votarem conforme sua vontade (a do coator) e em discordância com suas intenções originais de voto (as dos eleitores), certamente a integridade do resultado eleitoral se mostrará prejudicada no referente ao que seria esperado ante livre manifestação da intenção pública. E, a depender do âmbito de tal influência adulteradora, em um caso extremo pode-se inclusive conceber a inversão dos resultados de uma eleição, conferindo vitória a quem, sob livre e espontânea manifestação dos eleitores, talvez jamais viesse a ganhar o pleito em questão.

Tais aspectos ganham matizes consideravelmente mais complexos quando considerados sob a ótica da sociedade contemporânea do final do século XX e início do século XXI. Isto pois, tendo em vista a consolidação dos meios computacionais e a ubiquidade da informatização nos mais diversos aspectos do fazer humano, seria de se esperar que de fato chegasse o momento em que a humanidade questionasse: por que não aplicar as já comprovadas facilidades providas por computadores em tantas áreas também nos aspectos mais centrais da manutenção democrática? Em particular, por que não se utilizar de aparato computacional para realização dos processos eleitorais? Afinal, é de fácil percepção que computadores são capazes de resolver de maneira muito mais eficiente vários dos desafios mais óbvios comumente enfrentados por sistemas de tal natureza em atuais contextos democráticos. Dentre alguns desses desafios citam-se o tratamento de cédulas eleitorais extensas, a acessibilidade de eleitores em condições diversas (e.g. debilidades físicas, analfabetismo funcional) e o tempo demandado para a produção das totalizações de um pleito. Em todos esses aspectos, computadores mostram-se como prontas soluções, aptos a prover melhorias significativas em tais processos.

Em contexto brasileiro, o histórico de fraudes eleitorais do país proveu motivação adicional para que se corresse com um projeto de implementação nacional de um sistema puramente eletrônico no breve intervalo de 5 anos, o qual resultou em máquinas de votação de registro puramente digital cunhadas de *urnas eletrônicas*. O antigo processo eleitoral, que se utilizava de cédulas em papel e urnas de lona com apuração manual, era amplamente

vulnerável aos mais diversos tipos de fraude, além de se mostrar demasiadamente lento e sujeito a erros humanos na totalização. Em termos do sigilo, várias modalidades de violação foram criadas, do chamado "voto carreirinha" [1] às imposições de voto por parte de figuras regionais influentes como os coronéis, culminando inclusive na criação do termo "voto de cabresto" para indicar a prática recorrente de coerção eleitoral e compra ou venda de votos. Nas palavras do ministro Carlos Velloso, presidente do Tribunal Superior Eleitoral (TSE) à época da implantação do projeto eleitoral eletrônico brasileiro, tendo em vista o histórico eleitoral do país, o grande objetivo das urnas eletrônicas era o de "eliminar a fraude no processo eleitoral afastando a intervenção humana" [67].

Este breve histórico brasileiro, à luz da supracitada citação do ministro, provê ponto de partida ideal para a ilustração da complexidade intrínseca a sistemas eleitorais e dos perigos inerentes à consideração do emprego de aparato computacional como uma espécie de panaceia eleitoral. O histórico brasileiro de fraudes nas antigas urnas de lona com cédulas em papel é forte demonstrativo de como a complexidade oriunda da participação de múltiplos atores em suas mais variadas interações no contexto dos procedimentos necessários entre a definição de uma eleição e a publicação de sua totalização abrem margem a manipulações de naturezas diversas que podem culminar na deterioração da integridade eleitoral e prejuízo ao sigilo individual do eleitor. E a citação do ministro, por sua vez, desvela a pueril ideia de que a entrega do processamento eleitoral ao domínio completo e irrestrito de computadores supostamente tornaria o processo imune a quaisquer desígnios de mãos humanas, quando em realidade apenas lhes induz um nível adicional de indireção.

Assim sendo, o emprego de ferramental computacional a processos eleitorais não deve ser encarado de forma leviana. Afinal, a máquina apenas apresenta o potencial de agilidade e precisão necessárias para processar centenas de milhões de votos em conformidade com algum algoritmo em reduzido intervalo de tempo. A natureza e significado dos algoritmos que especificam tal processamento, no entanto, ainda passam pela intermediação humana, que, se em tal contexto aplicada para intentos desonestos, se apresenta com potencial, largamente ampliado pela máquina, no sentido de corroer de uma eleição suas tão necessárias garantias de integridade e sigilo. Isto posto, computadores, apesar de se mostrarem como óbvias soluções a várias das demandas eleitorais, também se mostram como potencial vetor de fraudes se incautamente empregados. E tal aspecto se mostra particularmente preocupante quando considerada a complexidade de um processo eleitoral, a qual viabiliza ampla superfície para introdução de vulnerabilidades, em caráter acidental ou não.

Faz-se mister, portanto, a cuidadosa análise das possibilidades de aplicação de maquinário eletrônico como suporte aos processos democráticos. Caso contrário, assume-se

[1] Modalidade de fraude em que cédulas de papel válidas são retiradas das sessões eleitorais e preenchidas por um fraudador, que então influencia eleitores válidos a depositarem tais cédulas por ele preenchidas.

o risco de que tal maquinário venha a se tornar mecanismo próprio à potencialização das vulnerabilidades que sua introdução originalmente visava mitigar ou coibir. Em tal esforço, necessariamente devem ser levadas em consideração as particularidades da democracia em questão. Nesse sentido, Dahl também propõe três observações ao notar a patente variedade de sistemas eleitorais nas diversas democracias [15, p.191]:

1. Um país democrático deve substituir um sistema eleitoral que claramente não lhe serve.

2. O sistema eleitoral de um país pode ser adaptado a seus mais diversos aspectos, culturais, sociais, históricos ou demais.

3. A adoção de alternativas eleitorais necessita ser cuidadosamente analisada com o auxílio de profissionais competentes.

Sob a ótica de tais observações, o projeto brasileiro parece ter se preocupado fortemente com os primeiros dois pontos expostos por Dahl, mas, por algum motivo, aparentemente se recusou a sequer considerar o terceiro ponto, optando de tal forma por um sistema feito quase que em caráter de total sigilo e sem o devido aval de técnicos, cientistas, acadêmicos e pesquisadores competentes nas áreas envolvidas. O resultado é um sistema eleitoral aplicado em escala nacional que se mostra incapaz de comprovar seus próprios resultados, exigindo confiança estrita na correta operação das máquinas empregadas, nas autoridades eleitorais, nos funcionários envolvidos e nas demais cadeias de produção e distribuição [68, p.9-10]. Tal realidade brasileira ilustra como o emprego de aparato computacional por si só não necessariamente resulta em um sistema que se mostre suficiente no sentido de prover garantias eleitorais básicas.

Tal capacidade de verificação independente dos resultados é hoje considerada mister a sistemas de tal natureza, e a busca pelo projeto e concepção de sistemas munidos de tal capacidade mostra-se como crucial à saúde democrática. À luz do exemplo brasileiro, observa-se que chegar a um projeto consistente e munido de tais capacidades não consiste em tarefa óbvia, estando à mercê da influência de forças que não unicamente do âmbito científico ou técnico. Não obstante, ante sua supracitada importância, o esforço no sentido de se chegar a um tal projeto certamente se mostra como necessário, principalmente no contexto de democracias crescentemente informatizadas e que buscam se utilizar de tal informatização também em seus processos democráticos.

1.1 Motivação

Em tal contexto de consideração e análise da aplicabilidade de soluções computacionais às demandas eleitorais, formaliza-se o conceito de *independência de software* [38]. Em si, o conceito resume a ideia de que um sistema eleitoral eletrônico deve ser capaz de produzir uma trilha de evidências eleitorais que possa ser verificada de forma independente do maquinário eletrônico empregado em uma eleição no intuito de se atestar a integridade dos resultados produzidos. Tal trilha, no entanto, não deve prejudicar o necessário sigilo dos votos individuais dos eleitores, sob risco de render o processo a influências indevidas (e.g. coerção, compra/venda de votos).

Tendo em vista o conceito da independência de software, viu-se nas últimas duas décadas o surgimento de diversos sistemas eleitorais eletrônicos e protocolos que buscam prover metodologias para condução de eleições que sejam capazes de produzir tal trilha comprobatória, dentre os quais os chamados *sistemas eleitorais criptográficos*. Tais sistemas criptográficos utilizam-se de princípios da criptografia no intuito de prover garantias probabilísticas aferíveis, ante um conjunto especificado de premissas, no referente ao resguardo dos princípios democráticos de sigilo e integridade intrínsecos a uma eleição.

Alguns exemplos de tais sistemas criptográficos são o seminal *Votegrity* [12], o *Prêt à Voter* [41], o *Scantegrity* [11], o *Wombat* [7] e o *STAR-Vote* [5]. Alguns destes, como o Scantegrity e o Prêt à Voter, já chegaram a ser empregados em algumas eleições em caráter oficial. Entretanto, sua utilização não progrediu de maneira significativa para novos âmbitos e contextos. Já outros, como o STAR-Vote, partiram de iniciativas governamentais em esforço conjunto com pesquisadores e acadêmicos. Infelizmente, o STAR-Vote aparenta findar como mais um sistema promissor que jamais viu uso [35]. Não obstante, tais exemplos servem como demonstrativos de que instituições democráticas em diversos países e âmbitos estão à procura de novos sistemas para respaldar seus processos eleitorais, mas também que há desconfianças e barreiras de origens diversas com relação à propulsão de sua aplicação.

Um aspecto interessante de sistemas eleitorais criptográficos é que aparentemente é necessário ser feito um compromisso inalienável entre *integridade incondicional* ou *sigilo incondicional*. Isto significa que apenas seria possível se reter a incondicionalidade de um desses atributos perante esforços de criptoanálise vindouros e a natural corrosão temporal das premissas criptográficas empregadas no protocolo. O outro atributo apenas é capaz de reter segurança *computacional*.

Assim sendo, é hoje lugar comum nos projetos de protocolos eleitorais criptográficos atribuir-lhes garantias de *integridade incondicional* e *sigilo computacional*. Isto significa que a integridade dos resultados eleitorais é mantida incondicionalmente ante respeito às premissas criptográficas empregadas, mas o sigilo individual dos votos é apenas resguar-

dado ante hipóteses computacionais estabelecidas em caráter corrente e projetado para um futuro finito. Naturalmente, com o passar do tempo, tais hipóteses computacionais se enfraquecem, tornando-se viável violação do sigilo dos votos. No entanto, a depender do contexto de aplicação do protocolo em questão, futuras violações de sigilo podem se mostrar de consequências graves à democracia ou aos eleitores individualmente (e.g. perseguição política, humilhação pública). E tais possibilidades podem inclusive se mostrar como fatores motivantes a que eleitores se sintam coagidos durante o voto, de tal forma violando de maneira indireta a integridade dos resultados eleitorais em comparação com aqueles que seriam obtidos caso eleitores se sentissem livres para expressar sua vontade.

Isto posto, e principalmente tendo-se em mente a realidade histórica brasileira de exemplos diversos de corrosão da integridade eleitoral pela violação do sigilo, faz sentido considerar-se o projeto de protocolos eleitorais criptográficos que prezam por capacidades de *sigilo incondicional* enquanto retêm garantias de *integridade computacional* durante um período específico de tempo durante o qual os resultados eleitorais possam ser contestados, a exemplo da metodologia descrita por [14]. Tal decisão carrega em si o prospecto de que é de limitada utilidade descobrir-se como adulterar a integridade de uma eleição quando esta já é legalmente considerada finalizada (e.g. descobrir no ano de 2038 como adulterar matematicamente os resultados de uma eleição ocorrida em 2026).

Assim, embora seja difícil se argumentar acerca de garantias absolutas de sigilo (e.g. é impossível se garantir a não existência de uma câmera ou *keylogger* registrando o voto do usuário), considera-se válido o empreendimento de projeto de um protocolo de votação no qual o próprio sistema não aja, por seu funcionamento próprio, também como agente reforçador da deterioração do sigilo dos votos individuais, sendo a descrição e análise de um tal protocolo o objeto principal deste trabalho.

1.2 Contribuições do Trabalho

O presente trabalho contribui com a descrição detalhada do cerne de um protocolo eleitoral para execução de eleições presenciais capazes de resguardar, mediante correta execução, o sigilo eterno dos votos individuais ao mesmo tempo em que provê garantias de integridade computacional durante todo um período eleitoral pré-definido.

Comum a outras propostas de protocolos eleitorais criptográficos, o protocolo proposto por este trabalho apresenta as seguintes propriedades:

- emprego de dispositivos computacionais, facilitando acesso de eleitores com limitações pessoais (e.g. motoras, visuais), além de conferir maior agilidade e assertividade

aos processos de votação e totalização.

- produção de uma trilha de evidência extrínseca ao âmbito das máquinas empregadas no processo, conferindo ao protocolo *independência de software*.
- capacidades de verificação do voto de ponta-a-ponta (*end-to-end verifiability*, ou E2E-V), permitindo a eleitores conferir o correto registro de seu voto e a devida inclusão deste na totalização, além de permitir a qualquer observador interessado, independente de sua vinculação ao processo, verificar que os resultados eleitorais são produzidos corretamente.

Em contraste com protocolos mais proeminentes, o protocolo aqui descrito se diferencia por sua forte ênfase de aplicação a cenários nos quais o sigilo pós-eleitoral dos votos seja condição rigidamente necessária aos pleitos realizados, tendo como foco o histórico e a realidade eleitoral brasileira. Além disso, mesmo ante propostas com maior ênfase em sigilo, como a de [14], o presente protocolo contribui com uma combinação única das seguintes características desejáveis:

- suporte a cédulas extensas (i.e. cédulas com grande número de opções) sem sobrecarga significativa ao desempenho do protocolo.
- fácil acoplamento a sistemas eleitorais pré-existentes, bastando que haja uma metodologia bem-definida de digitalização de votos.
- capacidade de *granularização*, viabilizando verificações com precisão configurável de locais de votação (e.g. seções eleitorais) nos quais erros de integridade porventura tenham ocorrido, de tal forma facilitando o empreendimento de procedimentos de auditoria e/ou corretivos.

A junção de tais características torna o presente protocolo apto a aplicação prática em contextos demandosos tanto em termos de volume de eleitores quanto de volume de candidatos aos cargos pleiteados, como ocorre, por exemplo, em eleições de âmbito nacional.

1.3 Organização da Dissertação

O restante da presente dissertação encontra-se organizada em cinco capítulos conforme abaixo dispostos:

- O Capítulo 2 explora o tópico da realização de eleições em âmbito informatizado, iniciando-se com o estabelecimento do conceito de *independência de software*, que tem sido usado como norte para o desenvolvimento de novos sistemas eleitorais eletrônicos. No capítulo explora-se ainda o estado da arte de sistemas eleitorais criptográficos por meio da análise de alguns dos protocolos mais proeminentes. Por fim, uma seção é dedicada à discussão do atual sistema eleitoral brasileiro.

- O Capítulo 3 estipula os principais conceitos matemáticos e criptográficos necessários ao entendimento do protocolo descrito neste trabalho.

- O Capítulo 4 apresenta a descrição detalhada do protocolo proposto, com ênfase em definições robustas e provas matemáticas de suas propriedades.

- O Capítulo 5 ilustra a execução do protocolo por meio de uma implementação deste utilizando-se de outros protocolos e esquemas criptográficos bem-estabelecidos na literatura.

- O Capítulo 6 apresenta as conclusões do trabalho e direções para possíveis melhorias e trabalhos futuros.

Capítulo 2

Sistemas Eleitorais Eletrônicos

Neste capítulo são discutidos exemplos de sistemas eleitorais eletrônicos, com ênfase no estado da arte de sistemas voltados a tal propósito. Partindo do conceito de *independência de software* e à luz deste, alguns sistemas criptográficos atuais já desenvolvidos e testados em campo são discutidos em termos de suas implementações e garantias. Por fim, uma análise do sistema eleitoral eletrônico brasileiro corrente é empreendida.

2.1 Independência de Software

Dada a complexidade associada a processos eleitorais, é de se esperar que a tradução destes para o âmbito computacional resulte em sistemas eleitorais eletrônicos que reflitam tal complexidade inerente ao processo em si. Parte dessa natureza deriva do conflito manifesto entre os necessários requisitos de *sigilo* e *integridade*, em que o requisito de sigilo diz respeito à inviabilidade de criação de um mapeamento entre um voto individual e o respectivo eleitor que o produziu, enquanto o requisito de integridade refere-se à inviabilidade de adulteração dos resultados de uma eleição de maneira indetectável. De modo geral, no projeto de sistemas eleitorais, fortalecer qualquer um desses requisitos impacta negativamente sobre o outro. Características como esta rendem sistemas eleitorais eletrônicos sujeitos a falhas e vulnerabilidades diversas, em caráter acidental ou não, que podem influenciar de forma degenerativa nos resultados observáveis de uma eleição ou na privacidade, muitas vezes garantida por lei, esperada por um eleitor para manifestação de suas opções de voto.

Assim sendo, sistemas eleitorais eletrônicos herdam características associadas a sistemas complexos em si: pequenos erros podem influenciar a operação do software com resultados imprevisíveis; ademais, os custos associados a provas extensivas de correção, testagem e auditoria de tais sistemas são consideráveis, possivelmente proibitivos. Dessa forma, o consenso científico da área é de que a ênfase da verificação eleitoral não deve pender sobre o correto funcionamento do sistema utilizado em si, mas sim sobre os resultados

eleitorais fornecidos por tais sistemas.

Nesse sentido, considera-se que a característica fundamental a ser oferecida por tais sistemas é que estes sejam capazes de prover resultados que sejam verificáveis de uma maneira tal que a verificação da correção de tais resultados independa de qualquer hipótese acerca do correto funcionamento do software empregado. Ou, de outra forma: a correção dos resultados eleitorais não pode depender estritamente do correto funcionamento do maquinário computacional empregado. Esta expectativa é resumida sob o conceito de *independência de software* [38], expresso na Definição 2.1.1 seguinte:

Definição 2.1.1 (Independência de Software)**.** *Um sistema de votação é dito ser* ***independente de software*** *se uma alteração não detectada em seu software é incapaz de causar uma alteração indetectável nos resultados eleitorais.*

Ressalta-se aqui que as *"alterações"* mencionadas na Definição 2.1.1 não se resumem estritamente àquelas de natureza proposital: erros (*bugs*) em programas são ocorrências comuns no processo de desenvolvimento de software, podendo ser nele embutidos em caráter acidental e mantidos imperceptíveis por consideráveis períodos de tempo. Dessa forma, a questão da independência de software não se lastreia puramente em questão de desconfiança ativa acerca da idoneidade dos desenvolvedores de um sistema de votação, mas também no reconhecimento da própria natureza do processo de desenvolvimento de software em si ao preconizar que não se deve respaldar toda a integridade eleitoral sobre este.

Outro aspecto relevante referente à independência de software é que esta não se caracteriza como um princípio teórico idealizado. Tampouco se pretende a advogar contra o uso de software em contextos eleitorais em plena ignorância das facilidades que são capazes de prover, como prevenção de falhas e inconsistências de natureza humana durante a votação, facilidades para eleitores com debilidades físicas e agilidade na totalização. Pelo contrário, exatamente por levar em consideração aspectos realistas intrínsecos ao software e a seu funcionamento, o princípio evidencia as limitações deste e prevê garantias que um sistema de propósito eleitoral deve oferecer de modo a tornar a integridade de seus resultados independente das particularidades do software utilizado.

Reflexo dessa característica prática do conceito é a possibilidade de se identificar algumas abordagens que conferem independência de software a um sistema de votação. Uma destas consiste em agregar a um sistema eletrônico um lastro físico que permita ao eleitor verificar que seu voto é depositado exatamente como pretendido. Este é o caso de sistemas que proveem um *registro em papel verificável pelo eleitor (VVPR)* ou um *rastro de auditoria verificado pelo eleitor (VVPAT)*, nos quais utiliza-se de um registro físico do voto, de interpretação prontamente compreensível pelo eleitor, como uma forma de tornar cada voto individual independente de qualquer máquina empregada no processo. Assim, alterações em software que possam adulterar resultados eleitorais são tornadas passíveis

de verificação por conta da existência física de cada voto fora do âmbito digital, que em si é invisível ao eleitor e controlado pelo maquinário empregado.

Outra abordagem que vem sendo alvo de constante estudo e demonstrando crescentes possibilidades de aplicação prática é a junção da criptografia ao contexto eleitoral, constituindo os chamados sistemas criptográficos de votação, também comumente denominados *sistemas de votação verificáveis de ponta-a-ponta (E2E-V)* pelas propriedades de verificação que são capazes de prover. Tal abordagem confere aos sistemas um fluxo de integridade e respectivas garantias que permitem acompanhar um voto em todo seu percurso pelo fluxo eleitoral, do depósito conforme intenção do eleitor até sua composição na totalização final. Alguns desses sistemas se utilizam de recibos em papel que são levados pelo eleitor após a votação. Estes recibos contêm material criptográfico que permite a um eleitor acompanhar a integridade de seu voto por todos os estágios eleitorais, sem, no entanto, revelar os conteúdos do voto. Dessa forma, são conferidas garantias criptográficas ao sigilo e à integridade de cada voto e, por conseguinte, também à integridade dos resultados finais em si.

Assim, à luz desse conceito, os diversos sistemas eleitorais que se utilizem de aparato eletrônico podem ser verificados como independentes de software ou não, a depender do relacionamento que estabelecem com o maquinário que empregam. Na seção seguinte alguns sistemas eleitorais independentes de software e com um viés de aplicação prática em contextos presenciais são brevemente analisados. Em seguida, um paralelo é traçado com a relação de dependência eletrônica estabelecida pelo sistema eleitoral eletrônico brasileiro.

2.2 O Estado da Arte

A presente seção discute brevemente uma seleção de exemplos de sistemas eleitorais eletrônicos criptográficos propostos nas últimas décadas, principalmente aqueles com finalidades de aplicação prática e em contextos estritamente voltados à votação presencial. Ênfase especial é dada ao STAR-Vote por conta de sua similaridade e influência no protocolo eleitoral proposto por este trabalho.

2.2.1 Votegrity

O *Votegrity* [12] foi um dos primeiros protocolos criptográficos E2E-V criados com um viés próprio para aplicação prática. Até então, tendo em vista a necessidade de sigilo dos votos individuais dos eleitores, tinha-se como princípio que sistemas eleitorais não poderiam emitir qualquer tipo de recibo de votação ao eleitor, sob o credo de que estes poderiam ser utilizados como forma de comprovação do voto a terceiros, abrindo assim margem à coerção e venda de votos. Entretanto, a carência de recibos dessa natureza também rende os eleitores a apenas suposições de correção dos resultados eleitorais, isto porque tolhe-se do eleitor e de observadores qualquer capacidade de atestar o devido registro dos votos individuais ou mesmo de que estes foram corretamente incluídos na totalização final publicada.

O Votegrity propôs então uma solução para tal dilema introduzindo o conceito de recibo cifrado: o recibo levado pelo eleitor consiste em uma versão de seu voto que é cifrada por múltiplas chaves criptográficas detidas por entidades ou indivíduos denominados *trustees*, que atuam conjuntamente como *guardiões eleitorais*. Um aspecto interessante do Votegrity é que a cédula eleitoral é composta por padrões visuais (i.e. pixels) que são cifrados e impressos em duas partes distintas. Ambas as partes codificam o mesmo voto, mas este apenas é tornado visível e legível quando ambas estas partes são fisicamente sobrepostas, fazendo com que os padrões aparentemente aleatórios componham conjuntamente um voto de forma compreensível pelo eleitor.

Assim, no ato da votação, o eleitor faz suas escolhas eleitorais em uma máquina, a qual imprime o voto em duas partes que são mostradas sobrepostas, tornando o voto legível e, por conseguinte, verificável pelo próprio eleitor como em conformidade com suas escolhas. O eleitor então escolhe uma das partes para levar como recibo, sendo tal parte também armazenada em formato digital pela máquina de votação. A versão impressa referente à outra parte (não escolhida pelo eleitor) é visivelmente destruída junto a um representante das autoridades eleitorais, sendo igualmente apagada da máquina de votação. Note-se, portanto, que apenas uma das partes do voto de cada eleitor é mantida, tanto em formato impresso como digital.

Finda a votação, a totalização é realizada por meio de múltiplas rodadas de redes de misturadores (*mixnets*) operadas pelos guardiões eleitorais. A entrada inicial de tal rede de misturadores são os próprios recibos dos eleitores, que podem ser individualmente verificados para se atestar sua inclusão. Cada qual das rodadas consiste na remoção de uma camada de proteção criptográfica dos recibos por um guardião, transformando-os em uma nova imagem e adicionalmente reordenando-lhes para impossibilitar um mapeamento entre entradas e saídas. Ao final das rodadas de decifração de todos os guardiões, o conjunto final obtido corresponde às imagens dos votos conforme observadas pelos eleitores

quando da votação, mas sem qualquer vinculação para com os eleitores que as produziram. A totalização pode então ser obtida por meio de um somatório simples dos votos contidos nas imagens finais.

2.2.2 Prêt à Voter

O *Prêt à Voter* [41] é um protocolo eleitoral eletrônico criptográfico baseado nas ideias originalmente propostas pelo Votegrity, chegando inclusive a ser empregado em eleições oficiais no estado de Vitória, na Austrália. Semelhante ao Votegrity, as cédulas eleitorais preenchidas pelos eleitores são compostas de duas partes distintas que em conjunto tornam um voto plenamente visível e legível, mas que o ofuscam quando separadas. No entanto, o Prêt à Voter difere do Votegrity por abandonar a ideia do uso de cédulas cifradas visualmente (i.e. uso de imagens) para se utilizar de um mecanismo mais simples, baseado em listagens aleatórias das opções disponíveis aos eleitores.

Dessa forma, um eleitor, ao chegar ao local de votação, seleciona uma das várias cédulas disponíveis no local de votação, sendo cada qual composta por duas partes: a esquerda, contendo uma listagem em ordenação aleatória das opções disponíveis para voto; e a direita, alinhada à primeira, contendo espaços para marcação física pelo eleitor de sua opção de escolha. Após marcar sua escolha, a cédula é separada pelo próprio eleitor nos dois supracitados componentes, e a parte esquerda da cédula (contendo a listagem de opções) é descartada. A parte direita, por sua vez, contém adicionalmente uma especificação cifrada da listagem exposta na parte esquerda, permitindo, via decifração, recuperação do voto específico do eleitor. Esta parte direita da cédula é então escaneada e registrada digitalmente, e o eleitor a leva consigo como recibo eleitoral.

Finalizada a votação, também de forma semelhante ao Votegrity, o Prêt à Voter se utiliza de uma rede de misturadores cuja entrada são os recibos de voto produzidos durante a votação, o que permite verificação individual por parte de cada eleitor da presença de seu voto na composição final. A rede de misturadores do Prêt à Voter, entretanto, se processa em duas etapas consecutivas: a primeira, de anonimização das cédulas, executada via recifrações dos recibos originais; e a segunda, de decifração das cédulas, executada para reconstruir cada qual dos votos originais. Com base na listagem final dos votos recuperados, a totalização é produzida via simples contagem.

2.2.3 Wombat

O Wombat [7] une a conveniência de registros em papel e criptografia em um sistema com capacidades de verificação de votos de ponta-a-ponta (E2E-V) e que também produz durante sua execução uma trilha de evidências em lastro físico, tornando-o independente de software. Busca-se assim manter uma forma híbrida de representação do voto que confere garantias criptográficas ao sistema enquanto se mostra familiar aos eleitores acostumados com o registro de votos em papel. O sistema chegou a ser utilizado em algumas eleições em caráter oficial.

Durante a realização de uma eleição com o Wombat, o eleitor opera uma máquina de votação, à qual informa suas opções por meio de uma interface intuitiva. O sistema então imprime uma cédula composta de duas partes: um voto em texto pleno, legível e verificável pelo próprio eleitor; e um código QR que consiste em uma versão cifrada e devidamente assinada do voto legível mostrado. Nesse ponto, o eleitor informa à máquina se deseja depositar o voto produzido ou desafiá-lo. Optando pelo desafio, o sistema imprime informação adicional que permite ao eleitor verificar se a cédula produzida foi corretamente construída pela máquina. O eleitor pode repetir esse processo de desafio quantas vezes queira e, quando satisfeito, opta por depositar sua cédula. O processo de depósito consiste em separar a cédula em seus dois componentes, texto pleno e voto cifrado, e depósito da porção em texto pleno. A porção em texto cifrado é escaneada por um representante das autoridades e entregue ao eleitor como recibo de seu voto.

Semelhante ao Prêt à Voter, o Wombat opta por se utilizar, como parte do processo de totalização, de redes de misturadores para anonimização dos votos via recifração destes. O sistema deixa a cargo de decisões das autoridades ou operadores do sistema a forma como a trilha de evidências será tratada: por exemplo, se os registros físicos serão absolutamente totalizados juntamente com os registros cifrados dos votos para fins comparativos ou se os registros físicos serão apenas verificados por amostragem para fins de construção de um percentual de certeza acerca dos resultados eleitorais produzidos. Igualmente aos demais sistemas descritos, o recibo entregue ao eleitor lhe permite verificar a presença de seu voto na totalização, a qual é produzida pela decifração dos votos cifrados já anonimizados.

2.2.4 STAR-Vote

O STAR-Vote [5] é um sistema nascido de uma colaboração entre a academia e membros do cartório eleitoral do condado de Travis, em Austin, no Texas (EUA). O condado já havia passado por experiências eleitorais prévias com outros sistemas, inclusive DREs, e com base nestas a administração eleitoral concluiu não haver nenhum sistema disponível para uso que fosse capaz de satisfazer a combinação particular de requisitos do condado. Dentre alguns desses requisitos específicos, alguns são de notável menção tanto devido a seu impacto particular no projeto quanto a sua similaridade com o contexto brasileiro:

- **Inteface DRE:** o uso de cédulas físicas abre margem a interpretações dúbias referentes a intenções de voto, algo que pode ser mitigado via uso de uma interface DRE. Além disso, o uso de computadores como interface com o eleitor apresenta outras claras vantagens, como tornar desnecessária a impressão prévia de cédulas físicas para marcação manual e facilitar o acesso a eleitores com deficiências. Trazendo este aspecto para uma interpretação no contexto brasileiro, o uso de uma interface DRE se aproveitaria das décadas de familiaridade do eleitor brasileiro com esses dispositivos, além de também facilitar o acesso dos vários eleitores com analfabetismo funcional pelas capacidades visuais que provê.

- **Aparato Comercial de Uso Geral:** equipamento eleitoral especializado, além da complexidade que lhe é própria, apresenta ainda custos surpreendentemente altos. Outro fator digno de menção é que dispositivos comerciais dessa natureza, tanto no que se refere a hardware quanto a software, frequentemente encontram-se sujeitos a imposições de sigilo de funcionamento sob a premissa de *"segredos comerciais"*, o que resulta na imposição de sérias restrições à transparência do sistema, além de dificultar procedimentos de análise e auditoria independentes. Assim sendo, o uso de aparato comercial de propósito geral mostra-se como uma alternativa de considerável potencial redutor dos custos, da complexidade e dos ciclos de atualização dos equipamentos. No Brasil faz-se uso de máquinas de propósito especificamente eleitoral e os referidos problemas associados ao uso de tais equipamentos se mostram na prática: em 2020 a justiça eleitoral aprovou a aquisição de até 180.000 novas urnas eletrônicas em uma licitação no valor de $R\$799,9$ milhões, no intuito de substituir urnas antigas fabricadas em 2006 e 2008 [60], custo consideravelmente alto devido à suposta especialização do aparato adquirido.

- **Cédulas Extensas:** o número de cargos e a quantidade de candidatos a estes influencia diretamente o tamanho das cédulas: quanto mais cargos e candidatos a estes,

maiores ficam as cédulas. Este é o contexto de aplicação do STAR-Vote, que suporta extensas listagens em suas cédulas. Este também é um requisito tradicionalmente associado à realidade eleitoral brasileira, em que a grande quantidade de candidatos a cada cargo inviabiliza qualquer possibilidade de produção de versões impressas da cédula para fácil marcação por eleitores. Em 2020, por exemplo, houve um total de 556.033 pedidos de registro de candidatura aos cargos de prefeito, vice-prefeito e vereador [56, p.24].

Esse conjunto de requisitos e restrições resultou em um interessante projeto de sistema eleitoral híbrido que provê capacidades E2E-V em conjunto com a produção de VVPATs. Os registros em papel permitem uma forma tangível e visual de identificação do voto produzido pelo eleitor, além de servirem como uma redundância física dos votos para fins de auditoria e recontagem em caso de necessidade. Já as capacidades criptográficas do sistema permitem ao eleitor acompanhar seu voto por todo o fluxo eleitoral, assegurando-se de que suas opções estão sendo corretamente consideradas na composição do resultado final.

O fluxo de votação do sistema é alimentado por um conjunto de máquinas isoladas em uma rede interna, sem qualquer via de acesso externo. O fluxo se inicia com a identificação do eleitor junto a um representante das autoridades eleitorais e com o auxílio de uma máquina conectada à internet. Esta conexão à internet permite consultas a uma base de dados eleitoral externa, viabilizando a devida verificação do eleitor e evitando votos em duplicidade. Por estar conectada à internet, esta máquina encontra-se necessariamente desconectada da rede interna montada no local de votação.

Devidamente identificado, o eleitor recebe um código de barras unidimensional que identifica sua seção de votação e o *estilo* de sua cédula: um dos requisitos do STAR-Vote é que qualquer eleitor do condado pudesse votar em qualquer dos locais disponíveis para votação, de modo que a composição das cédulas de cada eleitor é variável. Note-se assim que este código de barras não apresenta nenhuma informação sigilosa ou única ao eleitor em particular. O eleitor então dirige-se a uma máquina controladora conectada à rede interna e escaneia o código de barras recebido: esta é a única informação externa que adentra a rede local. A máquina controladora então gera um código numérico aleatório de 5 dígitos e o imprime em papel para o eleitor.

O eleitor finalmente procede para um dos terminais de votação, o qual apenas é ativado via fornecimento de um código de 5 dígitos válido, como o previamente recebido pelo eleitor. Apenas uma pequena quantidade desses códigos encontra-se ativa em um curto intervalo de tempo, o que dificulta tentativas aleatórias de acertar um código válido, além de inviabilizar tentativas de reuso do mesmo. O eleitor então faz suas escolhas via interface DRE do terminal, sendo apresentado a uma tela final de confirmação que o permite prosseguir com seu voto ou corrigi-lo, podendo refazer suas escolhas. Ao confir-

mar, o terminal de votação produz uma versão digital do voto cifrada homomorficamente em acordo com a metodologia descrita por [13], sendo tal versão armazenada em caráter temporário na rede interna.

Adicionalmente à versão digital cifrada, o terminal de votação imprime em papel uma versão física do voto dividida em dois artefatos: o primeiro, composto pelas escolhas legíveis do eleitor e um código numérico aleatório que identifica o voto em si; e o segundo, composto por meta-informações, como o identificador do terminal usado e horário de votação, junto a um *hash* do voto cifrado. O primeiro permite ao eleitor conferir, de maneira visual e tangível, se suas opções digitais correspondem ao que se encontra fisicamente impresso. Já o segundo opera como um compromisso (do inglês *commitment*) impresso da máquina para com o voto digital cifrado do eleitor. A versão digital é mapeada à versão física do voto por meio do código numérico aleatório impresso e mantido também na rede interna.

O eleitor pode então optar por depositar seu voto ou desafiar o sistema a mostrar que as duas versões, digital e física, correspondem uma à outra corretamente. Optando pelo desafio, o eleitor se dirige a um representante das autoridades eleitorais, que escaneia o código identificador aleatório do voto para informar à rede interna que o voto está sendo desafiado. O sistema então decifra a versão cifrada do eleitor, mostrando informações adicionais que permitem ao eleitor verificar a correspondência entre o voto digital e a versão impressa. Tendo optado pelo desafio, o voto previamente produzido é considerado agora inválido, devendo ao eleitor ser dada a oportunidade de produzir nova versão do voto com as escolhas de sua preferência. Os votos desafiados são mostrados também na página de resultados eleitorais como evidências adicionais do correto funcionamento do sistema, mas não são incluídos na contabilização da totalização final.

Estando satisfeito com o processo de desafio, o eleitor pode então optar pelo depósito de seu voto. Para isso, o eleitor segue para uma urna próxima e nela insere seu voto impresso, que tem seu número identificador escaneado, com isto informando à rede interna que o referido voto está sendo depositado e, portanto, fará parte da totalização. Note-se que o processo faz com que duas versões do voto sejam incluídas no montante de votos válidos: a versão digital cifrada, armazenada em um montante digital, e a versão física impressa, que possibilita verificações estatísticas e possíveis recontagens, caso necessárias. O eleitor leva consigo o segundo artefato impresso, contendo o *hash* de seu voto cifrado, que opera então como um recibo deste.

Finda a votação, as autoridades eleitorais divulgam a listagem dos votos cifrados em alguma mídia publicamente acessível (e.g. uma página web). O eleitor pode então se assegurar da presença de seu voto na totalização verificando se o voto correspondente ao *hash* em seu recibo se encontra listado. A totalização é realizada via composição homomórfica das versões digitais dos votos, sendo decifrado apenas o produto final de tal composição, o qual é matematicamente demonstrável corresponder ao resultado eleitoral.

Note-se assim que nenhum dos votos cifrados é decifrado em caráter individual, de modo que o eleitor não tem seu sigilo prejudicado no processo.

O STAR-Vote se utiliza ainda de processos de auditoria de limitação de riscos (*risk-limiting audits (RLA)*) [25], procedidos sobre os votos impressos no intuito de se garantir a correção estatística do resultado eleitoral obtido. Nesta, uma quantidade estatisticamente significativa dos votos físicos depositados é selecionada em caráter aleatório para verificação da correta correspondência entre as versões físicas e as versões cifradas desses votos e a produção de uma totalização parcial com base nestes. Esse mecanismo serve como uma garantia adicional da devida lisura do processo e da integridade dos resultados conforme obtidos e publicados.

Outro ponto digno de nota é o uso da rede interna pelo STAR-Vote como um mecanismo de manutenção de estado entre a máquina controladora, os terminais de votação e a urna. Isto permite comunicação entre estas máquinas no sentido de sincronizar ações dentro do fluxo de votação, como bem demonstrado pelos procedimentos de desafio e depósito.

Uma importante observação ilustrada pelo projeto do STAR-Vote é a de que sistemas eleitorais são fortemente sensíveis ao contexto pretendido de sua aplicação: os requisitos particulares do condado de Travis tiveram impacto direto na composição e organização finais do sistema desenvolvido, sendo determinantes para o que seria ou não aceitável ao projeto. Isto explicita que um sistema eleitoral, ainda que satisfatoriamente operacional em determinado contexto, pode se mostrar disfuncional quando levado a operar em um contexto distinto. Assim sendo, para além dos requisitos gerais inerentes a sistemas eleitorais, aqueles específicos ao contexto também interagem e contribuem de maneiras complexas, guiando o projeto dos mais diversos aspectos do sistema.

2.3 O Sistema Eleitoral Brasileiro

Desde 1996, a Justiça Eleitoral Brasileira, sob a figura central do *Tribunal Superior Eleitoral (TSE)*, vem adotando um projeto de sistema eleitoral crescentemente automatizado. Naquela ocasião, uma fração considerável do eleitorado, compreendendo mais de 32 milhões de eleitores, depositou seu voto em alguma das mais de 70 mil urnas eletrônicas colocadas em operação [45]. Este foi o marco público inicial de um projeto de substituição das urnas de lona usadas até então, feitas para receber votos em cédulas de papel e alvo de desconfiança de parte da população brasileira. Já no ano 2000, as eleições municipais foram realizadas em caráter puramente eletrônico, sendo todas as eleições a partir de então, tanto municipais quanto nacionais, executadas com lastro unicamente digital do

voto.

Inicialmente se utilizando de sistemas proprietários e com uma cultura de desenvolvimento absolutamente opaca à sociedade civil, a Justiça Eleitoral, sob pressão da comunidade técnica, migrou para sistemas GNU/Linux em 2008 e, a partir de 2009, adotou um modelo de testagem pública periódica dos sistemas desenvolvidos internamente, com edições geralmente realizadas proximamente à execução de novas eleições nacionais [55]. No mesmo ano de 2008 pôs-se em curso também o projeto de identificação biométrica do eleitor, com expectativa de cobertura biométrica total do eleitorado para as eleições nacionais de 2022 [52].

2.3.1 Composição e Funcionamento do Sistema

O Tribunal Superior Eleitoral desenvolve um conjunto consideravelmente complexo e diverso de softwares que compõem o sistema eleitoral eletrônico brasileiro, dentre os quais aqueles intrinsecamente associados a etapas críticas como o fluxo de votação e a totalização dos resultados, sendo tais artefatos de software responsáveis por manusear dados cruciais ao sigilo do eleitor e à integridade dos resultados. Em publicação própria do ano de 2014, o TSE informava que sua coleção total de sistemas digitais compreendia 60 artefatos distintos de software de finalidades variadas e ecossistemas distintos [47]. Desses artefatos de software, são listados um total de 28 como parte do ecossistema da urna eletrônica em si [61].

Tais componentes são mantidos e desenvolvidos internamente pela Justiça Eleitoral Brasileira, em caráter proprietário, sendo abertos para auditoria e/ou testes apenas em períodos determinados de tempo e sob condições ditadas pelo próprio TSE. Os sistemas desenvolvidos têm sua base de código supostamente *congelada* (i.e. mantida em um estado temporário em que não recebe modificações) para a realização de auditorias e testes. Posteriormente, tal base de código é modificada para acomodar eventuais correções apontadas pelos testes antes de novo congelamento prévio a um período eleitoral.

Uma vez que o software eleitoral seja considerado suficientemente maduro para ser posto em produção em uma eleição vindoura, uma versão deste é enviada aos *Tribunais Regionais Eleitorais (TREs)*, sendo armazenada em cartões de memória denominados *flashes ou mídias de carga*. Essas mídias de carga são as responsáveis pela instalação dos softwares: como a urna eletrônica brasileira não apresenta interface de rede, a comunicação com o dispositivo é feita com acesso físico e por intermédio de mídias plugáveis. Cada mídia de carga é reportadamente responsável pela instalação do software em algumas dezenas de urnas.

Após instalação, as interfaces plugáveis do dispositivo são lacradas com lacres supostamente invioláveis e as urnas são armazenadas sob controle da Justiça Eleitoral para, às vésperas de uma eleição, serem transportadas para as respectivas seções nas quais serão usadas [65]. Já no dia da eleição, antes de serem utilizadas para votação, as urnas são comandadas a imprimir um registro físico denominado *zerésima*. Este registro é emitido logo após o procedimento de inicialização da urna eletrônica, e supostamente atesta que a máquina não possui qualquer registro de voto nela armazenado até então [64] [57, p.11]

Na etapa de votação, cada eleitor é individualmente autenticado, procedimento necessário no intuito de se verificar se o mesmo se encontra habilitado a votar. O procedimento é executado por um representante das autoridades eleitorais por meio de um dispositivo denominado *terminal do mesário*. Tal dispositivo é diretamente ligado por um cabo ao chamado *terminal do eleitor*, o qual é usado pelo eleitor para informar seu voto à máquina. A justiça eleitoral deixa claro que o supracitado cabo é usado como canal de comunicação de via dupla entre o terminal do eleitor e o terminal do mesário [66] [57, p.11].

Durante a etapa de votação, cada voto depositado é registrado apenas digitalmente em uma tabela mantida pela urna e denominada *registro digital do voto (RDV)* [63], que supostamente também provê um mecanismo de recontagem dos votos caso esta se faça necessária. No intuito de resguardar o sigilo de cada eleitor, essa tabela é mantida embaralhada, o que, em meio digital, implica na utilização de alguma metodologia de geração de números aleatórios. O RDV foi implantado em 2003 como um mecanismo substitutivo ao voto impresso e produzido de maneira tal a não comprometer o sigilo do eleitor [46] [9], embora a edição dos testes públicos de 2012 tenha demonstrado uma situação consideravelmente distinta, em que, via manutenção de uma ordenação externa dos eleitores de uma seção, seria possível e trivial produzir um mapeamento entre eleitores e votos depositados devido à metodologia infantil de embaralhamento utilizada [3].

Ainda durante a votação, urnas eletrônicas são selecionadas em caráter aleatório para serem submetidas a um processo simulado e controlado de votação denominado *votação paralela* [49]. Nesse processo, uma urna selecionada recebe votos idênticos aos reais, mas preparados em cédulas físicas e sem premissa de sigilo. Tais votos preparados são então informados tanto à urna quanto a um segundo sistema eletrônico de apoio desenvolvido exclusivamente para fins de auditoria, de modo que são realizadas duas eleições paralelas sobre o mesmo conjunto conhecido de votos. Ao final, ambas estas eleições, na urna e na máquina de apoio, são totalizadas, devendo ambos os resultados se mostrarem idênticos. A argumentação para tal procedimento é o de que a observação de um comportamento honesto por parte das urnas selecionadas supostamente atesta que as demais máquinas utilizadas no pleito, com grande probabilidade, se comportarão também de maneira honesta, o que atestaria a integridade dos resultados reais obtidos ao final da votação.

Finda a votação, o encerramento do sistema é realizado por operação manual do terminal do mesário em conjunção com a operação do terminal do eleitor. Nesse processo os votos registrados são agregados em totalizações parciais, as quais são impressas em cópias do chamado *boletim de urna (BU)*. Além de diversos totais referentes aos votos depositados na urna em questão, este documento registra ainda informações de identificação de urna e seção e o horário de encerramento da eleição [62] [57, p.14]. Outros dois boletins são também impressos: o *boletim de justificativas (BUJ)*, contendo justificativas de ausência de eleitores, e o *boletim de identificação de mesários (BIM)*, contendo o registro dos mesários presentes na seção eleitoral. Um dos boletins de urna deve ser afixado na porta da própria seção eleitoral.

Impressos os boletins, os mesários são então instruídos a remover uma mídia específica, denominada *mídia de resultados (MR)*, que fica acoplada à urna eletrônica. Além de uma cópia digital assinada do próprio boletim de urna, esta mídia armazena diversos artefatos digitais produzidos durante a operação da urna no dia de votação. A mídia de resultados é então fisicamente transportada para um *centro de transmissão*. Neste centro as mídias de resultados são acopladas a máquinas da justiça eleitoral e os resultados parciais são então transmitidos, via rede privada, para totalização final. Aparentemente a estratégia de totalização varia entre eleições: como exemplo, em 2020 a totalização foi centralizada no próprio TSE, enquanto em eleições anteriores os *TREs* conduziam suas totalizações regionais [58].

2.3.2 Os Testes Públicos

Conforme menção prévia, a pressão da comunidade técnica fez com que, em 2009, o TSE adotasse um procedimento de testagem pública que permitisse verificação do sistema por parte de interessados externos ao âmbito da própria justiça eleitoral. Com isto passaram a ser realizados os chamados *Testes Públicos de Segurança (TPS)*, eventos periódicos, geralmente realizados proximamente aos períodos eleitorais, nos quais alguns artefatos do sistema eleitoral são abertos à análise de membros da comunidade acadêmica, peritos e demais interessados que satisfaçam determinadas restrições [44].

O objetivo destes eventos é permitir que os participantes tentem executar planos de ataque à urna eletrônica brasileira e alguns de seus componentes no intuito de demonstrar a viabilidade prática de adulterações no sistema, em seu funcionamento ou em artefatos produzidos por sua execução. Para participar, os interessados se inscrevem junto ao TSE e, tendo aprovada sua inscrição, devem se submeter a um termo de confidencialidade antes de visualizar qualquer tipo de código-fonte do sistema.

Cabe aqui mencionar que se mostra como fator estranho a um processo público de testagem a assinatura de um termo de tal natureza, haja vista o efeito nocivo de tal exigência sobre possíveis divulgações e análises técnicas de achados por parte dos participantes do processo de testagem, ficando a publicação destas sob domínio de análise da própria justiça eleitoral. Ademais, as justificativas dadas pelo próprio TSE para a existência de tal termo [53] lastreiam-se em argumentações que vão contra preceitos básicos da segurança computacional contemporânea, ignorando aspectos estabelecidos, como o *Princípio de Kerckhoffs* [33], em favor de um projeto lastreado na chamada *"segurança por obscuridade"* [21, p.8], noção antiquada e historicamente demonstrada como falha a menos de possível e limitada valia em contextos extremos (e.g. situações emergenciais de guerra).

Aceitando se submeter ao termo de confidencialidade obrigatório, o participante ganha finalmente acesso, por alguns dias, aos códigos-fonte dos artefatos que compõem o sistema eleitoral, um universo de dezenas de milhões de linhas de código de distintos sistemas espalhadas por uma coleção de milhares de arquivos. Com base nessa análise limitada, o participante deve construir uma bateria de planos de testes e submetê-la ao TSE para uma análise por meio da qual a justiça eleitoral define quais serão aceitos ou não com base em critérios relativamente abstratos como "viabilidade técnica", "clareza suficiente" e "atendimento de objetivos específicos", os quais são definidos pela própria autoridade eleitoral [54, p.9, 14].

Um fator limitante crítico associado aos testes, imediatamente depreendido de suas supracitadas descrições, é que estes seguem um modelo fortemente restritivo, totalmente controlado pela própria justiça eleitoral, que reserva-se ainda ao direito de excluir componentes do próprio sistema eletrônico do processo de testagem, a exemplo da identificação biométrica do eleitor, procedimentos de processamento de arquivos das urnas e artefatos referentes à totalização dos resultados, além de outros [54, p.2]. Outras possibilidades de exclusão também são previstas, como aquelas baseadas nas propostas de planos de testes [54, p.9,14] e na nacionalidade dos adversários do sistema [51, p.18]. Restrições de tais naturezas, entretanto, não refletem as condições de um ataque realista ao sistema por parte de adversários competentes, em que a questão de viabilidade técnica de um ataque é sensível ao aporte financeiro, técnico ou político do atacante e a nacionalidade do mesmo é fator indeterminante no que se refere a sua capacidade de compreender o código.

Não obstante tal modelo restritivo, as poucas edições dos testes públicos já foram suficientes para revelar falhas de segurança no sistema eleitoral que são de significativa importância. Na edição de 2012, dentre demais achados comprometedores, foi descoberta uma vulnerabilidade grave associada ao registro digital do voto, o qual então fazia uso criptográfico infantil de um gerador impróprio a tal fim para embaralhar os votos armazenados. Tal vulnerabilidade comprometia o sigilo do voto de maneira tal que possibilitava um mapeamento total entre votos e eleitores de uma mesma seção eleitoral fazendo uso

apenas de informação que poderia ser obtida em caráter público, ou seja, sem necessidade de invasões ao sistema ou adulterações do software [3].

Já na edição de 2017, uma possibilidade crítica de ataque ao sistema da urna eletrônica foi descoberta, oriunda da decisão de projeto duvidosa de armazenar chaves criptográficas dentro do próprio código-fonte. A disponibilidade de uma chave criptográfica em texto-pleno dentro das mídias de instalação da urna permitiu inspeção de seus conteúdos e detecção de duas bibliotecas cujas assinaturas digitais não eram corretamente checadas, o que por sua vez permitiu inclusão e execução de código arbitrário pela urna eletrônica [2]. Dada a decisão de projeto de compartilhar, entre todas as urnas usadas, tal chave criptográfica trivialmente descoberta, a vulnerabilidade detectada viabiliza tecnicamente a construção de um ataque facilmente escalável e de âmbito generalizado a tais dispositivos.

É importante mencionar que a prática de armazenamento de chaves criptográficas em código-fonte já havia sido flagrada em artefatos do sistema eleitoral em ocasiões prévias [3] [36] [48], tendo as autoridades eleitorais sido devidamente alertadas com relação aos perigos associados a tal decisão de projeto. Desse modo, a recidiva observada denota aparente incapacidade ou indisposição por parte dos mantenedores do sistema de seguir orientações básicas de ordem técnica e científica com relação às práticas de segurança adotadas no desenvolvimento e manutenção dos artefatos de software que compõem o sistema eleitoral brasileiro.

2.3.3 As Auditorias Pré-Eleitorais

Seis meses antes da realização de eleições, o TSE oferece a possibilidade de verificação dos softwares que serão utilizados no pleito. De caráter muito mais restritivo que os testes públicos de segurança, essas auditorias pré-eleitorais apenas permitem representantes indicados por entidades fiscalizadoras seletas [49, Artigo 8º]. Ademais, são restritos quaisquer tipos de testes sobre os hardwares ou softwares, de modo que o processo se assemelha mais a uma inspeção pura do que a um processo rigoroso de auditoria em si. Assim sendo, qualquer tipo de tentativa de análise dos programas implementados pela justiça eleitoral tem de se limitar à leitura das dezenas de milhões de linhas de código-fonte em si, sem possibilidade de compilação ou execução do código.

A auditoria pré-eleitoral, no entanto, resguarda algumas similaridades com os testes públicos. Nela, os inspetores são também obrigados a se submeter a um termo de confidencialidade previamente ao primeiro acesso aos códigos-fonte, cabendo aqui as mesmas críticas já notadas para tal requerimento no contexto dos TPSs. Outra semelhança

reside na determinação de que os próprios inspetores são vedados de fazer qualquer tipo de anotação ou consulta a manuais durante o processo de fiscalização. Assim, fica aqui também claramente ilustrado o rígido controle do processo de inspeção, controle este definido e exercido por parte da própria entidade fiscalizada [49, Artigo 9º].

2.3.4 A Única Auditoria Pós-Eleitoral

Em 2014, após uma eleição acirrada e uma vitória presidencial por pequena margem pelo Partido dos Trabalhadores (PT), o Partido da Social Democracia Brasileira (PSDB), perdedor do pleito, formalizou um pedido de auditoria pós-eleitoral ao TSE. Pedidos de auditoria dessa sorte são hoje citados pela justiça eleitoral como um dos mecanismos de atestação de transparência e integridade oferecidos para se averiguar a devida lisura do pleito finalizado [59], mas nenhuma auditoria do tipo tinha sido feita até então, de modo que esta se mostrava como algo inédito.

Embora a auditoria de 2014 demonstrasse motivações políticas indiscutíveis, a possibilidade de realização de um procedimento de pós-auditoria de viés independente é requisito crucial de um sistema eleitoral que se pretenda minimamente transparente. No entanto, a solicitação feita pelo PSDB foi mal-vista por parte da sociedade, interpretada como tentativa de reversão dos resultados do pleito, por isso tendo sido intensamente criticada [26].

Não obstante a motivação política e a resistência enfrentada, o procedimento resultou em um detalhado relatório técnico de auditoria, o qual chegou a uma conclusão consideravelmente interessante com relação à implementação eleitoral eletrônica brasileira: o sistema é inauditável e permite apenas resultados inconclusivos no referente a possíveis manipulações ou fraudes [36]. Dentre as diversas críticas listadas no documento produzido encontram-se o controle excessivo do processo de auditoria por parte das próprias autoridades eleitorais (algo definido no relatório como uma espécie de *"auditoria comandada pelo auditado"*), a imaturidade do processo interno de desenvolvimento do sistema, a absoluta negligência para com a possibilidade de ataques de origem interna, a desatenção a padrões internacionais de segurança, decisões de projeto questionáveis (e.g. armazenamento de chaves criptográficas dentro do próprio código-fonte) e, essencialmente, um sistema final que ignora por completo o princípio da independência de software [36, p.209-211].

Apesar das diversas críticas ao sistema e à condução restritiva da auditoria, juntamente com o veredicto de inconclusividade apresentado pelos auditores, o TSE manifestou-se a respeito de maneira consideravelmente distinta, focando-se na aparente argumentação

de que a ausência de evidências cabais de fraude denotaria prova da suposta lisura do processo eleitoral, postura até hoje mantida pelas autoridades e reforçada via propaganda extensiva junto à sociedade civil. Para além disto, a própria justiça eleitoral fez questão de transparecer seu desagrado para com o processo de auditoria, dedicando uma sessão de sua manifestação pública a um suposto caráter oneroso do processo, contando inclusive com referência a uma frase proferida por um representante do Ministério Público Eleitoral (MPE), que dizia *"para que se saiba quanto custa essa aventura para que ela não se torne uma rotina na Justiça Eleitoral brasileira"* [50].

Essa experiência única serve como demonstrativo de que um grande desafio referente ao procedimento de auditoria pós-eleitoral encontra-se em sua ligação temporal intrínseca com o pleito recém-finalizado, o que leva tal procedimento a ser confundido pela sociedade civil como mera tentativa de reversão dos resultados eleitorais. Tal visão é reforçada pelo fato de que aqueles que geralmente se apresentam como os maiores interessados na realização de auditorias dessa natureza são precisamente os partidos e políticos perdedores do pleito. Isto posto, a normalização da auditoria pós-eleitoral como procedimento recorrente de verificação necessário à transparência eleitoral e à averiguação da integridade de seus resultados, com engajamento dos diversos partidos políticos, inclusive os vencedores, pode auxiliar na mitigação de tal sensação. Ademais, a atitude negativa demonstrada pela justiça eleitoral para com o procedimento requer séria revisão da postura interna da entidade e é claramente incompatível com a forma como a mesma costuma propagandeá-lo.

2.3.5 Resumo Crítico do Sistema

O modelo de votação eletrônica implantado no Brasil a partir de 1996 foi um grande marco democrático para o país, auxiliando na estabilização da então jovem e imatura democracia nacional ao substituir um sistema de votação em cédulas de papel acerca do qual havia grande descrédito por parte da população. Entretanto, analisando-se a situação sob uma ótica contemporânea, o que se revela na atual realidade eleitoral brasileira são os contornos de um sistema eletrônico de caráter obscurantista, vulnerabilizado por decisões de projeto questionáveis e modelo de ataque insuficiente, que persiste na manutenção de um lastro puramente eletrônico combinado a metodologias de desenvolvimento e auditoria predominantemente monocráticas com um viés de ignorância aos avanços científicos da área e às orientações técnicas daqueles que o avaliam.

Os artefatos gerados pela urna eletrônica, rotineiramente citados como atestados da lisura da operação da máquina, na prática mostram-se insuficientes nesse sentido, isto

predominantemente por conta da decisão de projeto de se manter um sistema de natureza puramente eletrônica. Em termos de tais artefatos, a situação do sistema eleitoral presente é a que se segue:

- A zerésima, primeiro documento produzido pela própria urna eletrônica após sua inicialização, é citada como um atestado da honestidade da máquina ao informar que esta não possui, em seus registros, votos armazenados para quaisquer candidatos. Tal argumento, entretanto, não se mantém em sentido técnico: uma urna desonesta é trivialmente capaz de produzir uma zerésima falsa na qual mente a respeito de sua suposta idoneidade de modo plenamente convincente. Isto posto, tal documento nada prova além da capacidade da máquina de produzir uma zerésima compatível com aquela que seria esperada de uma máquina honesta.

- O registro digital do voto (RDV) é um artefato digital no qual os votos recebidos pela urna eletrônica são armazenados de forma embaralhada. É também citado como um mecanismo de transparência que auxilia em possível recontagem dos votos em caso de necessidade. No entanto, por ser produzido e mantido sob controle da própria máquina que registra os votos, sofre de problemas similares aos da zerésima no sentido de que uma máquina desonesta pode produzir um RDV inconsistente com os votos por ela recebidos e ainda assim passar uma falsa impressão de integridade do registro. Ademais, o RDV, se indevidamente implementado como bem ilustraram os testes públicos de 2012, pode se mostrar como um mecanismo fragilizador do sigilo do voto. A possibilidade de atuação desonesta sobre o RDV poderia ser tornada detectável pela introdução de um registro físico do voto na forma de um *VVPAT*, dificultando adulterações imperceptíveis dos votos. Isto posto, a despeito das declarações do TSE e das alegadas motivações para sua introdução, o RDV claramente se mostra como mecanismo incapaz de atuar como um substituto do lastro físico do voto.

- O boletim de urna (BU) é um registro físico impresso pela urna após o encerramento da votação, apresentando totais diversos supostamente referentes aos votos registrados pela máquina. Dada sua origem digital e de impressão controlada pela própria urna eletrônica, tal documento segue uma natureza semelhante à da zerésima e do RDV: é plenamente plausível a uma máquina desonesta produzir um BU totalmente compatível com um RDV armazenado, mas que não representa a realidade dos votos fornecidos ao dispositivo. Novamente, o problema principal reside na inexistência de um mecanismo alheio à máquina que confira lastro físico e independente de software aos votos registrados em uma seção eleitoral.

No referente ao procedimento de votação paralela, sua realização não é capaz de simular com precisão as condições reais de votação, isto por conta de seu caráter controlado

com vistas puramente à conferência dos resultados finais. Como exemplo, é tecnicamente possível se implementar em código uma sequência oculta de comandos no software das urnas que poderia ativar um comportamento indevido por parte destas. Note-se que tal sequência oculta dificilmente seria detectada pelo procedimento de votação paralela. É interessante observar, entretanto, que a justiça eleitoral, no parágrafo único do artigo 69 de sua resolução de número 23.603, aparenta admitir a importância de um lastro físico para os votos ao prever, em caso de divergências de resultados, a recontagem por meio da conferência minuciosa das cédulas físicas fornecidas aos sistemas eletrônicos de votação [49].

Por sua vez, os processos de auditoria, tanto pré quando pós-eleitoral, são fortemente dominados e limitados pela própria justiça eleitoral. As limitações impostas por vezes não fazem sentido prático, mostrando-se muito mais como formas de prejuízo gratuito ao trabalho dos auditores envolvidos. Além disso, as descobertas reveladas pela auditoria, mesmo aquelas de impacto catastrófico à íntegra e correta operação do sistema, são costumeiramente desmerecidas pelo TSE como irrelevantes ou de impacto desprezível, em clara atitude de resistência às críticas e contribuições oferecidas pela comunidade técnica e científica.

Por fim, o TSE necessita urgentemente rever o modelo de ataque que motiva as decisões de projeto dos softwares de seu sistema eleitoral. Há clara e completa ignorância a qualquer possibilidade de ataque de origem interna, seja por parte de um membro ou funcionário do próprio tribunal, seja por parte de alguém com recursos suficientes para de alguma forma obter acesso interno. Tal desconsideração é patente na prática recorrente de armazenamento de chaves criptográficas e senhas críticas no interior de códigos-fonte diversos, dos quais podem ser trivialmente obtidas por qualquer um que a eles tenha acesso. Tal omissão mostra-se particularmente preocupante quando a história recente demonstra que ataques na tentativa de influenciar resultados eleitorais nacionais são empreendidos ou financiados por entidades com considerável aporte financeiro, robusto conhecimento técnico e forte capacidade de penetração política [40].

Ademais, o próprio Tribunal Superior Eleitoral demonstra que suas afirmações auto-proclamadas da existência de sistemas *"100% seguros"* e *"invioláveis"* não encontram respaldo mesmo em seus próprios sistemas internos, vítimas de ataques cibernéticos no ano de 2020 [17]. Apesar de tais ataques supostamente não terem afetado os sistemas ou resultados eleitorais do pleito daquele ano, servem como demonstrativo adicional da falsidade do mito do sistema inviolável. E, apesar da insistência do TSE em tentar vender a urna eletrônica como um tal sistema, os resultados dos diversos testes e auditorias demonstram que tais afirmações tampouco se aplicam ao sistema eleitoral eletrônico vigente no Brasil.

Isto posto, apesar das claras mazelas do sistema eleitoral brasileiro, estas apenas constituem reflexos de uma já exposta constatação científica: sistemas eleitorais eletrô-

nicos são consideravelmente complexos, com requisitos próprios que muitas vezes atuam de forma contraditória em seu projeto e passíveis de erros ou adulterações que podem influenciar resultados eleitorais de maneira imperceptível. Disto deriva a preconização, formalizada pelo *princípio da independência de software*, de que um sistema eleitoral pode fazer uso das benesses oferecidas por maquinário eletrônico, mas necessitam de mecanismos a este extrínsecos no sentido de aferir a integridade de seus resultados de maneira que fraudes, caso ocorridas, sejam passíveis de detecção.

Em tal sentido, portanto, o presente trabalho se mostra como uma possibilidade ilustrativa de como o atual sistema eleitoral eletrônico brasileiro poderia se munir de garantias aferíveis de integridade e sigilo lastreadas em premissas criptográficas e hipóteses sólidas para viabilizar a realização de eleições independentes de software e passíveis de comprovação de seus resultados.

Capítulo 3

Preliminares Criptográficas

Antes de proceder à descrição do protocolo em si, este capítulo apresenta conceitos matemáticos e criptográficos considerados centrais tanto para a construção quanto para o entendimento das propriedades associadas ao protocolo eleitoral proposto por neste trabalho. Tópicos e definições considerados centrais ao protocolo são explorados em maiores detalhes, e resultados úteis a eles associados, os quais serão evocados posteriormente, são demonstrados. Em contraste, alguns resultados algébricos já amplamente disponíveis na literatura são simplesmente postulados ou mencionados sem o empreendimento de esforço comprobatório adicional.

3.1 Grupos Algébricos e Logaritmo Discreto

Um *grupo* é um construto algébrico abstrato de definição relativamente simples, dado por um par ordenado $(\mathbb{G}, \cdot)$, em que $\mathbb{G}$ é um conjunto e '$\cdot$' é uma *operação binária* definida entre elementos de $\mathbb{G}$, i.e., uma função $\mathbb{G} \times \mathbb{G} \to \mathbb{G}$. O conceito é mais precisamente formalizado na Definição 3.1.1 que se segue, na qual é enfatizado o possivelmente mais sutil aspecto de *fechamento* associado à operação '$\cdot$':

Definição 3.1.1 (Grupo)**.** *Sejam $\mathbb{G}$ um conjunto e '$\cdot$' uma operação binária definida sobre elementos de $\mathbb{G}$. O par ordenado $(\mathbb{G}, \cdot)$ é um* ***grupo*** *se a operação '$\cdot$' satisfaz as seguintes propriedades:*

- ***Fechamento:*** $\forall a, b \in \mathbb{G} : a \cdot b \in \mathbb{G}$
- ***Existência de identidade:*** $\exists e \in \mathbb{G} : \forall a \in \mathbb{G} : a \cdot e = e \cdot a = a$
- ***Existência de inversos:*** $\forall a \in \mathbb{G} : \exists a^{-1} \in \mathbb{G} : a \cdot a^{-1} = a^{-1} \cdot a = e$
- ***Associatividade:*** $\forall a, b, c \in \mathbb{G} : (a \cdot b) \cdot c = a \cdot (b \cdot c)$

Se um grupo $(\mathbb{G}, \cdot)$ adicionalmente satisfaz a propriedade de *comutatividade* com relação à sua operação, então tal grupo é chamado *abeliano*, conceito formalizado na Definição 3.1.2 seguinte:

Definição 3.1.2 (Grupo Abeliano)**.** *Seja um grupo* $(\mathbb{G}, \cdot)$ *tal que a operação* '$\cdot$' *satisfaz também a seguinte propriedade:*

- ***Comutatividade:*** $\forall a, b \in \mathbb{G} : a \cdot b = b \cdot a$

  Nesse caso, o grupo $(\mathbb{G}, \cdot)$ *é dito ser um* ***grupo abeliano****.*

Quando a operação associada a um grupo é dedutível a partir do contexto, é comum referir-se ao grupo unicamente por seu conjunto. Ou seja, quando não há margem a dúvidas acerca da operação associada a um grupo, o mesmo costuma ser notacionalmente simplificado para $\mathbb{G}$ em lugar da notação mais densa $(\mathbb{G}, \cdot)$. Tal simplificação notacional será aplicada no decorrer do texto quando conveniente e em contextos inequívocos.

Outro recurso notacional importante refere-se à *exponenciação* de um elemento de um grupo: sendo $a \in \mathbb{G}$, denota-se por $a^x, x \in \mathbb{Z}^+$, a aplicação repetida da operação '$\cdot$' sobre x cópias do mesmo elemento a (e.g. $a^3 = a \cdot a \cdot a$). Por outro lado, a exponenciação negativa a^{-1} é usada para denotar *o inverso de* a *em* $\mathbb{G}$, notação esta já usada quando da exposição da Definição 3.1.1. De igual maneira, utiliza-se a notação a^{-x} como significando o mesmo que $(a^{-1})^x$, ou seja, a aplicação da operação '$\cdot$' sobre x cópias do inverso de a, sendo de fácil verificação que $a^{-x} = (a^x)^{-1}$. Por consistência, define-se ainda $a^0 \stackrel{def}{=} e$, em que e, como já exposto, é a identidade do grupo.

Ainda com relação à notação empregada, e conforme bem ilustram as definições previamente estabelecidas, o presente trabalho emprega uma notação *multiplicativa* para grupos algébricos, em que a operação associada a um grupo é descrita como um "produto" e a aplicação repetida a um mesmo elemento é denotada como uma "exponenciação". Ressalta-se, entretanto, que este é um mero recurso notacional, não incorrendo este em alteração do significado da operação associada a um grupo: por exemplo, para o grupo $(\mathbb{Z}, +)$, em que '+' denota a adição tradicional sobre inteiros, tem-se que a notação multiplicativa se traduz para '$\cdot$' = '+', $a^x = x \cdot a$ e $a^{-1} = -a$.

No referente ao conjunto associado a um grupo $\mathbb{G}$, se este apresenta um número finito de elementos então ele é dito ser *finito*, sendo tal número de elementos denominado *ordem* de $\mathbb{G}$ e denotado por $|\mathbb{G}|$. Se o grupo possuir infinitos elementos então ele é dito ter ordem *infinita*. Tal conceito de ordem é estendido a um elemento $a \in \mathbb{G}$ como *o menor inteiro positivo* q *tal que* $a^q = e$, em que e, conforme mencionado, é a identidade da operação no grupo. De maneira semelhante, denota-se a ordem de a por $|a|$. Se tal valor q não existe então a ordem de a é dita ser *infinita*.

Intrinsecamente ligado a tal definição de ordem de um elemento, um outro conceito básico refere-se ao *conjunto gerado por* a, em que a é um elemento pertencente a um

grupo $\mathbb{G}$. Tal conjunto é dado por $\langle a \rangle = \{a^n, n \in \mathbb{Z}\}$, isto é, o conjunto formado por todas as potências inteiras do elemento a. Sob a definição de que $a^0 = e$, e lembrando que $a^{-x} = (a^x)^{-1}$, é de fácil verificação que $\langle a \rangle$ forma um subgrupo de $\mathbb{G}$. Quando $|a| = q$, tem-se que $\langle a \rangle = \{a^0, a^1, ..., a^{q-1}\}$. Por sua vez, quando $\langle a \rangle = \mathbb{G}$ diz-se que a é um *gerador* de $\mathbb{G}$ e, com isso, tem-se $|a| = |\mathbb{G}|$. Um grupo que possua ao menos um gerador é dito ser *cíclico*.

Postas as supracitadas definições, o presente trabalho tem enfoque sobre grupos abelianos cíclicos finitos, sendo os grupos doravante considerados, salvo menção explícita em contrário, todos considerados de tal natureza, a qual é formalizada na Definição 3.1.3 seguinte:

Definição 3.1.3 (Grupos abelianos cíclicos finitos)**.** *Seja* $(\mathbb{G}, \cdot)$ *um grupo abeliano, cíclico e finito. Isto implica que* $(\mathbb{G}, \cdot)$ *necessariamente satisfaz todas as seguintes propriedades:*

1. ***Abeliano****:* $(\mathbb{G}, \cdot)$ *satisfaz a Definição 3.1.2*
2. ***Cíclico****:* $\exists g \in \mathbb{G} : \langle g \rangle = \mathbb{G}$
3. ***Finito****:* $|\mathbb{G}| = n, n \in \mathbb{N}$

Assim, dada a Definição 3.1.3, e sendo g um gerador do grupo $\mathbb{G}$ de ordem n, tem-se que $\langle g \rangle = \{g^0, g^1, ..., g^{n-1}\} = \mathbb{G}$. Isto implica que $\forall a \in \mathbb{G} : \exists x \in \mathbb{Z}_n : a = g^x$, em que $\mathbb{Z}_n = \{0, 1, ..., n-1\}$. Note-se assim que, dada a natureza do gerador de um grupo, para cada elemento a de $\mathbb{G}$ não apenas um valor x conforme acima definido deve existir, mas tal valor também necessariamente deve ser único em $\mathbb{Z}_n$. Assim sendo, tem-se que $\forall a \in \mathbb{G}$, se g é um gerador de $\mathbb{G}$, então a igualdade $a = g^x$ é bem definida, sendo o valor x conforme acima exposto denominado de *logaritmo discreto de a na base g*. Este conceito é formalizado na Definição 3.1.4 que se segue:

Definição 3.1.4 (Logaritmo Discreto – DLOG)**.** *Seja* $\mathbb{G}$ *um grupo abeliano cíclico finito de ordem* n *e* $g \in \mathbb{G} : \langle g \rangle = \mathbb{G}$ *(i.e.* g *é um gerador de* $\mathbb{G}$*). Sendo* $a \in \mathbb{G}$*, dá-se o nome de* ***Logaritmo Discreto (DLOG) de a na base g*** *ao único valor* $x \in \mathbb{Z}_n$ *tal que* $a = g^x$*, e denota-se tal valor por* $x = \log_g a$*.*

Ressalta-se aqui que nem sempre o logaritmo discreto entre dois elementos de um grupo é bem definido. Como um exemplo, seja considerado o grupo multiplicativo módulo 7, dado por $\mathbb{Z}_7^* = \{1, 2, 3, 4, 5, 6\}$ e cuja operação é a multiplicação módulo 7. Embora seja facilmente verificável que $\log_3 2 = 2$ e $\log_3 6 = 3$, tem-se que $\nexists x \in \mathbb{Z}_7 : 2 = 6^x$, de modo que $\log_6 2$ é indefinido nesse grupo. De maneira semelhante, tem-se que $\log_2 6$ também é indefinido para o mesmo grupo. Note-se, entretanto, que isto não invalida a Definição 3.1.4: quando a base logarítmica é um gerador do grupo, o logaritmo discreto se

encontra garantidamente bem definido para qualquer elemento desse grupo. No exemplo dado, é de fácil verificação que nem 6 nem 2 são geradores de $\mathbb{Z}_7^*$.

Esta última afirmação ilustra a motivação por trás da restrição da base a geradores empregada na Definição 3.1.4, mas cabe observar que o logaritmo discreto *pode* existir mesmo que a base *não* seja um gerador: aproveitando-se do exemplo com $\mathbb{Z}_7^*$, note-se que, embora 2 não seja gerador, tem-se que $\log_2 4 = 2$. A restrição imposta na Definição 3.1.4 visa enfatizar que, salvo menção em contrário, logaritmos discretos no contexto do presente trabalho são considerados como tendo geradores por base.

Uma vez definidos um grupo $\mathbb{G}$ de ordem n e um gerador g desse mesmo grupo, é considerado ser um problema de fácil resolução computar a exponenciação $g^x, x \in \mathbb{Z}_n$. Classifica-se aqui por "*problema de fácil resolução*" um problema para o qual exista e seja conhecido ao menos um algoritmo determinístico polinomial capaz de computar uma solução válida para o mesmo, sendo um tal algoritmo denominado "eficiente". Embora a eficiência real de tal computação seja dependente do grupo em questão, existem diversos grupos para os quais algoritmos eficientes são conhecidos para a computação da exponenciação. Para $(\mathbb{Z}, +)$, por exemplo, $a^x = x \cdot a$, ou seja, a exponenciação consiste em computar uma multiplicação inteira. Já para um grupo como $(\mathbb{Z}_p^*, \cdot_p)$, em que p é primo, $\mathbb{Z}_p^* = \{1, 2, 3, ..., p-1\}$ e '$\cdot_p$' denota a *multiplicação módulo p*, algoritmos eficientes como o *"Square and Multiply"* são conhecidos de longa data [19].

Entretanto, dados um grupo $\mathbb{G}$ de ordem n, um gerador g desse mesmo grupo e um elemento $a \in \mathbb{G} : a = g^x, x \in \mathbb{Z}_n$, é possível se identificar grupos para os quais é considerado ser de difícil resolução o problema de computar o valor do logaritmo discreto $x = \log_g a$. Por "*problema de difícil resolução*" denota-se um problema para o qual não são conhecidos algoritmos determinísticos polinomiais capazes de computar uma solução válida para o mesmo a menos de uma probabilidade negligenciável. O problema da computação do logaritmo discreto é formalizado na Definição 3.1.5, derivada daquela dada por Stinson [43, p.234], mas aqui, por fins de conformidade com definições prévias e com o contexto deste trabalho, aplicada considerando-se geradores como base logarítmica:

Definição 3.1.5 (Problema do Logaritmo Discreto – DLP)**.** *Seja um grupo $\mathbb{G}$ de ordem n, um gerador g desse grupo e um elemento $a \in \mathbb{G}$. O* ***Problema do Logaritmo Discreto (DLP)*** *consiste em encontrar o único inteiro $x \in \mathbb{Z}_n$ tal que $a = g^x$. Note-se que, pela Definição 3.1.4, $x = \log_g a$.*

Assim sendo, considerando-se o grupo $(\mathbb{Z}_p, +_p)$, em que p é primo, $\mathbb{Z}_p = \{0, 1, 2, ..., p-1\}$ e '$+_p$' denota a *adição módulo p*, note-se tratar-se este de um grupo no qual o problema do logaritmo discreto é de simples resolução. Afinal, sendo g um gerador deste grupo, tem-se que o problema do logaritmo discreto para $a \in \mathbb{Z}_p$ consiste em computar o valor x tal que $a = g^x = g +_p ... +_p g \quad = (x \cdot g) \bmod p$, o que implica que $x = (a \cdot g^{-1}) \bmod p$, em que g^{-1} é o inverso multiplicativo de g, cuja existência é garantida se g é gerador de

$(\mathbb{Z}_p, +_p)$. Assim sendo, o problema do logaritmo discreto em tal grupo consiste em computar uma multiplicação entre o valor a considerado e o inverso multiplicativo do gerador usado como base, o que sabidamente pode ser resolvido de forma eficiente.

Entretanto, quando considerados, por exemplo, os grupos multiplicativos $\mathbb{Z}_p^*$ já previamente mencionados, embora para valores p suficientemente pequenos uma metodologia de força-bruta possa ser empregada na resolução do DLP, não são até o momento conhecidos algoritmos determinísticos polinomiais em um modelo clássico de computação que sejam capazes de resolver o problema do logaritmo discreto no caso geral. Desse modo, tal problema é hipotetizado como sendo de difícil resolução para valores de p suficientemente grandes. Esta hipótese é formalizada na Definição 3.1.6 seguinte:

Definição 3.1.6 (Hipótese do Logaritmo Discreto)**.** *Existem grupos $\mathbb{G}$ nos quais o DLP é intratável.*

A grande utilidade criptográfica da hipótese do logaritmo discreto reside no fato de que, embora computar uma exponenciação possa ser uma operação eficiente, computar o logaritmo discreto é hipoteticamente um problema intratável em grupos $\mathbb{G}$ específicos e de tamanho suficiente. Assim sendo, quando aplicada a elementos de grupos nos quais tal hipótese seja válida, a exponenciação opera como uma *função de sentido único*, caracterizada pela facilidade de computação de seu valor em contraste com a dificuldade de computação de sua inversa.

Uma outra classe de grupos nos quais a hipótese do logaritmo discreto é considerada válida é aquela dos grupos definidos por curvas elípticas, descritos em maiores detalhes na seção seguinte.

3.2 Curvas Elípticas

Para além dos grupos baseados em aritmética inteira descritos previamente, outros grupos amplamente aplicados, principalmente em contextos criptográficos, são aqueles definidos por pontos de curvas elípticas. Uma curva elíptica $\mathbb{E}$ pode ser definida sobre um corpo matemático como sendo o conjunto de pontos (x, y) que satisfazem a equação $y^2 = x^3 + ax + b$, em que a, b pertencem ao corpo de definição da curva, junto a um outro ponto especial denominado *ponto infinito* [22]. Ou seja, denotando-se por $\mathcal{O}$ o ponto infinito, define-se uma curva elíptica $\mathbb{E}$ como:

$$\mathbb{E} \stackrel{def}{=} \{(x, y) : y^2 = x^3 + ax + b\} \cup \{\mathcal{O}\}$$

Sobre os pontos de uma curva elíptica $\mathbb{E}$ define-se uma operação binária $(\cdot)$ que, de maneira simplificada e interpretando-se a curva sobre os reais $\mathbb{R}$, pode ser compreendida como operando da seguinte forma [21, p.326]:

- sendo $\mathcal{O}$ o ponto infinito e P qualquer ponto sobre a curva elíptica, define-se $\mathcal{O} \cdot P = P \cdot \mathcal{O} = P$. Note-se, assim, que o ponto infinito $\mathcal{O}$ opera como identidade da operação.
- mediante análise da equação de uma curva elíptica $\mathbb{E}$, observe-se que a mesma apresenta uma simetria com relação ao eixo X, i.e., $(x, y) \in \mathbb{E} \iff (x, -y) \in \mathbb{E}$. Dessa forma, define-se o inverso de um ponto $P = (x, y)$ como seu simétrico $P^{-1} = (x, -y)$.
- sendo dois pontos $P_1, P_2 \in \mathbb{E} : P_1, P_2 \neq \mathcal{O}$, é possível se demonstrar que, traçando-se uma reta $\mathcal{R}$ que passe sobre P_1 e P_2 e considerando-se o ponto infinito $\mathcal{O}$ como pertencente a qualquer reta vertical, tal reta $\mathcal{R}$ necessariamente intercepta a curva em um terceiro ponto P_3. Assim sendo, define-se a operação entre dois pontos P_1 e P_2 de uma curva como:

  $$P_1 \cdot P_2 \stackrel{def}{=} P_3^{-1},$$

  em que $P_1, P_2, P_3 \in \mathcal{R}$. Ou seja, a operação entre P_1 e P_2 é o inverso do terceiro ponto obtido pelo traçado da reta $\mathcal{R}$ que passa por P_1 e P_2.
- sendo $P \in \mathbb{E}$ e $m \in \mathbb{Z}$, denota-se por P^m a aplicação da operação $(\cdot)$ sobre m instâncias do ponto P. Se m é negativo, a operação é aplicada sobre m instâncias do inverso de P (i.e., P^{-1}).

É possível se demonstrar que o conjunto de pontos de uma curva elíptica junto à operação binária acima definida caracterizam um grupo abeliano. Ou seja, existem curvas $\mathbb{E}$ tais que o par $(\mathbb{E}, \cdot)$ forma um grupo abeliano. É também possível se definir curvas elípticas que formem grupos que, além de abelianos, sejam também cíclicos e finitos. Diversas curvas com tais características se encontram padronizadas, inclusive em caráter oficial para uso governamental [28, p.87–101].

Assim sendo, com base na supracitada operação binária sobre pontos de curvas elípticas, formaliza-se aqui o chamado *problema do logaritmo discreto em curvas elípticas*:

Definição 3.2.1 (Problema do Logaritmo Discreto em Curvas Elípticas – ECDLP)**.** *Seja uma curva elíptica $\mathbb{E}$ de ordem n, um ponto gerador G dessa curva e um ponto $P \in \mathbb{E}$. O* ***Problema do Logaritmo Discreto em Curvas Elípticas (ECDLP)*** *consiste em encontrar o único inteiro $x \in \mathbb{Z}_n$ tal que $P = G^x$.*

De maneira semelhante à sua contrapartida em grupos multiplicativos inteiros, o ECDLP exposto na Definição 3.2.1 é também hipotetizado como um problema de difícil resolução para algumas curvas elípticas. A grande importância criptográfica do ECDLP em contraste com o DLP, no entanto, é que os algoritmos conhecidos de resolução do ECDLP são ainda menos eficientes que aqueles conhecidos para a resolução do DLP [21, p.325]. Isto implica que grupos lastreados em curvas elípticas exigem parâmetros menores em termos de número de bits para um mesmo nível pretendido de segurança.

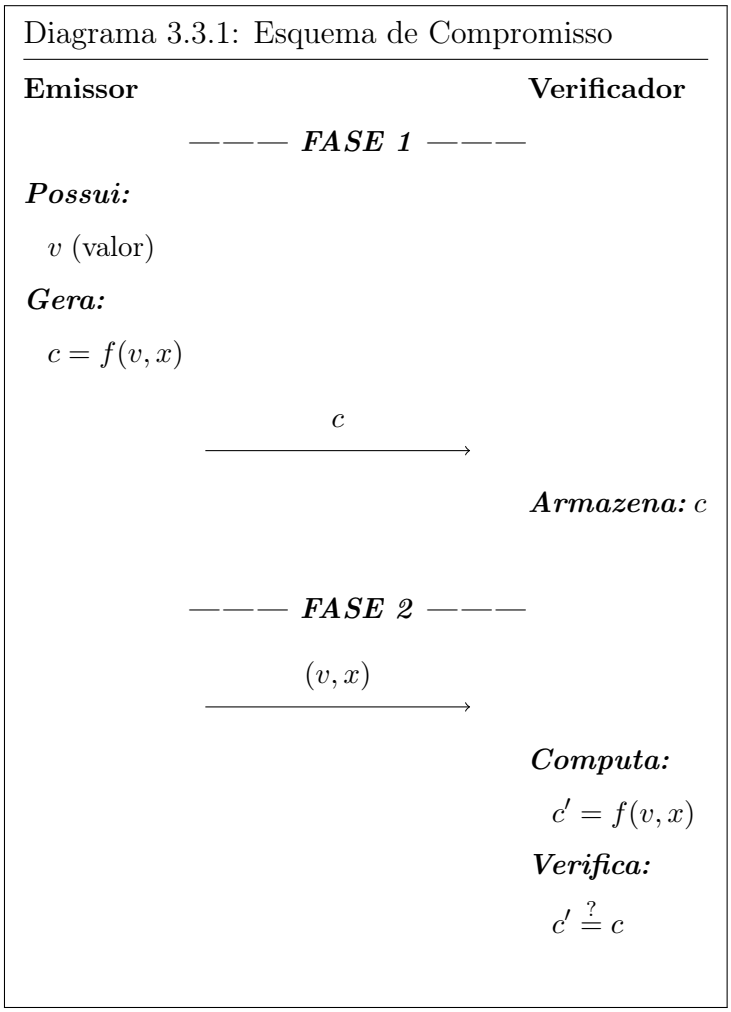
Assim, conhecidos grupos para os quais a hipótese do logaritmo discreto é tida como válida, esquemas criptográficos nela lastreados podem ser engendrados de modo a se obter construtos matemáticos com propriedades de interesse para diversas aplicações e parâmetros de segurança variados. Dentre estes tem-se o chamado *Compromisso de Pedersen*, que consiste em um esquema criptográfico de compromisso (*commitment*) de valores, conceito brevemente explanado na seção seguinte.

3.3 Esquemas de Compromisso

Esquemas de Compromisso [1] constituem um conceito criptográfico básico que permite a uma das partes executoras de um protocolo se comprometer com um determinado valor diante de alguma outra parte executora desse mesmo protocolo de modo tal que o sigilo sobre o valor comprometido seja mantido [18, p.223]. O Diagrama 3.3.1 seguinte explicita o funcionamento geral de um esquema de compromisso genérico:

[1] No decorrer deste texto utiliza-se o termo *"compromisso"* em português como tradução para o termo *"commitment"* em inglês.

Diagrama 3.3.1: Esquema de Compromisso



Conforme expõe o Diagrama 3.3.1, um esquema de compromisso geral é constituído de duas fases: a **Fase 1**, de *compromisso*, na qual um *emissor* se compromete com um valor v, sem revelá-lo, por meio da divulgação de algum artefato c (e.g. um resultado de alguma função matemática) a um *verificador*; e a **Fase 2**, de *revelação*, na qual o *emissor* revela o valor v comprometido ao mesmo tempo em que demonstra ao *verificador* como o artefato c previamente divulgado se relaciona com o valor comprometido agora revelado. O Diagrama 3.3.1 deixa implícito que a função f de geração do artefato c é de conhecimento geral das partes envolvidas no esquema e pode se utilizar de parâmetros adicionais, no diagrama coletivamente representados por x.

É importante notar que um esquema de compromisso deve operar de uma maneira tal que deduzir o valor v comprometido seja um problema intratável para um *verificador* que conhece unicamente o artefato c divulgado. Do contrário, um *verificador* pode, antes da fase de revelação do esquema de compromisso, se utilizar de um algoritmo determinístico polinomial para desvelar o valor v comprometido pelo *emissor*. Quando tal intratabilidade se mostra verdadeira apenas quando da imposição de limitações computacionais sobre o *verificador*, então o esquema de compromisso é dito ter *sigilo computacional*. Em contraste, quando o sigilo de v é inviolável a partir de c mesmo diante de um *verificador*

computacionalmente ilimitado, então o esquema é dito ter *sigilo incondicional*.

De maneira semelhante, a capacidade de produzir uma tupla (v', x') distinta de (v, x), mas tal que $f(v, x) = f(v', x')$, deve ser um problema intratável para um *emissor*. Do contrário, tal *emissor* pode, antes da fase de revelação, computar um valor distinto daquele originalmente comprometido, mas que também passará pelo crivo do *verificador*, permitindo ao *emissor* selecionar para revelação aquele que mais lhe aprouver no contexto do protocolo em execução. Também de forma semelhante, quando a intratabilidade de tal problema exige a imposição de limitações computacionais sobre o *emissor*, o esquema é dito ter *compromisso computacional*. Quando tal problema se mostra intratável mesmo para um *emissor* sem restrições computacionais então diz-se que o esquema é de *compromisso incondicional*.

Como um exemplo ilustrativo, seja considerado um esquema de compromisso no qual o *emissor* se compromete com um valor $v \in \mathbb{Z}_p$ divulgando $c = f(v) = g^v$, em que g é um gerador conhecido de um grupo multiplicativo inteiro $\mathbb{G}$ de ordem p um número primo suficientemente grande de modo que o DLP seja considerável intratável. O esquema é demonstrado no Diagrama 3.3.2 seguinte:

Diagrama 3.3.2: Esquema de Compromisso (Exemplo)

Emissor		**Verificador**
	——— ***FASE 1*** ———	
Possui:		
$v \in \mathbb{Z}_p$ (valor)		
Gera:		
$c = g^v$		
	$c \longrightarrow$	
		Armazena: c
	——— ***FASE 2*** ———	
	$v \longrightarrow$	
		Computa:
		$c' = g^v$
		Verifica:
		$c' \stackrel{?}{=} c$

Dada a validade da hipótese do logaritmo discreto em $\mathbb{G}$, um *verificador* consegue revelar v antes da fase de revelação computando $\log_g c$, um problema intratável sob

consideração de limitações computacionais sobre o *verificador*. Por outro lado, dada a existência de um único valor $v \in \mathbb{Z}_p$ tal que $c = g^v$, mesmo um *emissor* sem restrições computacionais é incapaz de produzir um valor v' diferente de v que satisfaça o crivo do *verificador*. Logo, o Diagrama 3.3.2 representa um esquema de *compromisso incondicional* e *sigilo computacional*.

Em contraste, um outro exemplo de esquema de compromisso que se utiliza também de propriedades do logaritmo discreto, mas apresentando propriedades de *compromisso computacional* e *sigilo incondicional*, é dado por artefatos conhecidos como *Compromissos de Pedersen*, estudados em detalhes na seção que se segue.

3.4 Compromissos de Pedersen

Em um de seus escritos, Pedersen descreve uma metodologia de compartilhamento de segredos que se utiliza de um esquema de compromisso computacional com sigilo incondicional [30]. Tal esquema lastreia-se na complexidade da resolução do problema do logaritmo discreto em grupos multiplicativos $\mathbb{Z}_p^*$ suficientemente grandes, em que p é um número primo. Isto posto, os *Compromissos de Pedersen*, como passaram a ser conhecidos tais construtos, são definidos conforme exposto no que se segue:

Definição 3.4.1. *Sejam p e q números primos grandes tais que $q|p-1$ e denote-se por $\mathbb{G}_q$ o subgrupo de $\mathbb{Z}_p^*$ de ordem q. Sendo g e h dois geradores de $\mathbb{G}_q$ e $r \xleftarrow{\$} \mathbb{Z}_q$, um* Compromisso de Pedersen *para um valor $v \in \mathbb{Z}_q$ é dado por:*

$$E(v, r) = g^v \cdot h^r \pmod{p}$$

Esta definição básica resguarda duas propriedades que conferem aos Compromissos de Pedersen suas características intrínsecas de sigilo incondicional e compromisso computacional. O sigilo incondicional de tais compromissos é garantido pelo Teorema 3.4.2, exposto a seguir:

Teorema 3.4.2. *Seja $E(v, r)$ um Compromisso de Pedersen para um valor $v \in \mathbb{Z}_q$. Então:*

$$\forall v' \in \mathbb{Z}_q : \exists r' \in \mathbb{Z}_q : E(v', r') = E(v, r)$$

Demonstração. Seja f um gerador qualquer de $\mathbb{G}_q$. Por se tratar de um gerador, certamente existem valores $x, y \in \mathbb{Z}_q$ tais que:

$$g = f^x \pmod p$$
$$h = f^y \pmod p$$

Assim sendo, é possível se reescrever um compromisso de Pedersen $E(v, r)$ em termos desse gerador f da seguinte maneira:

$$E(v, r) = g^v \cdot h^r = (f^x)^v \cdot (f^y)^r = f^{xv} \cdot f^{yr} = f^{xv+yr} \pmod p$$

Seja agora considerado um valor $v' \in \mathbb{Z}_q$ qualquer a ser utilizado para produção de um compromisso $E(v', r')$. Lembrando que $\mathbb{Z}_q = \{0, 1, ..., q-1\}$, observe-se que o seguinte predicado certamente é válido:

$$\forall v' \in \mathbb{Z}_q : \exists w \in \mathbb{Z}_q : (v + w) \bmod q = v'$$

Isto é, qualquer valor v' pode ser escrito em termos de outro valor v como uma adição modular em q. Considere-se então o compromisso de Pedersen $E(v', r')$ seguinte, também escrito em termos do gerador f previamente definido e utilizando-se da relação modular entre v' e v acima estabelecida:

$$E(v', r') = g^{v'} \cdot h^{r'} = (f^x)^{v+w} \cdot (f^y)^{r'} = f^{xv+xw+yr'} \pmod p$$

Como h é gerador tem-se necessariamente que $h \neq 1$, o que, sendo $h = f^y$, implica que $y \neq 0$. Dessa forma, como q é primo, existe o inverso multiplicativo y^{-1} de y em $\mathbb{Z}_q$. Consequentemente, deve também existir o valor $r' \in \mathbb{Z}_q : r' = (y^{-1}) \cdot (yr - xw) \pmod q$. Aplicando-se tal valor para r' no compromisso $E(v', r')$, tem-se:

$$\begin{aligned} E(v', r') &= f^{xv+xw+yr'} \\ &= f^{xv+xw+y\cdot[(y^{-1})\cdot(yr-xw)]} \\ &= f^{xv+xw+yr-xw} = f^{xv+yr} \\ &= E(v, r) \end{aligned}$$

□

O Teorema 3.4.2 diz que um mesmo compromisso E pode ser resultado de qualquer valor $v \in \mathbb{Z}_q$ mediante escolha de um valor $r \in \mathbb{Z}_q$ apropriado. Assim, a observação do compromisso E nada revela acerca do valor v comprometido. Disso decorre o sigilo incondicional dos *Compromissos de Pedersen*, propriedade esta formalizada no Corolário 3.4.3:

Corolário 3.4.3. *Um* Compromisso de Pedersen *apresenta sigilo (*hiding*) incondicional.*

Observação. *Note-se que o Corolário 3.4.3 assume tacitamente que r é uniformemente escolhido em $\mathbb{Z}_q$, sendo o resultado exposto válido apenas sob tal hipótese. Tal aspecto não é explicitamente ressaltado aqui dado que se encontra claramente assumido na Definição 3.4.1.*

Já a qualidade computacional dos compromissos de Pedersen deriva da necessidade de computação de logaritmos discretos para adulteração de um valor previamente comprometido, aspecto este demonstrado pelo Teorema 3.4.4 seguinte:

Teorema 3.4.4. *Sejam conhecidos valores $v, v', r, r' \in \mathbb{Z}_q$, com $v \neq v'$ e $r \neq r'$, tais que $E(v,r) = E(v',r')$. Então, o logaritmo discreto $\log_h g$ é computável em tempo polinomial.*

Demonstração. Sejam os compromissos $E(v,r)$ e $E(v',r')$ produzidos por valores conforme enunciado do teorema e tais que $E(v,r) = E(v',r')$. Por definição, tem-se que:

$$E(v,r) = E(v',r') \implies g^v \cdot h^r = g^{v'} \cdot h^{r'}$$

Note-se que, como $g, h \in \mathbb{G}_q$, tem-se que $\forall x \in \mathbb{Z}_q : g^x, h^x \in \mathbb{G}_q$, de modo que certamente devem existir inversos multiplicativos $g^{-x}, h^{-x} \in \mathbb{G}_q$. Ademais, como $v \neq v'$ e q é primo, certamente o valor $v - v'$ possui inverso multiplicativo $(v - v')^{-1}$ em $\mathbb{Z}_q$. Assim sendo, tem-se:

$$\begin{aligned} g^v \cdot h^r = g^{v'} \cdot h^{r'} &\implies g^v \cdot g^{-v'} = h^{-r} \cdot h^{r'} \\ &\implies g^{v-v'} = h^{r'-r} \\ &\implies \log_h g^{v-v'} = r' - r \\ &\implies (v - v') \cdot \log_h g = r' - r \\ &\implies \log_h g = (r' - r) \cdot (v - v')^{-1} \bmod q \end{aligned}$$

□

O Teorema 3.4.4 demonstra que o conhecimento de dois Compromissos de Pedersen iguais para valores v e v' distintos possibilita computar trivialmente o logaritmo discreto $\log_h g$ entre os dois geradores escolhidos. Sob a hipótese da dificuldade de computação do logaritmo discreto para grupos específicos e suficientemente grandes, conclui-se o caráter computacional associado a tais esquemas de compromisso no contexto de tais grupos, aspecto este formalizado no Corolário 3.4.5 a seguir:

Corolário 3.4.5. *Sob a hipótese de dificuldade de resolução do problema do logaritmo discreto no grupo escolhido, um* Compromisso de Pedersen *apresenta amarração (i.e.* binding*) computacional.*

Observe-se ainda que a escolha de $\log_h g$ no enunciado e na demonstração do Teorema 3.4.4 é arbitrária no sentido de que $\log_g h$ é também trivialmente computável sob as mesmas hipóteses, e de igual maneira incorrendo nas mesmas conclusões. Afinal, verifica-se facilmente que o conhecimento de $\log_h g$ implica em conhecimento de $\log_g h$:

$$\begin{aligned}
\log_h g = x &\implies g = h^x \\
&\implies g^{(x^{-1})} = (h^x)^{(x^{-1})} = h^{x\cdot(x^{-1})} = h \\
&\implies \log_g h = x^{-1} \\
&\implies \log_g h = (\log_h g)^{-1}
\end{aligned}$$

Observe-se que, como g e h são assumidos como geradores tais que $g \neq h$, o valor x certamente existe em $\mathbb{Z}_q$ assim como seu inverso multiplicativo x^{-1}.

Uma terceira característica associada aos *Compromissos de Pedersen* lhes é conferida pelo aspecto exponencial associado à forma como são computados: o esquema de compromisso é *homomórfico*. Esse aspecto é formalizado no Teorema 3.4.6 a seguir:

Teorema 3.4.6. *Se $E(v, r)$ e $E(v', r')$ são dois* Compromissos de Pedersen*, então tem-se que:*

$$E(v, r) \cdot E(v', r') = E(v + v', r + r')$$

Demonstração. A demonstração do homomorfismo dos compromissos é trivial e verificável pela aplicação de sua própria definição:

$$\begin{aligned}
E(v, r) \cdot E(v', r') &= (g^v \cdot h^r) \cdot (g^{v'} \cdot h^{r'}) \\
&= (g^v \cdot g^{v'}) \cdot (h^r \cdot h^{r'}) \\
&= g^{v+v'} \cdot h^{r+r'} \\
&= E(v + v', r + r')
\end{aligned}$$

□

O Teorema 3.4.6 mostra que o produto de dois compromissos de Pedersen resulta em um compromisso do somatório dos valores originalmente comprometidos. Note-se ainda que a extensão de tal propriedade a qualquer produtório de compromissos sob os mesmos parâmetros é também facilmente verificável como válida.

Como visto, *Compromissos de Pedersen* lastreiam-se em dois geradores $g, h \in \mathbb{G}_q$, os quais devem ser devidamente definidos de modo a se aplicar o esquema. Isto posto, é importante ressaltar a necessidade de que $\log_h g$ seja desconhecido. Caso contrário, torna-se trivial à parte emissora, em conhecendo tal logaritmo, produzir informação que

possibilite revelar seu compromisso para um valor distinto daquele originalmente comprometido. Esta propriedade pode ser tacitamente inferida pelo Teorema 3.4.4, mas, por clareza, é formalizada pelo Teorema 3.4.7:

Teorema 3.4.7. *No contexto dos compromissos de Pedersen, um emissor que conheça* $\log_h g$ *pode revelar seu compromisso para qualquer valor, independentemente daquele por ele originalmente comprometido.*

Demonstração. Suponha-se conhecido o valor $\log_h g$, por conveniência aqui denotado por x. Seja um compromisso de Pedersen $E(v, r) = g^v \cdot h^r \bmod p$, por meio do qual o emissor se compromete com o valor v. Se tal emissor, no entanto, desejar comprovar esse compromisso como referente a um valor v' tal que $v' \neq v$, ele poderá fazer uso de seu conhecimento acerca de $\log_h g$ para produzir um novo valor $r' \in \mathbb{Z}_q, r' \neq r$, que satisfaça a igualdade $E(v, r) = E(v', r')$. Pelo Teorema 3.4.4, sabe-se que:

$$x = \log_h g = (r' - r) \cdot (v - v')^{-1} \bmod q$$

Note-se, entretanto, que:

$$\begin{aligned} x = (r' - r) \cdot (v - v')^{-1} &\implies (r' - r) = x \cdot (v - v') \\ &\implies r' = [x \cdot (v - v')] + r \bmod q \end{aligned}$$

Como o emissor, além dos parâmetros públicos associados ao esquema, conhece também os valores v, r, v' e $x = \log_h g$, a computação de r' mostra-se trivial. Note-se que, utilizando-se de tal valor r' para construir um compromisso a v', tem-se:

$$\begin{aligned} E(v', r') &= g^{v'} \cdot h^{r'} \\ &= g^{v'} \cdot h^{[x \cdot (v - v')] + r} \\ &= g^{v'} \cdot h^{x \cdot (v - v')} \cdot h^r \\ &= g^{v'} \cdot (h^x)^{v - v'} \cdot h^r \\ &= g^{v'} \cdot g^{v - v'} \cdot h^r \\ &= g^{v' + v - v'} \cdot h^r \\ &= g^v \cdot h^r \\ &= E(v, r) \end{aligned}$$

□

O Teorema 3.4.7 explicita a importância da devida cautela quando da definição dos parâmetros a serem utilizados para geração de compromissos de Pedersen. Em particular,

os geradores usados devem ser escolhidos de modo que a relação logarítmica entre ambos seja desconhecida e, portanto, resguardadas as devidas dimensões do grupo subjacente ao esquema, apresente uma probabilidade de computação negligenciável sob a hipótese do logaritmo discreto. Caso contrário, isto é, se conferido a uma das partes o poder de escolha dos geradores sem qualquer interferência ou contribuição das demais partes envolvidas, a esta parte é oferecida a oportunidade de produzir geradores tais que a relação logarítmica entre ambos seja conhecida, conferindo-lhe de tal forma a capacidade de revelar um compromisso para um valor distinto daquele originalmente comprometido.

Dado o caráter de sigilo incondicional associado aos compromissos de Pedersen, uma questão de interesse a ser analisada diz respeito às probabilidades associadas à geração de compromissos específicos, aspecto este formalizado pelo Teorema 3.4.8 exposto a seguir:

Teorema 3.4.8. *Seja q a ordem do grupo $\mathbb{G}_q$ sobre o qual opera o esquema de compromisso de Pedersen. Sob hipótese de uniformidade na escolha dos valores aleatórios r, a probabilidade de produção de um compromisso c específico é o inverso de q, independentemente da distribuição probabilística ou restrições impostas sobre os valores v a serem comprometidos.*

Demonstração. Seja $c \in \mathbb{G}_q$ um compromisso de Pedersen. Assim, denota-se por $P[C = c]$ a probabilidade de que tal compromisso seja produzido, em que C é uma variável aleatória associada ao compromisso. Definindo V como uma variável aleatória associada ao valor comprometido, tem-se:

$$\begin{aligned} P[C = c] = {} & P[C = c|V = v_1] \cdot P[V = v_1] + \\ & + P[C = c|V = v_2] \cdot P[V = v_2] + ... \\ & + P[C = c|V = v_i] \cdot P[V = v_i] + ... \\ & + P[C = c|V = v_q] \cdot P[V = v_q] \end{aligned}$$

Note-se que, conforme demonstra o Teorema 3.4.2, $\forall v \in \mathbb{Z}_q : \exists r \in \mathbb{Z}_q : g^v \cdot h^r = c$. Isto posto, e sendo R uma variável aleatória associada ao valor r utilizado, tem-se que:

$$\forall v \in \mathbb{Z}_q : P[C = c|V = v] = P[R = r] = 1/q$$

Portanto, conclui-se que:

$$\begin{aligned} P[C = c] = 1/q \cdot P[V = v_1] + \\ +1/q \cdot P[V = v_2] + ... \\ +1/q \cdot P[V = v_i] + ... \\ +1/q \cdot P[V = v_q] = 1/q \cdot \sum_{v \in \mathbb{Z}_q} P[V = v] = 1/q \end{aligned}$$

□

Embora possa parecer óbvio dada a incondicionalidade do sigilo de compromissos de Pedersen, o Teorema 3.4.8 busca enfatizar o aspecto de que não é necessário que todos os valores em $\mathbb{Z}_q$ sejam plausíveis de serem comprometidos. Em particular, os compromissos de Pedersen não perdem sua incondicionalidade com relação ao sigilo mesmo quando os valores comprometidos são restritos a um subconjunto ínfimo de $\mathbb{Z}_q$ sob qualquer distribuição de probabilidade. Note-se assim que mesmo sob tais limitações todos os possíveis compromissos mantêm-se equiprováveis, isto por conta da hipótese de uniformidade associada à geração dos valores aleatórios utilizados. Tal característica é particularmente útil quando os valores a serem comprometidos são votos imprescindíveis de sigilo: os compromissos de Pedersen lhes resguardam privacidade mesmo que suas restrições de validade particulares impliquem na delimitação de seus valores a apenas um pequeno subgrupo de $\mathbb{Z}_q$, além de provável distribuição não-uniforme e potencialmente previsível.

Lastreado no exposto pelo Teorema 3.4.8, outra questão de interesse a ser analisada diz respeito à probabilidade de colisões de compromissos produzidos no contexto de grupos $\mathbb{G}_q$ suficientemente grandes, o que é mostrado pelo Teorema 3.4.9 seguinte:

Teorema 3.4.9. *Seja q a ordem do grupo sobre o qual opera o esquema de compromisso de Pedersen. Seja ainda T um limite superior estimado para o número desses compromissos a serem gerados no contexto de alguma aplicação específica. Sob a hipótese de uniformidade de escolha dos valores aleatórios r, se $q \gg T$, isto é, se $q - T \approx q$, então a probabilidade de que ao menos dois recibos c de mesmo valor sejam produzidos no contexto da aplicação é negligenciável.*

Demonstração. Denotando por COL o evento de colisão de ao menos dois compromissos no contexto da aplicação, tem-se que $P[COL] = 1 - P[\overline{COL}]$, em que o evento complementar $\overline{COL}$ denota a inexistência de qualquer colisão. Estipulando-se para cada um dos T compromissos gerados um índice $i, 1 \leq i \leq T$, denote-se por E_i o evento no qual o i-ésimo recibo não colide com nenhum dos recibos com índice $j < i$. Assim, assumindo-se que a geração de cada recibo é independente, tem-se que:

$$P[\overline{COL}] = P[E_1] \cdot P[E_2] \cdot ... \cdot P[E_i] \cdot ... \cdot P[E_T]$$

Sob a hipótese de uniformidade na escolha dos valores aleatórios r, o Teorema 3.4.8 mostra que a probabilidade de geração de um compromisso específico é $1/q$. Logo:

$$
\begin{aligned}
P[E_1] &= q \cdot \frac{1}{q} = \frac{q}{q} \\
P[E_2] &= (q-1) \cdot \frac{1}{q} = \frac{q-1}{q} \\
&\ldots \\
P[E_i] &= (q-i+1) \cdot \frac{1}{q} = \frac{q-i+1}{q} \\
&\ldots \\
P[E_T] &= (q-T+1) \cdot \frac{1}{q} = \frac{q-T+1}{q}
\end{aligned}
$$

Isto posto, em termos da probabilidade de inexistência de colisões e utilizando-se da hipótese de que $q - T \approx q$, tem-se:

$$
\begin{aligned}
P[\overline{COL}] &= P[E_1] \cdot P[E_2] \cdot \ldots \cdot P[E_i] \cdot \ldots \cdot P[E_T] \\
&= \frac{q}{q} \cdot \frac{q-1}{q} \cdot \ldots \cdot \frac{q-i+1}{q} \cdot \ldots \cdot \frac{q-T+1}{q} \\
&= \left(\frac{1}{q}\right)^T \cdot \prod_{i=1}^{T} (q-i+1) \\
&\approx \left(\frac{1}{q}\right)^T \cdot \prod_{i=1}^{T} q = \left(\frac{1}{q}\right)^T \cdot q^T = 1
\end{aligned}
$$

Portanto, $P[\overline{COL}] \approx 1$ e, com isto, $P[COL] \approx 0$.

□

Note-se que o aspecto negligenciável referente à probabilidade de colisão descrita pelo Teorema 3.4.9 deriva diretamente da cardinalidade do total de compromissos produzidos, assumido ser de caráter polinomial, diante do parâmetro de segurança do esquema expresso pela ordem q do grupo algébrico usado, de caráter exponencial.

Utilizando-se da desigualdade matemática $1 - x < e^{-x}$ é possível escrever uma aproximação para a probabilidade de colisão de compromissos $P[COL]$ da seguinte forma:

$$
\begin{aligned}
P[COL] &= 1 - P[\overline{COL}] \\
&= 1 - (P[E_1] \cdot P[E_2] \cdot \ldots \cdot P[E_i] \cdot \ldots \cdot P[E_T]) \\
&= 1 - \left(\frac{q}{q} \cdot \frac{q-1}{q} \cdot \ldots \cdot \frac{q-i+1}{q} \cdot \ldots \cdot \frac{q-T+1}{q} \right) \\
&= 1 - \prod_{i=0}^{T-1} \left(1 - \frac{i}{q} \right) > 1 - \prod_{i=0}^{T-1} \left(e^{-\frac{i}{q}} \right) = 1 - e^{-\frac{1}{q} \cdot \sum_{i=0}^{T-1}(i)} \\
&= 1 - e^{-\frac{1}{q} \cdot \frac{T \cdot (T-1)}{2}} = 1 - e^{-\frac{T \cdot (T-1)}{2 \cdot q}} \\
P[COL] &> 1 - e^{-\frac{T \cdot (T-1)}{2 \cdot q}}
\end{aligned}
$$

Resolvendo a equação de aproximação probabilística acima demonstrada em termos do número T de compromissos, obtém-se para uma probabilidade 1/2 de colisão a inequação $T^2 - T < 2 \cdot q \cdot \ln 2$. Esta equação revela por exemplo que, em um grupo da ordem de 256 bits, de modo a se obter uma probabilidade de colisão maior que 1/2, é necessário produzir uma quantidade de compromissos aproximadamente da ordem de 2^{128}. Ainda considerando-se esta mesma probabilidade, o número de compromissos necessários para colisão cresce para a ordem aproximada de 2^{256} em grupos de 512 bits e para 2^{512} em grupos de 1024 bits. Em geral, para uma quantidade de compromissos gerados na ordem de $\sqrt{q}$, a probabilidade de colisão já pode ser considerada como não negligenciável [6, Apêndice, p.35]. Entretanto, dado o aspecto exponencial de q, para grupos suficientemente grandes a quantidade $\sqrt{q}$ de compromissos gerados é inatingível em tempo polinomial.

Expostas tais propriedades, ao se interpretar cada valor comprometido em um compromisso de Pedersen como um voto individual de um eleitor no contexto de alguma eleição, tais compromissos, proveem uma estrutura básica que viabiliza a construção de um protocolo eleitoral capaz de prover garantias de sigilo incondicional dos votos ao mesmo tempo em que lhes confere integridade computacional protegida pela intratabilidade do DLP em um grupo cautelosamente escolhido para a aplicação. A descrição de um tal protocolo é mostrada no capítulo seguinte.

Capítulo 4

O Protocolo

Descreve-se aqui um protocolo eleitoral de sigilo incondicional por meio do qual votações presenciais podem ser preparadas, executadas, totalizadas e auditadas de maneira transparente e independente de software, resguardando computacionalmente a integridade dos resultados eleitorais ante validade de algumas hipóteses matemáticas bem estabelecidas.

O protocolo proposto oferece capacidades de verificação de ponta-a-ponta que preservam o fluxo transitivo de integridade do voto, ou seja, permitindo a um eleitor individualmente conferir que seu voto é depositado conforme pretendido, registrado conforme depositado e totalizado conforme registrado. Ademais, o protocolo viabiliza que qualquer observador interessado possa se assegurar matematicamente de que todos os votos foram contabilizados na totalização conforme se encontram oficialmente registrados, independentemente de filiação partidária, participação direta no processo ou acesso privilegiado a informações restritas.

Para além de tais características, o protocolo descrito viabiliza ainda a realização de totalizações parciais que fruem das mesmas premissas e garantias da totalização final. Tal possibilidade confere a uma eleição uma granularização configurável da etapa de totalização que permite a autoridades e auditores identificar matematicamente pontos de ocorrência de falhas que porventura venham a ocorrer, podendo tais pontos então serem submetidos a auditorias específicas, recontagens manuais de votos ou quaisquer outras medidas corretivas porventura consideradas apropriadas.

Por fim, o protocolo apresenta características simbióticas que lhe permitem fácil acoplamento a processos eleitorais pré-existentes, mitigando esforços migratórios associados à evolução de sistemas eleitorais ao não exigir a completa e imediata reformulação destes. Isto permite fortalecer sistemas mais básicos, atribuindo-lhes capacidades E2E-V de robustez matemática e criptográfica.

Isto posto, a descrição do protocolo se inicia com a Seção 4.1, na qual constrói-se um arcabouço de definições com vistas a estabelecer com precisão a terminologia empregada no texto. Em seguida, a Seção 4.2 detalha minuciosamente o protocolo em termos de suas etapas individuais, estabelecendo e explorando os resultados matemáticos e criptográficos que embasam seu funcionamento. Logo após, a Seção 4.3 reitera as propriedades do protocolo à luz da descrição recém-estabelecida. Por fim, a Seção 4.4 descreve alguns

ataques de interesse ao protocolo em sua forma básica e estabelece possíveis formas de mitigação destes.

4.1 Definições Básicas

De modo a se lastrear com maior precisão a discussão acerca do protocolo eleitoral proposto, é interessante se estabelecer um conjunto básico e suficientemente robusto de definições. Ressalta-se, entretanto, que as seguintes definições conforme aqui colocadas não têm a pretensão de se ater a uma terminologia específica conforme utilizada por uma autoridade eleitoral em particular ou trabalho em contexto extrínseco ao deste escrito. Tampouco têm a pretensão de se apresentar como forma melhor ou mais correta de definirem-se os termos aos quais se referem. Conforme mencionado, apenas buscam oferecer uma base sólida para as descrições e discussões referentes ao protocolo descrito.

O conceito básico principal a ser definido na presente discussão é o que vem a ser uma *eleição*. Para os propósitos do protocolo proposto, uma eleição é compreendida da seguinte forma:

Definição 4.1.1 (Eleição)**.** *Uma* ***eleição*** *é o processo completo do fazer eleitoral no referente a seu cerne enquanto processo decisório, compreendendo suas etapas gerais de definição, preparação, votação, totalização e auditoria.*

É importante observar que, a depender do contexto, uma eleição em sua concepção mais geral pode abranger outras etapas, como cadastro de eleitores, propaganda eleitoral de candidatos, campanhas educativas direcionadas ao eleitor, entre outras. A Definição 4.1.1, no entanto, foca naquilo que concerne mais propriamente à execução do processo decisório participativo em si. Assim sendo, etapas adicionais como os supracitados exemplos, por sua considerável sensibilidade ao contexto de aplicação, são consideradas como subjacentes ao cerne do fazer eleitoral propriamente dito e, por conseguinte, também ao protocolo aqui descrito, podendo a ele ser agregadas conforme necessidade sem interferências nocivas a suas premissas de segurança.

Via de regra, espera-se que todas as etapas de uma eleição sejam iniciadas, executadas e concluídas dentro de um período temporal bem definido. Dessa forma, assume-se fazer sentido se referir a uma eleição específica como *ainda não iniciada*, *em execução* ou *já concluída*, predicados estes dependentes do instante de tempo em que são analisados. Este aspecto temporal associado a uma eleição é capturado pela definição do que vem a ser o *período eleitoral*, conceito definido a seguir:

Definição 4.1.2 (Período Eleitoral)**.** *O* ***período eleitoral*** *é o intervalo de tempo compreendido entre o instante de início da execução do cerne eleitoral e o instante a partir do qual o resultado eleitoral pode ser considerado oficialmente válido e legalmente incontestável.*

Como um exemplo advindo da realidade brasileira, assume-se como plausível, considerando-se o instante da escrita deste trabalho ao ano de 2021, referir-se a uma possível eleição nacional realizada no ano de 2042 como *ainda não iniciada*. Da mesma forma, assume-se plausível referir-se à eleiçào nacional do ano de 2014 como *já concluída*.

Note-se que, de modo a se evitar qualquer tipo de viés com relação àquilo que seriam os períodos de início e término de uma eleição em termos de interesses político-partidários, a definição exposta apenas engloba o referente ao cerne do processo decisório juntamente aos aspectos oficiais e legais que o governam, os quais são fortemente sensíveis ao contexto e, por isso, deixados em aberto na definição visando maior adaptabilidade do protocolo.

Uma eleição possui diversos parâmetros a ela associados e que, dada a generalidade do termo, podem ser dependentes do contexto eleitoral ou mesmo de requisitos oficiais e legais. Isto posto, diferencia-se aqui aqueles parâmetros que são intrinsecamente associados ao protocolo proposto, denominados *parâmetros eleitorais* e formalmente definidos a seguir:

Definição 4.1.3 (Parâmetro Eleitoral)**.** *Um* ***parâmetro eleitoral*** *é todo e qualquer parâmetro que de alguma forma tenha impacto matemático e/ou criptográfico sobre o protocolo eleitoral no que se refere à eleição.*

Observe-se assim que a definição de *parâmetro eleitoral* apenas abarca parâmetros que tenham impacto nas premissas de segurança, transparência, sigilo ou integridade exigidas para a correta execução do protocolo proposto, não tendo a conotação genérica de "qualquer parâmetro relacionado à eleição". Dessa forma, um grupo algébrico associado a compromissos de votos, uma função de hash a ser aplicada sobre descrições digitais dos votos produzidos e uma chave pública de assinatura de recibos de votos constituem parâmetros eleitorais. Por outro lado, o número de urnas a serem empregadas, a quantidade de mesários e seus respectivos nomes ou a descrição de um protocolo de segurança a ser adotado no dia da votação para evitar tumultos locais não constituem parâmetros eleitorais a priori.

Note-se, entretanto, que o que constitui ou não um parâmetro eleitoral não é estritamente inerente à natureza do parâmetro em si, mas sim ao seu relacionamento para com o protocolo eleitoral. Como exemplo, uma autoridade eleitoral pode exigir a inclusão da quantidade de mesários e seus nomes em uma representação digital da eleição que será passada por uma função de hash definida para gerar um resumo criptográfico público dessa eleição. Dessa forma, a alteração de um único mesário adulteraria o resumo

criptográfico resultante para algo diferente do originalmente publicado, o que poderia obrigar as autoridades, por exemplo, a prover satisfações acerca da motivação para a alteração promovida. Isto elevaria a quantidade e o nome dos mesários à condição de parâmetros eleitorais, visto que sua adulteração impacta no resumo criptográfico associado à eleição.

Assim sendo, embora alguns parâmetros sejam claramente classificáveis como *parâmetros eleitorais* a priori, outros parâmetros podem ser inclusos sob tal definição a depender do contexto de aplicação e demais exigências extrínsecas ao âmbito do protocolo. A especificação proposta busca ser relativamente genérica nesse sentido, incluindo apenas parâmetros eleitorais necessários a seu correto funcionamento no intuito de ser mais facilmente adaptável a contextos eleitorais diversos.

Dentre os vários parâmetros eleitorais possíveis, verificar-se-á que sobre alguns recaem requisitos de sigilo enquanto sobre outros recaem requisitos de transparência. Assim sendo, é possível separarem-se os diferentes parâmetros eleitorais em dois subgrupos distintos:

Definição 4.1.4 (Parâmetro Eleitoral Privado)**.** *Um* ***parâmetro eleitoral privado*** *é aquele sobre o qual as premissas de integridade e/ou transparência da eleição requerem que este seja mantido em sigilo.*

Definição 4.1.5 (Parâmetro Eleitoral Público)**.** *Um* ***parâmetro eleitoral público*** *é aquele sobre o qual as premissas de integridade e/ou transparência da eleição requerem que este seja divulgado ou mesmo publicamente gerado de forma aferível.*

A chave privada usada pelas autoridades eleitorais para assinatura dos votos válidos é um exemplo de parâmetro eleitoral privado: caso divulgada, qualquer um que tenha seu conhecimento pode assinar votos genéricos, fazendo-os passar por votos legítimos. Já a chave pública correspondente constitui um exemplo de parâmetro eleitoral público: caso não divulgada, não é possível a um auditor ou observador interessado verificar se os votos registrados e contabilizados são legítimos de fato.

A existência de parâmetros eleitorais privados evoca a necessidade de responsáveis por sua manutenção que sejam capazes de resguardar o sigilo esperado destes. Isto preconiza a existência de *guardiões eleitorais* de parâmetros privados:

Definição 4.1.6 (Guardião Eleitoral)**.** *Um* ***guardião eleitoral*** *é uma entidade responsável pela manutenção de um parâmetro eleitoral privado e das premissas de sigilo a ele inerentes.*

Dessa forma, a um guardião eleitoral é atribuído algum parâmetro eleitoral privado, o qual ficará sob sua responsabilidade. Como exemplo, a autoridade eleitoral responsável pela eleição é um guardião, haja vista ser ela a responsável pela chave de assinatura que confere legitimidade aos votos por ela assinados. No âmbito geral do protocolo, ver-se-á

que os guardiões desempenham conjuntamente um papel fundamental na manutenção do sigilo dos votos, tolhendo de qualquer entidade a possibilidade de violação unilateral do sigilo individual e, mediante correta execução do protocolo, conferindo aos votos sigilo incondicional.

Neste trabalho, faz-se uma diferenciação explícita entre os termos *eleição*, já anteriormente descrito na Definição 4.1.1, e *votação*, a qual é definida conforme segue:

Definição 4.1.7 (Votação)**.** *A **votação** é a etapa específica de uma eleição na qual os eleitores comunicam suas escolhas eleitorais por meio do depósito de seus votos individuais.*

Note-se aqui a distinção que se faz entre ambos os termos. Embora no dizer popular ambos sejam costumeiramente usados de maneira indiscriminada, no correr do presente escrito tais termos servem a propósitos intrinsecamente distintos, embora certamente relacionados: a *eleição* é o processo em si como um todo; a *votação* é a etapa particular desse processo em que o eleitor participa direta e ativamente com seu voto.

Tal definição clama por uma formalização mais estrita do que vem a ser considerado um *voto*. No contexto do presente trabalho, um voto é definido conforme o exposto na Definição 4.1.8 seguinte:

Definição 4.1.8 (Voto)**.** *Seja $\mathbb{O}$ o conjunto de todas as possíveis escolhas disponíveis a um eleitor em uma eleição. Um **voto** nessa eleição é um subconjunto $v \subseteq \mathbb{O}$.*

A Definição 4.1.8, apesar de sua aparente simplicidade, guarda em si uma complexidade sutil de consideráveis implicações: o *voto* em si é compreendido unicamente por um subconjunto particular de escolhas, mas não carrega consigo qualquer informação referente a quem o produziu. Esta separação é importante devido aos requisitos de sigilo inerentes a sistemas de votação, de modo que tal desvencilhamento identitário mostra-se de valia na concepção e análise desses sistemas.

No entanto, intuitivamente é de fácil constatação que nem todo subconjunto de opções necessariamente forma um voto *válido*: a depender das regras eleitorais, algumas combinações de opções podem não ser aceitas. Como um exemplo, um voto que contemple dois candidatos a um mesmo cargo em uma eleição na qual apenas um candidato pode ser escolhido não constituiria um voto válido. Isto implica na possibilidade de construção de um conjunto composto por subconjuntos de opções que são passíveis de produção em conformidade com as regras eleitorais vigentes, conceito este estabelecido na Definição 4.1.9 a seguir:

Definição 4.1.9 (Espaço de Votos)**.** *Seja $\mathbb{O}$ o conjunto previamente definido de opções disponíveis para escolha em uma eleição $\mathscr{E}$. Um **espaço de votos** é um conjunto V tal que $V = \{v | v \subseteq \mathbb{O}\}$, em que v representa um agregado de opções que constitui um voto em conformidade com as regras da eleição $\mathscr{E}$ considerada.*

Com base na Definição 4.1.9, torna-se agora possível melhor formalizar o que vem a ser considerado como um voto *válido*, conceito este estabelecido na Definição 4.1.10:

Definição 4.1.10 (Voto Válido)**.** *Seja V o espaço de votos de uma eleição $\mathscr{E}$. Um voto v é dito ser um **voto válido** se $v \in V$. Do contrário, v é dito ser um **voto inválido**.*

Note-se que o conceito de *espaço de votos* apresenta natureza fortemente regimentar, sendo intrínseco às regras de uma eleição particular, de modo que os conceitos de votos *válidos* ou *inválidos*, por definição, também o são. Assim, votos considerados válidos de acordo com as regras de uma eleição podem ser inválidos para uma outra eleição, ainda que o conjunto $\mathbb{O}$ considerado para ambas seja possivelmente idêntico.

Cabe ainda observar que os conceitos expostos de validade ou invalidade de votos constituem juízos técnicos acerca da aderência destes às regras da eleição à qual se referem. Não obstante, é possível se identificar votos que satisfazem tais regras, mas cujo depósito ou inclusão em totalizações são considerados indevidos e adulteram a integridade dos resultados eleitorais. Exemplos práticos destes são votos depositados por indivíduos que não eleitores válidos, ou votos depositados em duplicidade por um mesmo eleitor válido, ou ainda votos depositados em nome de um eleitor válido e à revelia deste. Tais votos, embora válidos segundo o crivo exposto na Definição 4.1.10, apresentam um caráter de indevibilidade em sua própria existência, sendo por isso classificados conforme a Definição 4.1.11 que se segue:

Definição 4.1.11 (Voto Indevido)**.** *Sendo v um voto válido, v é dito ser um **voto indevido** se seu depósito, registro ou contemplação na totalização infringem as regras da eleição.*

Um outro tipo interessante de voto que pode ser identificado, principalmente no contexto de sistemas eleitorais criptográficos, é aquele que se pretende como fiel às intenções de um eleitor, mas que em realidade representa um conjunto de opções que não reflete tais intenções. O conceito é formalmente estabelecido na Definição 4.1.12 a seguir:

Definição 4.1.12 (Voto Inconforme)**.** *Seja v um voto válido cujo conjunto de opções manifesta precisamente as escolhas pretendidas por um eleitor $\mathscr{e}$. Um voto $v' \neq v$ é dito ser um **voto inconforme** se o eleitor $\mathscr{e}$ é de alguma forma induzido, à sua revelia, a depositar v' em lugar de v.*

Como exemplo, seja um sistema eleitoral criptográfico em que, por questões de sigilo, votos são apenas apresentados aos eleitores de forma cifrada. Sendo v o voto pretendido por um eleitor $\mathscr{e}$, o sistema poderia apresentar a $\mathscr{e}$ um criptograma $\mathsf{ENC}(v')$ como seu suposto voto cifrado, em que $v' \neq v$. Dada a camada criptográfica sobre o voto, o eleitor é incapaz de identificar a adulteração, sendo de tal forma induzido a depositar um voto distinto daquele por ele originalmente pretendido.

É importante notar que a definição de *voto inconforme* não diz respeito a uma característica absoluta associada a um voto $v \in V$, mas sim à sua ligação com relação às intenções de voto de um eleitor $\mathscr{e}$ específico. Isto posto, note-se que um voto v pode ser inconforme para um eleitor $\mathscr{e}_1$, mas não para um eleitor $\mathscr{e}_2$.

As definições mostradas nessa seção constituem um conjunto básico da terminologia que será empregada na descrição do protocolo e discussões a este referentes no contexto deste trabalho. Definições adicionais serão expostas durante a descrição conforme necessárias.

4.2 Descrição

Esta seção descreve o protocolo eleitoral proposto por este trabalho em maiores detalhes, com ênfase nas propriedades matemáticas e criptográficas que lastreiam seu funcionamento e suas garantias. Definições adicionais são providas conforme necessidade e resultados associados às propriedades do protocolo são descritos conforme venham a ser discutidos.

O protocolo eleitoral é subdividido em cinco etapas distintas, as quais são executadas em caráter sequencial dentro de um intervalo de tempo específico denominado *período eleitoral*, conforme expõe a Definição 4.1.2. As etapas do protocolo são sucintamente mostradas na listagem a seguir:

1. **Definição:** etapa em que os aspectos básicos associados à eleição são devidamente definidos e publicados.

2. **Preparação:** etapa em que são gerados parâmetros públicos e privados necessários à devida execução do protocolo.

3. **Votação:** etapa em que ocorre a participação direta dos eleitores pelo depósito individual de votos nas opções disponíveis para escolha na eleição.

4. **Totalização:** etapa de agregação dos votos individuais em totalizações que constituirão os *resultados eleitorais*.

5. **Auditoria:** etapa pós-eleitoral que permite verificação individual e pública da correção dos resultados eleitorais.

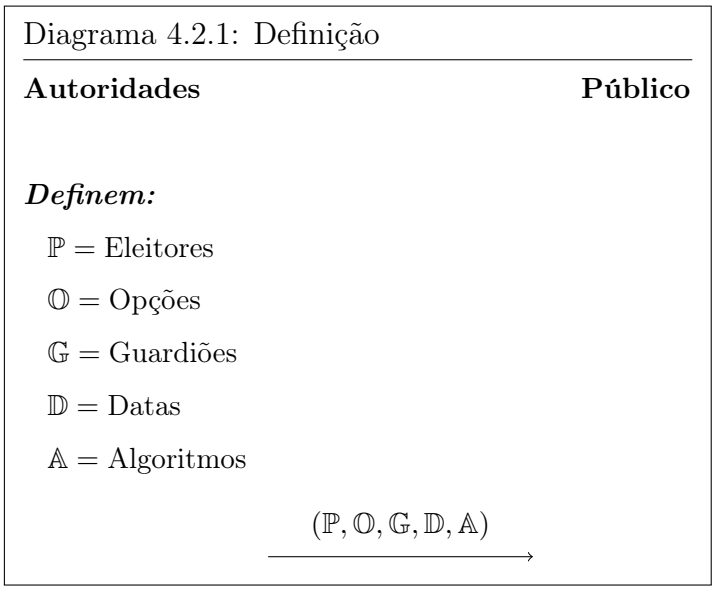
Cada uma dessas etapas é explorada individualmente e em maiores detalhes nas seções que se seguem.

4.2.1 Definição

A etapa de *definição* compreende a determinação de aspectos básicos associados a uma eleição. O que vem a ser considerado "básico" e quais aspectos se enquadram em tal categoria são questões fortemente intrínsecas ao contexto de aplicação, legislações vigentes e possivelmente outros requisitos de natureza diversa. A intenção desta etapa é definir todos os aspectos que precisam já estar estabelecidos quando do início da etapa de preparação eleitoral seguinte.

Assim sendo, não se busca aqui prover uma listagem exaustiva dos aspectos que possam ser definidos nesta etapa, mas sim apresentar um subconjunto destes que se enquadram como *parâmetros eleitorais* e, portanto, cuja definição é considerada sensível à correta execução do protocolo. O Diagrama 4.2.1 mostrado a seguir explicita tais aspectos:

Diagrama 4.2.1: Definição



Na etapa de *definição*, mostrada no Diagrama 4.2.1, as autoridades eleitorais utilizam-se de dados (e.g. listagens de eleitores aptos a votar, listagens de candidaturas) e legislações vigentes para determinar os seguintes conjuntos:

- **Eleitores ($\mathbb{P}$):** o conjunto de entidades (e.g. pessoas, cidadãos) que podem depositar votos nas opções disponíveis para escolha na eleição em definição.
- **Opções ($\mathbb{O}$):** o conjunto de opções (e.g. candidatos) disponíveis para escolha por parte dos eleitores da eleição.

- **Guardiões ($\mathbb{G}$):** o conjunto de entidades (e.g. indivíduos, instituições etc) responsáveis pela guarda e sigilo de parâmetros eleitorais privados necessários à correta execução do protocolo.

- **Datas ($\mathbb{D}$):** o conjunto das datas que especificam prazos para a execução das diversas atividades associadas à eleição em definição.

- **Algoritmos ($\mathbb{A}$):** o conjunto dos algoritmos e implementações computacionais que serão empregados no decorrer da execução eleitoral.

A necessidade da definição do conjunto de eleitores ($\mathbb{P}$) e do conjunto de opções ($\mathbb{O}$) nesta primeira etapa motiva-se por conta de a cardinalidade de tais conjuntos poder apresentar impacto direto sobre os parâmetros de segurança a serem usados pelo protocolo. Ademais, tais conjuntos certamente precisam estar totalmente determinados previamente à etapa de votação, quando serão usados para controle de acesso de eleitores e para determinar as opções a serem mostradas a um eleitor em particular.

Por sua vez, os guardiões eleitorais serão necessários já na etapa seguinte, de preparação da eleição, quando a estes será atribuída a guarda de parcelas de uma chave privada de decifração de *códigos de segurança* associados a cada voto individual, como se há de ver em maiores detalhes quando da descriçao da etapa de *votação*. Desse modo, faz-se necessário formalizar a composição do conjunto de guardiões ($\mathbb{G}$) em caráter prévio.

Com relação às datas, apesar de estas se tratarem de aspecto consideravelmente sensível ao contexto de aplicação, assume-se aqui ser possível estabelecer um conjunto $\mathbb{D}$ que seja necessariamente capaz de definir, de maneira inequívoca, o chamado *período eleitoral*, conforme consta na Definição 4.1.2. A definição desse período de tempo é crucial para a determinação dos parâmetros de segurança associados ao protocolo eleitoral, haja vista este depender de premissas que devem se manter válidas, sob hipóteses computacionais realistas, durante todo este período.

No que se refere ao conjunto de algoritmos ($\mathbb{A}$), o devido estabelecimento destes se mostra de crucial importância para a transparência do processo, possibilitando ao público geral a avaliação, ou livre terceirização desta, quanto ao âmbito computacional da eleição. Ademais, é importante ressaltar que os aspectos de segurança do pleito estão intrinsecamente ligados aos diversos parâmetros eleitorais produzidos para sua realização, parâmetros estes advindos diretamente da execução desses algoritmos. Dessa forma, sua publicação se mostra como elemento de fundamental importância tanto para os supracitados fins de transparência quanto para análises referentes a sua correção e/ou aplicabilidade aos fins propostos.

Após definição dos supracitados conjuntos, estes são devidamente publicados em canais oficiais e deixados acessíveis. No referente a aspectos que, a depender do contexto de aplicação, possam ser considerados de caráter sigiloso, como os elementos particulares

do conjunto $\mathbb{P}$ de eleitores, a forma de publicação pode ficar a cargo de determinações ou legislações locais, (e.g. publicação apenas de nomes ou números de registro eleitoral parciais). Em último caso, apenas a cardinalidade de $\mathbb{P}$ pode ser divulgada ao público geral. Não obstante, a plena determinação do conjunto de eleitores válidos é de crucial importância no sentido de determinar aqueles que podem ou não participar da etapa de *votação*.

Publicados os conjuntos decorrentes da etapa de definição, e após tempo suficiente para possíveis contestações legais ou procedimentos afins que venham a alterar a composição de tais conjuntos, prossegue-se à etapa de *preparação* eleitoral.

4.2.2 Preparação

A *preparação* compreende a etapa seguinte do protocolo, em que diversos parâmetros eleitorais de caráter matemático e criptográfico são definidos para uso nas etapas posteriores do processo eleitoral. Conforme previamente mencionado, tais parâmetros, a depender das premissas que sobre eles recaem, podem ser classificados como *públicos* ou *privados*. Dado que a etapa de *preparação* exige a geração de diversos parâmetros eleitorais públicos, esta deve ser executada por meio da realização de uma cerimônia formal, preferencialmente presencial e transmitida ao vivo ao público geral, da qual devem participar indivíduos e entidades com distintas filiações e interesses políticos no processo eleitoral e seus resultados. Assim sendo, a preparação eleitoral no contexto de uma tal cerimônia é esquematizada no Diagrama 4.2.2 seguinte, cujos componentes serão devidamente explanados posteriormente:

Diagrama 4.2.2: Preparação (Cerimônia)

Autoridades	**Guardiões**	**Público**
Definem:		
$\mathbb{W}$ (grupo)		
$\alpha, \beta \in \mathbb{W}$ (geradores)		
$H : V \to \mathscr{V}$ (hash de votos)		
Geram:		
$pub, priv = \mathsf{GEN}_{\mathsf{HOM}}(1^n)$		
$k = \mathsf{GEN}_{\mathsf{SYM}}(1^n)$		
$c_{priv} = \mathsf{ENC}_k(priv)$		
$\mathbb{S} = \mathsf{SHARE}(k, \lvert\mathbb{G}\rvert, t)$		
	$\xrightarrow{(s \in \mathbb{S}, c_{priv})}$	
Calculam:		
$h = \mathsf{HASH}(\mathbb{W}, \alpha, \beta, H, pub)$		
	$\xrightarrow{(\mathbb{W}, \alpha, \beta, H, pub, h)}$	

Com base nos parâmetros produzidos pela etapa de *definição*, as autoridades eleitorais, seguindo metodologia (e.g. máquinas, algoritmos) previamente especificada, iniciam a realização da cerimônia pública pela definição dos seguintes parâmetros eleitorais públicos:

- Um *grupo algébrico* $\mathbb{W}$;
- Dois *geradores independentes* $\alpha, \beta \in \mathbb{W}$;
- Um *hash de votos* $H(v)$.

Em seguida, junto aos guardiões eleitorais determinados durante a etapa de *definição*, as autoridades geram *parcelas* que serão distribuídas entre estes guardiões em acordo com o seguinte processo:

- Gera-se um par de chaves $(pub, priv)$ de uma cifra aditivamente homomórfica (aspecto denotado pelo subíndice HOM) em acordo com um parâmetro de segurança n associado à eleição;

- Adicionalmente, gera-se uma chave k para cifração simétrica (aspecto denotado pelo subíndice SYM), também em acordo com o parâmetro de segurança n da eleição;
- A chave simétrica k é então usada para cifrar a chave privada $priv$, produzindo um criptograma c_{priv};
- Um conjunto $\mathbb{S}$ de $|\mathbb{G}|$ parcelas da chave simétrica k é gerado utilizando-se de um esquema de compartilhamento de segredos;
- Cada uma das parcelas de $\mathbb{S}$ é entregue unicamente a um guardião, o qual fica responsável por sua tutela e sigilo.

Ao final, os parâmetros eleitorais públicos gerados são agregados em um resumo criptográfico e publicados em caráter oficial. O significado de tais parâmetros no contexto do protocolo eleitoral, bem como os requisitos impostos sobre o processo de geração/definição de cada qual, são melhor explorados nas subseções seguintes.

4.2.2.1 Definição do Grupo Algébrico

O primeiro parâmetro público a ser definido na cerimônia é um grupo algébrico $\mathbb{W}$ abeliano, cíclico e finito, no qual a hipótese do logaritmo discreto (c.f. Definição 3.1.6) seja considerada válida. Dado o inevitável crescimento do poder computacional disponível, tanto no sentido de aparato computacional mais potente quanto de possíveis algoritmos mais eficientes, é esperado que, decorrido um intervalo de tempo suficientemente longo, a hipótese do logaritmo discreto no grupo em questão deixe de ser válida e o DLP passe a ser considerado como solúvel nesse grupo. Isto preconiza a necessidade de uma definição mais precisa do grupo a ser gerado no referente à intratabilidade do DLP. Assim, a menção a um *grupo* no contexto do protocolo necessariamente refere-se ao conceito formalizado pela Definição 4.2.1 a seguir:

Definição 4.2.1 (Grupo)**.** *Um* ***grupo*** *no contexto do protocolo é um grupo algébrico abeliano, cíclico e finito no qual a hipótese do logaritmo discreto é matematicamente considerada válida, ante aparato computacional vigente e projeções satisfatoriamente realistas daquele que há de vir, durante todo o período eleitoral.*

Observe-se que, em acordo com a Definição 4.2.1, o que vem ou não a ser um "grupo" no protocolo é sensível à eleição particular na qual o protocolo é empregado, mais especificamente no referente ao período eleitoral considerado. Assim sendo, a definição deixa implícita a existência de um parâmetro de segurança associado ao grupo a ser usado:

eleições com um período eleitoral de maior duração exigem um parâmetro de segurança mais robusto de modo a garantir que o grupo algébrico sobre o qual o protocolo opera seja de fato um grupo em respeito à Definição 4.2.1.

Dessa forma, o grupo algébrico $\mathbb{W}$ a ser gerado durante a cerimônia deve necessariamente levar em consideração o período eleitoral previamente definido por $\mathbb{D}$ no intuito de prover garantias satisfatórias de que $\mathbb{W}$ se comporte como um *grupo* no âmbito do protocolo. Assim sendo, note-se que, à luz da Definição 4.2.1, o DLP entre quaisquer dois elementos obtidos uniformemente de um grupo apenas pode ser resolvido com probabilidade negligenciável durante o intervalo de tempo do período eleitoral em questão.

4.2.2.2 Definição dos Geradores Independentes

Produzido o grupo $\mathbb{W}$, dois geradores desse grupo, aqui denominados α e β, devem ser obtidos. Tais geradores devem necessariamente resguardar a propriedade de que o DLOG entre ambos seja desconhecido, aspecto formalizado na Definição 4.2.2 seguinte:

Definição 4.2.2 (Geradores Independentes)**.** *Seja $\mathbb{W}$ um grupo e sejam dois elementos $\alpha, \beta \in \mathbb{W}$ tais que $\langle\alpha\rangle = \langle\beta\rangle = \mathbb{W}$ e $\log_\alpha \beta = x$. Seja ainda $\mathscr{A}_{PPT}$ um algoritmo probabilístico polinomial qualquer que recebe como entrada dois geradores de $\mathbb{W}$ e retorna como saída o logaritmo discreto entre ambos. Os elementos α e β são **geradores independentes** de $\mathbb{W}$ se a seguinte propriedade é válida durante todo o período eleitoral:*

$$\forall \mathscr{A}_{PPT} : Pr[\mathscr{A}_{PPT}(\alpha, \beta) = x] \leq negl(n),$$

em que $negl(n)$ é uma função negligenciável no parâmetro de segurança n associado ao grupo $\mathbb{W}$ em questão.

Em complemento à Definição 4.2.2 anterior, coloca-se a seguir, para fins de completeza e melhor definição da terminologia usada, a Definição 4.2.3:

Definição 4.2.3 (Geradores Dependentes)**.** *Sendo $\mathbb{W}$ um grupo, dois elementos $\alpha, \beta \in \mathbb{W}$ tais que $\langle\alpha\rangle = \langle\beta\rangle = \mathbb{W}$ são **geradores dependentes** se, e somente se, **não** são geradores independentes, ou seja, existe algoritmo probabilístico polinomial capaz de computar $\log_\alpha \beta$ durante o período eleitoral.*

Embora talvez trivialmente observável, explicita-se aqui que a Definição 4.2.1 por si só não encerra as Definições 4.2.2 e 4.2.3. Isto significa que o fato de serem α e β geradores de um grupo algébrico apto à aplicação no protocolo não garante a independência entre ambos, como ressalta o Teorema 4.2.4:

Teorema 4.2.4. *Seja $\mathbb{W}$ um grupo de ordem q e um elemento $\alpha \in \mathbb{W} : \langle\alpha\rangle = \mathbb{W}$. Então um elemento $\beta \in \mathbb{W} : \langle\beta\rangle = \mathbb{W}$ pode ser facilmente obtido de modo que α e β sejam geradores dependentes.*

Demonstração. Suponha-se um grupo $\mathbb{W}$ de ordem q e um gerador α desse grupo. Um gerador β pode ser obtido a partir de α usando-se do seguinte fato algébrico:

$$\forall x \in \mathbb{Z}_q : \gcd(x, q) = 1 \implies \langle\alpha^x\rangle = \mathbb{W}$$

Como certamente $\gcd(q-1, q) = 1$, tem-se que $\beta = \alpha^{q-1}$ é um gerador de $\mathbb{W}$. Mas isto implica que $\log_\alpha \beta = q - 1$ e, dado que q é conhecido, computa-se facilmente o logaritmo entre os geradores α e β. Assim sendo, por definição, tem-se que α e β são geradores dependentes. □

Isto posto, é crucial que os geradores produzidos nesta etapa da cerimônia sejam *geradores independentes*. A motivação para tal necessidade tornar-se-á mais clara adiante na descrição do protocolo. Cabe ainda ressaltar que a independência diz respeito à improbabilidade polinomial de determinação da relação entre os dois gerados por meio de seu logaritmo, mas não estipula a forma como os dois geradores serão produzidos. Por exemplo, não há problemas em que α seja fixado, desde que β seja aleatório, de relação logarítmica desconhecida com relação a α.

4.2.2.3 Definição do Hash de Votos

Por fim, uma função $H : V \to \mathscr{V}$ de mapeamento de votos para *representações binárias* destes deve ser precisamente definida, em que V é o *espaço de votos* (c.f. Definição 4.1.9) e $\mathscr{V}$ denota o conjunto das representações binárias destes. Esta função deve ser capaz de mapear qualquer voto válido passível de produção durante a etapa de *votação* em uma *representação binária* univocamente correpondente a este.

Em eleições nas quais os votos em si sejam armazenados digitalmente, o espaço de votos já corresponde a representações digitais dos mesmos, de modo que é possível definir-se $V = \mathscr{V}$ e H como a função identidade. Já em contextos em que não exista uma etapa inerente de digitalização do voto (e.g. voto em papel), tem-se necessariamente $V \neq \mathscr{V}$, de modo que uma representação binária deve ser derivada a partir da composição de cada voto físico em si. Neste caso, cada voto físico necessariamente precisa ser submetido a uma etapa intermediária de "digitalização" seguindo-se alguma metodologia bem estabelecida.

Entretanto, em ambos os casos, uma digitalização pura do voto pode render sua representação digital sujeita a variações inerentes à própria composição dos votos em si (e.g.

representações digitais com tamanhos diferentes a depender de suas opções componentes). Assim sendo, no intuito de mitigar esta possível variabilidade de forma, é de interesse ao protocolo eleitoral aqui estabelecido restringir as representações digitais a sequências binárias de tamanhos definidos e uniformes por meio da aplicação de alguma função de *hash* escolhida em acordo com os parâmetros eleitorais (e.g. $\mathscr{V} = \{hash(v) : \forall v \in V\}$).

Isto posto, independentemente da metodologia empregada na produção do conjunto $\mathscr{V}$ de representações binárias, é de crucial importância que tal conjunto retenha certas propriedades, as quais são formalizadas sob o conceito de *espaço digital de votos*, exposto na Definição 4.2.5 seguinte:

Definição 4.2.5 (Espaço Digital de Votos)**.** *Seja V o espaço de votos de uma eleição $\mathscr{E}$. O conjunto $\mathscr{V}$ é dito ser um* ***espaço digital de votos*** *de $\mathscr{E}$ se, e somente se, $\forall \boldsymbol{v} \in \mathscr{V} : \boldsymbol{v} = \{0,1\}^m$, $m \in \mathbb{N}$, e existe uma função injetora $H : V \to \mathscr{V}$. Adicionalmente, $\mathscr{V}$ é dito ser* ***uniforme*** *se, e somente se, $\forall \boldsymbol{v}_i, \boldsymbol{v}_j \in \mathscr{V} : |\boldsymbol{v}_i| = |\boldsymbol{v}_j|$.*

A Definição 4.2.5 em sua essência especifica que um espaço digital de votos $\mathscr{V}$ é uma codificação binária do espaço de votos V, devendo resguardar a unicidade esperada de cada voto válido de uma eleição. Assim sendo, dois votos válidos que apresentem ao menos uma opção distinta entre si devem ter representações necessariamente distintas em $\mathscr{V}$, refletindo o fato de que codificam subconjuntos diferentes de opções eleitorais. Por outro lado, dois votos válidos referentes ao mesmo conjunto de opções devem sempre ser codificados como uma mesma representação binária em $\mathscr{V}$. A característica de uniformidade associada a um espaço digital de votos apenas visa garantir que todas as codificações apresentam o mesmo tamanho binário (e.g. 256 bits).

Note-se assim que o espaço digital de votos especifica precisamente o contradomínio da supracitada função H de mapeamento de votos para representações digitais. Assumindo-se $\mathscr{V}$ como sendo uniforme, a função de mapeamento H opera de forma similar a uma função de *hash* criptográfico (e.g. SHA256). Isto posto, é possível identificar alguns requisitos comuns às funções de *hash* e à função H. Em particular, para fins de aplicabilidade ao protocolo, é crucial a H ser *resistente à segunda pré-imagem* e *resistente a colisões*. A resistência à segunda pré-imagem visa tornar intratável, dado um voto $v_1 \in V$, encontrar um segundo voto $v_2 \in V$ tal que $H(v_1) = H(v_2)$. Por sua vez, a resistência a colisões dificulta a obtenção de votos $v_1, v_2 \in V$ tais que $H(v_1) = H(v_2)$.

Observando-se que a resistência a colisões implica em resistência à segunda pré-imagem [39], uma função H que satisfaça a primeira dessas propriedades é dita ser um *hash de votos*, conceito formalmente estabelecido na Definição 4.2.6 seguinte:

Definição 4.2.6 (Hash de Votos)**.** *Seja $H : V \to \mathscr{V}$ uma função que mapeia o espaço de votos V em um espaço digital de votos $\mathscr{V}$ para uma dada eleição $\mathscr{E}$. A função H é dita ser um* ***hash de votos*** *se, durante todo o período eleitoral, ela resguardar a propriedade de resistência a colisões.*

Tomando por base a Definição 4.2.6, note-se que um hash de votos resguarda propriedades similares àquelas esperadas de funções de hash criptográficas. Embora estas últimas devam ainda resguardar a propriedade de *resistência à pré-imagem*, tal propriedade não é necessária a um hash de votos no contexto do protocolo aqui descrito. Isto é, dado um valor obtido pela aplicação do hash de votos, não há qualquer prejuízo ao protocolo ou suas premissas de segurança em ser de fácil dedução qual voto originou tal valor. A não exigência da resistência à pré-imagem para o hash de votos mostra-se como uma decisão de projeto razoável considerando-se que o espaço de votos V que constitui o domínio dessa função é em geral de caráter suficientemente limitado a ponto de possibilitar buscas extensivas por mapeamentos entre votos e hashes. Hipóteses de segurança lastreadas no desconhecimento desses mapeamentos, portanto, mostrar-se-iam como potencialmente danosas às premissas do protocolo. Contudo, não há qualquer prejuízo à finalidade de um hash de votos caso este se utilize de uma função que por base se mostre como resistente a esforços de computação de pré-imagens.

Isto posto, funções de hash criptográficas tradicionais que ainda retenham suas propriedades de hash, por terem passado pelo crivo do tempo e da criptoanálise pública, provavelmente proveem um respaldo mais aferível no referente a aspectos de segurança das propriedades delas esperadas se comparadas com hashes eleitorais criados em caráter *ad-hoc* para uma eleição individual. Isto porque, dado o tempo, possivelmente exíguo, reservado ao período eleitoral, um hash criado especificamente para o propósito da eleição em questão pode carecer de testes, verificações e comprovações matematicamente satisfatórias que lhe atestem como sendo, de fato, um hash de votos. Assim, funções de hash criptográficas fornecem um lastro suficientemente robusto que pode ser aproveitado quando da definição de uma função de hash de votos a ser empregada em uma eleição particular.

4.2.2.4 Geração das Parcelas dos Guardiões

Como se haverá de ver quando da descrição da etapa de *votação*, o sigilo dos votos individuais dos eleitores é resguardado por um *código de segurança*, que consiste em um valor inteiro aleatório gerado em conformidade com a ordem do grupo $\mathbb{W}$ definido para a eleição. Também como se haverá de ver, o conjunto desses códigos de segurança adicionalmente habilita uma equação de auditoria universal que, a menos de uma probabilidade negligenciável, deve ser respeitada matematicamente por qualquer eleição íntegra. Isto posto, é importante que cada qual desses códigos de segurança seja protegido por um esquema de cifração que deve possibilitar tratamento homomórfico aditivo dos criptogramas

produzidos.

Assim, finda a geração dos parâmetros eleitorais previamente descritos, as autoridades devem gerar, junto aos guardiões eleitorais, um par de chaves criptográficas mediante execução de um algoritmo previamente acordado referente a um esquema de cifração homomórfico de chave pública. Isto é, gera-se o par $(pub, priv) = \mathsf{GEN}_{\mathsf{HOM}}(1^n)$, em que $\mathsf{GEN}_{\mathsf{HOM}} \in \mathbb{A}$ é o algoritmo estabelecido para geração das chaves, n é um parâmetro de segurança associado à eleição e o par $(pub, priv)$ corresponde às chaves geradas em si. A chave pública pub é extraída para publicação e a chave privada $priv$ é mantida unicamente em memória para cifração posterior. O subíndice HOM denota que o esquema de cifração utilizado é aditivamente homomórfico, permitindo operação de adição de artefatos cifrados via operações em seus respectivos criptogramas.

Finda a geração do par de chaves para cifração homomórfica, uma chave simétrica k é gerada também por um algoritmo previamente acordado. Desse modo, produz-se a chave $k = \mathsf{GEN}_{\mathsf{SYM}}(1^n)$, em que $\mathsf{GEN}_{\mathrm{SYM}} \in \mathbb{A}$ é um algoritmo de geração estabelecido, n é um parâmetro de segurança associado à eleição (não necessariamente idêntico ao previamente aplicado em $\mathsf{GEN}_{\mathsf{HOM}}$) e k é a chave simétrica em si. Esta chave k é então usada junto a um algoritmo de cifração $Enc_k \in \mathbb{A}$ previamente estabelecido para cifrar a chave privada $priv$, gerando-se assim um criptograma $c_{priv} = Enc_k(priv)$. Note-se que a decifração de c_{priv} requer a chave k e permite a obtenção da chave privada $priv$. Como será descrito posteriormente, a obtenção de $priv$ em um contexto controlado será necessária para habilitar a verificação da equação de auditoria universal.

Isto feito, um algoritmo $\mathsf{SHARE} \in \mathbb{A}$ de compartilhamento de segredos, a exemplo daquele descrito por [42], é aplicado sobre a chave simétrica k, produzindo assim um conjunto $\mathbb{S} = SHARE(k, |\mathbb{G}|, t)$ de parcelas dessa chave tal que $|\mathbb{S}| = |\mathbb{G}|$. De modo a evitar que um guardião ou pequeno grupo destes possa impedir a verificação universal da integridade eleitoral, é prudente que seja definido um valor $t \leq |\mathbb{G}|$ tal que uma congregação de ao menos t guardiões consiga obter a chave k parcelada em $\mathbb{S}$. O valor de t é um aspecto demasiadamente sensível e possivelmente de forte dependência para com um contexto eleitoral particular, de modo que estipular valores precisos para este parâmetro é uma tarefa que foge ao escopo da descrição do presente protocolo. Idealmente, t deve ser suficientemente grande para impedir pequenas congregações de guardiões conspirantes de obter a chave $priv$, mas também suficientemente pequeno de modo a não render a possibilidade de verificação da integridade eleitoral aos desígnios dessas mesmas pequenas congregações.

Produzido esse conjunto $\mathbb{S}$, cada parcela individual $s \in \mathbb{S}$ deve ser unicamente entregue, junto a uma cópia de c_{priv}, a um dos guardiões previamente estipulados quando da etapa de *definição*. Isto deve ser feito de modo que nenhum guardião, ou mesmo as autoridades eleitorais, tenha acesso ou conhecimento das parcelas de quaisquer dos demais guardiões. Assim, cada guardião recebe um par (s, c_{priv}), em que s é a parcela particular

daquele guardião e c_{priv} é uma cópia do criptograma gerado pela cifração de $priv$ com k. Cada guardião deve ficar responsável pelo devido armazenamento e sigilo de suas parcelas e, posteriormente, findo o período eleitoral, pela correta destruição da mesma.

4.2.2.5 Conclusão da Cerimônia

A cerimônia é concluída pela formalização dos parâmetros eleitorais públicos produzidos em seu âmbito. Para tanto, representações digitais padronizadas desses parâmetros devem ser agregadas, juntamente a versões digitais de quaisquer documentos pertinentes ou legalmente exigidos, e um hash criptográfico desse conjunto é produzido. Tal valor de hash, bem como o supracitado conjunto, são o resultado final da etapa cerimonial preparatória. Note-se que o hash de votos previamente especificado é semanticamente distinto do hash aqui utilizado: enquanto aquele visa mapear um espaço de votos para um espaço digital de votos, este visa unicamente prover um resumo criptográfico aferível do agregado público correspondente à eleição em preparação. Nada em princípio impede, entretanto, que ambos façam uso de um mesmo lastro criptográfico (e.g. SHA256).

Tal valor de hash resultante do processo deve então ser divulgado em canais oficiais de comunicação, digitais ou fisicamente impressos. Adicionalmente, o conjunto compreendido por todos os parâmetros eleitorais públicos produzidos deve ser disponibilizado digitalmente para acesso, pelo menos durante todo o período eleitoral, via algum canal oficial e publicamente acessível. Este conjunto de parâmetros é de crucial importância para as etapas eleitorais seguintes, e, como se há de ver, sua publicação é fundamental para fins de auditoria universal das totalizações produzidas.

4.2.3 Votação

Executada a cerimônia de preparação da eleição e decorrido algum tempo para a devida divulgação de seus parâmetros e possíveis contestações legais, inicia-se a etapa de *votação*. Nesta etapa, eleitores registrados como participantes da eleição são chamados a comparecer presencialmente a locais designados de votação para depositar seus votos nos candidatos ou opções disponíveis para escolha.

O processo de votação aqui descrito exige a operação, em caráter individual, de uma máquina devidamente preparada para receber as opções do eleitor e gerar uma cédula

contendo parâmetros intrínsecos ao voto nela representado. Esta cédula é impressa via uma impressora acoplada diretamente à máquina de votação e é então entregue ao eleitor, que tem a opção de depositar seu voto e encerrar o processo ou desafiar a máquina a demonstrar a conformidade da cédula gerada com o voto pretendido.

Tais procedimentos são executados por cada qual dos eleitores oficializados para a eleição. Mais precisamente, na etapa de votação cada eleitor deve cumprir em sucessão o seguinte fluxo sequencial, em que a etapa de desafio da cédula é opcional:

1. **Identificação do eleitor:** cada eleitor que pretenda depositar um voto precisa ser devidamente identificado como um eleitor oficializado, isto é, constante no conjunto $\mathbb{P}$ de eleitores previamente definido;

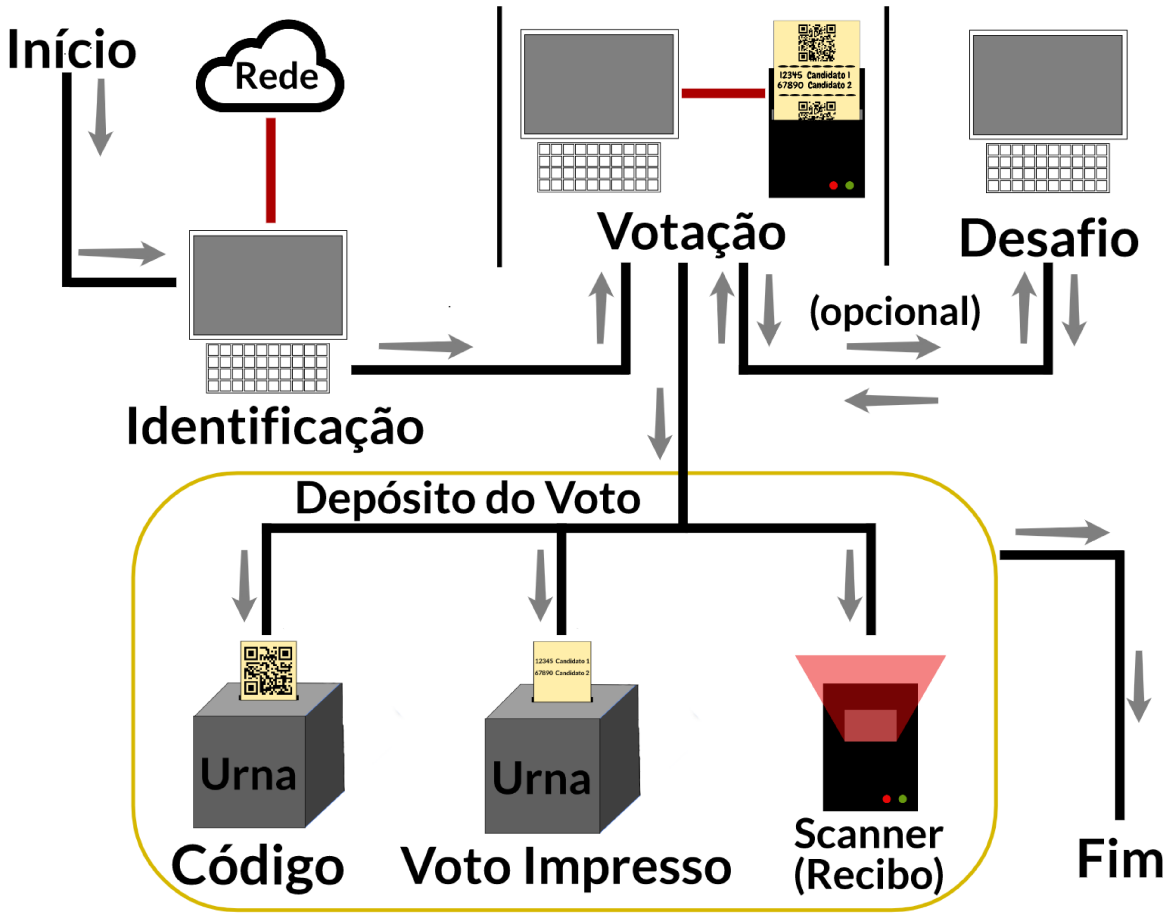
2. **Geração da cédula:** o eleitor devidamente identificado manifesta suas opções a uma máquina, que então gera uma cédula impressa;

3. **Desafio da cédula (opcional):** de posse de sua cédula, o eleitor pode optar por desafiar o sistema a demonstrar que sua cédula está correta;

4. **Depósito do voto:** estando devidamente satisfeito com suas escolhas, o eleitor deposita sua cédula para que esta seja considerada na etapa posterior de totalização.

Esta sequência, aqui coletivamente denominada de *fluxo de votação*, é mostrada no diagrama exposto na Figura 4.1:

Figura 4.1: O fluxo de votação



Fonte: Elaborado pelo autor.

A Figura 4.1 reúne todos os supracitados passos sucessivos que compreendem o fluxo de votação. Os segmentos de reta pretos, junto a suas setas direcionais associadas, indicam o movimento do eleitor pelo fluxo eleitoral. Os segmentos de reta vermelhos, por sua vez, ressaltam possíveis interconexões críticas esperadas em diferentes máquinas utilizadas no processo. Por fim, o contorno em amarelo que engloba as urnas e o scanner de recibos ressalta a atomicidade necessária ao processo de depósito.

As etapas individuais do fluxo eleitoral, à luz do diagrama exposto na Figura 4.1, são exploradas em maiores detalhes nas discussões colocadas nas subseções que se seguem.

4.2.3.1 Identificação do Eleitor

O primeiro passo do processo de votação compreende a necessidade intrínseca de identificação do eleitor no intuito de garantir-lhe acesso ao restante do fluxo de votação. Isto é importante por dois motivos principais: primeiro, para impedir que indivíduos não

reconhecíveis como eleitores oficializados tenham acesso ao restante do fluxo; e segundo, para evitar que eleitores oficializados sejam capazes de depositar mais de um voto.

Isto posto, espera-se que a metodologia particular de análise decisória para definir o acesso de cada eleitor ao fluxo de votação seja fortemente sensível a aspectos técnicos e legais, de modo que o presente protocolo não se propõe a especificar um padrão rígido para tanto. De maneira geral, assume-se aqui a existência de alguma forma de base de dados, física ou digital, sob controle das autoridades eleitorais e em conformidade com o conjunto $\mathbb{P}$ de eleitores previamente definido e publicado. Tal base de dados é assumida como mantendo uma listagem extensiva dos eleitores que podem ter acesso ao fluxo e é consultada de modo a garantir acesso à votação apenas a tais eleitores oficializados, devendo ser imediatamente atualizada de forma a refletir o depósito do voto por cada eleitor em particular.

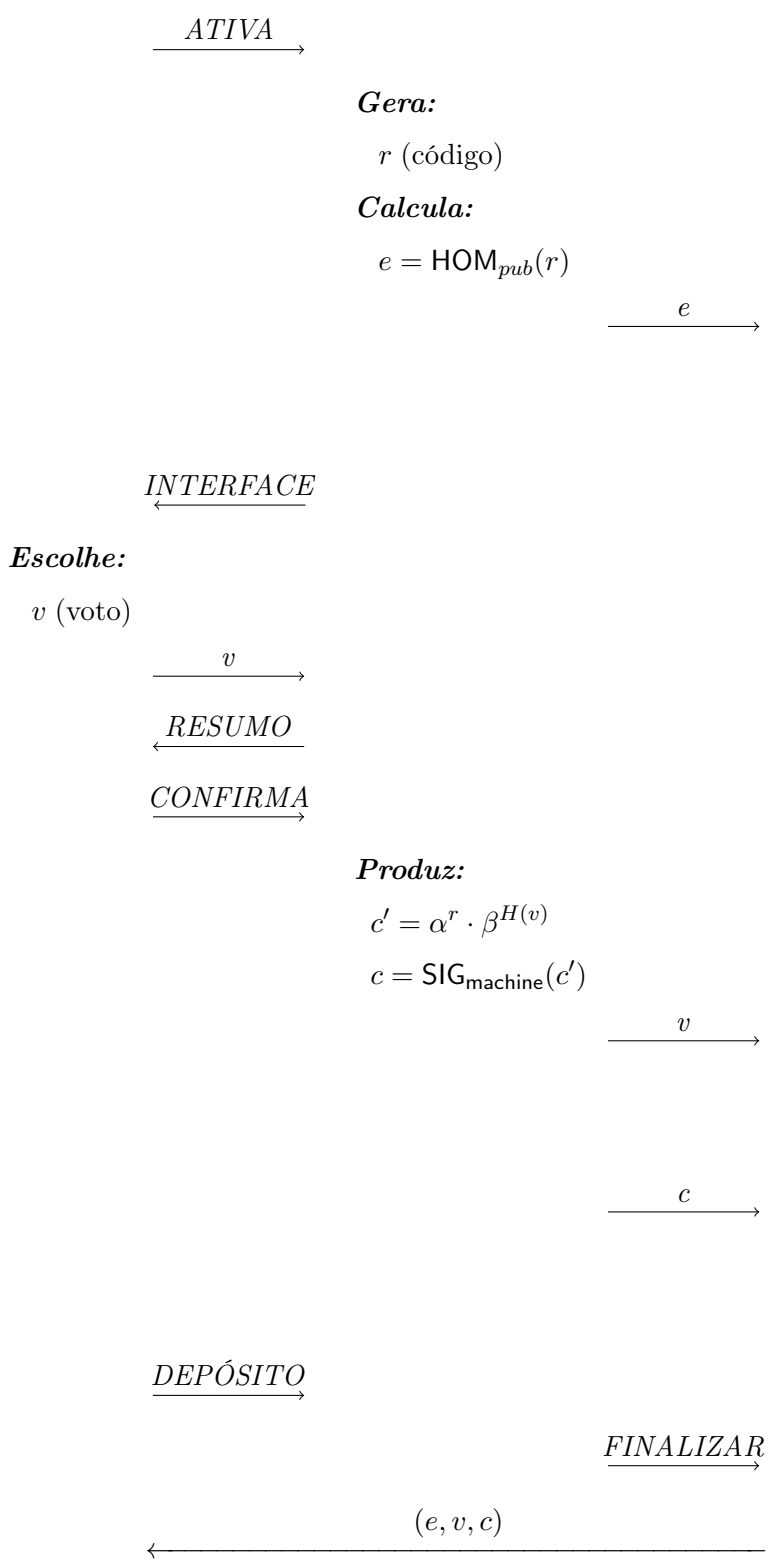
Não obstante a sensibilidade desta etapa ao contexto de aplicação do protocolo, considera-se como de crucial importância à devida lisura do processo que a metodologia de identificação, independente de tal contexto, não deve requerer qualquer tipo de comunicação direta com a máquina de votação em si. A existência de tal canal de comunicação, além de tecnicamente viabilizar violação do sigilo do voto, abre também margem à possibilidade de identificação pessoal do eleitor por parte da máquina que produz a cédula, o que permitiria a esta agir de forma particular ao eleitor que dela faz uso em determinado momento. Em um contexto de votação, tal capacidade poderia ser explorada no sentido de empreender tentativas de manipulação das escolhas do eleitor ou dificultar-lhe o voto de alguma forma.

4.2.3.2 Geração da Cédula

Tendo o eleitor passado pelo crivo necessário do processo de identificação, a ele é concedido acesso ao restante do fluxo eleitoral. Nesse momento, o eleitor é direcionado à máquina de votação em si, dispositivo este responsável pela geração, via uma impressora acoplada, de uma cédula eleitoral impressa em papel que, em última instância, será depositada em urnas para contabilização em uma totalização eleitoral. Assume-se aqui que a máquina de votação usada é dotada de uma tela, sendo operada pelo próprio eleitor por meio de alguma interface conveniente e suficientemente intuitiva, como um teclado numérico simples ou uma tela sensível ao toque.

O Diagrama 4.2.3 demonstra o fluxo de votação executado por um eleitor que não tenha interesse no passo opcional de desafio da cédula gerada (i.e. um eleitor que opta diretamente pelo depósito de seu voto):

Diagrama 4.2.3: Votação



Como demonstra o Diagrama 4.2.3, o eleitor utiliza-se da máquina de votação para gerar um conjunto de três parâmetros que, quando impressos, coletivamente correspon-

dem àquilo que vem a ser a sua *cédula* eleitoral. Este conceito de cédula é formalmente estabelecido na Definição 4.2.7:

Definição 4.2.7 (Cédula)**.** *Sejam $\mathbb{W}$ um grupo de ordem q, dois elementos $\alpha, \beta \in \mathbb{W}$ geradores independentes deste grupo e $H : V \to \mathcal{V}$ um hash de votos. Uma* ***cédula*** *é uma tripla (r, v, c) que satisfaz as seguintes propriedades:*

1. $r \xleftarrow{\$} \mathbb{Z}_q$
2. $v \in V$
3. $c = \alpha^r \cdot \beta^{H(v)}$

O parâmetro aleatório r é denominado ***código de segurança****, o parâmetro v é denominado* ***voto legível*** *do eleitor e o construto c é denominado* ***recibo****.*

Note-se assim que uma cédula, conforme definida no contexto do presente protocolo, apresenta parâmetros específicos que devem necessariamente respeitar uma relação matemática entre si. Embora a Definição 4.2.7 explicite que uma cédula necessariamente respeita a equação matemática exposta, é de trivial observação ser possível se agregar uma tripla (r, v, c) que não apresenta qualquer correlação segundo a supracitada equação, mas que passa uma suposta aparência de cédula. Isto posto, e para fins de ênfase discursiva, colocam-se as seguintes extrapolações terminológicas na forma das Definições 4.2.8 e 4.2.9, complementares entre si:

Definição 4.2.8 (Cédula Válida)**.** *Uma* ***Cédula Válida*** *é uma cédula cujos parâmetros (r, v, c) necessariamente respeitam a Definição 4.2.7.*

Definição 4.2.9 (Cédula Inválida)**.** *Uma* ***Cédula Inválida*** *é uma suposta cédula cujos parâmetros (r, v, c) são tais que $c \neq \alpha^v \cdot \beta^{H(v)}$.*

Assim sendo, em contextos no presente escrito nos quais a correlação matemática exposta na Definição 4.2.7 seja posta em xeque, a terminologia colocada nas Definições 4.2.8 e 4.2.9 será empregada para fins de diferenciação explícita. Por outro lado, em contextos nos quais seja assumida como válida, o uso puro do termo *cédula* necessariamente faz referência a uma tripla (r, v, c) que constitui uma *cédula válida*.

Voltando à Definição 4.2.7, o termo *segurança* associado ao parâmetro aleatório r refere-se às garantias de sigilo incondicional e integridade computacional do voto v do eleitor dado seu recibo c associado. A natureza de tais garantias é demonstrada pelo Teorema 4.2.10 seguinte:

Teorema 4.2.10. *Seja uma cédula formada por uma tripla (r, v, c) específica. Se o parâmetro r é uniformemente aleatório e desconhecido, então o voto v do eleitor, mediante unicamente observação do recibo c associado, apresenta sigilo incondicional e integridade computacional.*

Demonstração. As referidas garantias de sigilo e integridade providas pelo recibo c decorrem diretamente das características dos parâmetros eleitorais associados à produção da cédula em conjunto com a observação de que c corresponde a um compromisso de Pedersen em si. Neste, o código de segurança r é um valor aleatório utilizado para produzir um compromisso com o valor $H(v)$ correspondente à representação digital do voto v expresso na cédula. Isto posto, o *sigilo* do voto diz respeito à dificuldade de dedução de v dada observação do recibo c correspondente. Sob as hipóteses de aleatoriedade de r e seu total desconhecimento em âmbito extrínseco ao da máquina de votação que o gerou, o Corolário 3.4.3 garante que o sigilo do voto é incondicionalmente resguardado por c. Por sua vez, a *integridade* de v dado o recibo c diz respeito à dificuldade de se encontrar valores r', v', com $v' \neq v$, tais que $c' = \alpha^{r'} \cdot \beta^{H(v')} = c$. Note-se que fazê-lo com $r' = r$ exigiria encontrar uma colisão em H, o que, pela Definição 4.2.6, não é computacionalmente factível durante o período eleitoral. Por outro lado, encontrar um valor $r' \neq r$ que produza o mesmo recibo para um voto diferente implica em resolução do problema do logaritmo discreto, conforme exposto pelo Teorema 3.4.4, de modo que o Corolário 3.4.5 garante a integridade computacional do voto v. □

Com base no exposto pelo Teorema 4.2.10, note-se, entretanto, que a integridade computacional, resguardadas as devidas premissas, é garantida apenas durante o período eleitoral, sendo esperado que, decorrido tempo suficiente após a finalização do pleito, torne-se computacionalmente factível produzir recibos idênticos para votos distintos, seja pela exploração de colisões no hash de votos, seja pela capacidade de resolução do DLP no grupo utilizado. Portanto, o protocolo assume tacitamente que a produção de recibos colidentes findo o período eleitoral é de impacto limitado sobre o processo eleitoral e seus resultados dado o contexto de aplicação do protocolo.

Por outro lado, considerando-se o espaço de votos V como sendo suficientemente limitado de modo a possibilitar computação polinomial de $H(v)$ para todos os votos $v \in V$, a revelação de um código de segurança r e de seu recibo c correspondente permite com alta probabilidade revelar o voto comprometido por tal recibo. Este fato é exposto no Teorema 4.2.11 a seguir:

Teorema 4.2.11. *Sob a hipótese de que V é polinomialmente limitado, o conhecimento de r e c correspondentes viabiliza, a menos de uma probabilidade negligenciável, revelar v em tempo polinomial.*

Demonstração. Sob a supracitada hipótese de limitação do espaço de votos V, é computacionalmente plausível o percorrimento de V em tempo polinomial, conforme expõe o seguinte algoritmo:

Entrada: *par (r,c) correspondente*
Saída: *voto v*

for $v \in V$ **do**
  $c^* \leftarrow \alpha^r \cdot \beta^{H(v)}$
  if $c^* = c$ **then**
    return v
return *ERRO*

O algoritmo percorre todo o espaço de votos computando recibos para cada qual utilizando-se do valor r provido como entrada. A única possibilidade de o algoritmo retornar *ERRO* é no caso em que o par (r, c) provido como entrada não é correspondente. Do contrário, certamente um voto v tal que $\alpha^r \cdot \beta^{H(v)} = c$ será encontrado. Observe-se ainda que o voto v retornado apenas poderá estar incorreto em caso de colisão no hash de votos H, o que, por definição, apenas ocorre com probabilidade negligenciável.

□

O Teorema 4.2.11, portanto, explicita a importância de que o parâmetro r seja devidamente mantido em sigilo. Há de se notar ainda que esse processo de revelação dos votos é tornado ainda mais fácil diante do conhecimento de permutações aleatórias de todos os códigos de segurança, todos os votos e todos os recibos para uma mesma eleição, algo plausível de ocorrência diante de descuidos em contexto pós-eleitoral. Isto é exposto pelo Teorema 4.2.12:

Teorema 4.2.12. *Seja uma eleição $\mathscr{E}$ para a qual sejam conhecidas três permutações aleatórias: $\mathbb{R}$, referente a todos os códigos de segurança r utilizados; $\mathbb{V}$, referente a todos os votos v depositados; e $\mathbb{C}$, referente a todos os recibos c gerados. Assim, sendo V o espaço de votos, é possível produzir em tempo polinomial, a menos de uma probabilidade negligenciável, um mapeamento unívoco $\mathscr{M} : \mathbb{C} \to V$ entre recibos e os respectivos votos aos quais se referem.*

Demonstração. Assumindo-se que cada eleitor deposita uma única tripla (r, v, c) e que o número de eleitores é polinomialmente limitado, o percorrimento exaustivo de cada um dos elementos das referidas permutações pode ser feito em tempo polinomial. Assim sendo, considere-se o seguinte algoritmo, em que $\mathscr{M}$ é o mapeamento entre recibos e votos a ser produzido, α, β são os geradores independentes e H é o hash de votos:

$\mathscr{M} \leftarrow NULL$
for $r \in \mathbb{R}$ **do**
  for $v \in \mathbb{V}$ **do**
    $c \leftarrow \alpha^r \cdot \beta^{H(v)}$
    if $c \in \mathbb{C}$ **then**
      $\mathscr{M}[c] \leftarrow v$
      break
return $\mathscr{M}$

O algoritmo mostrado faz um percorrimento exaustivo pelos códigos de segurança, combinando cada qual destes com algum voto registrado e verificando se tal combinação resulta em algum dos recibos. Sob a hipótese da devida correspondência entre as três permutações, certamente existirá um par $(r, v) \in \mathbb{R} \times \mathbb{V}$ que produzirá um recibo $c \in \mathbb{C}$ correspondente, o qual é relacionado com seu respectivo voto em um mapeamento $\mathscr{M}$ retornado ao fim da execução. Este mapeamento apenas poderá estar errado ou incompleto em dois casos:

- Por colisão no hash de votos H, o que por definição apenas ocorre com probabilidade negligenciável.
- Por colisão de recibos em $\mathbb{C}$, o que pelo Teorema 3.4.9 também apenas ocorre com probabilidade negligenciável para grupos de ordem suficientemente grande.

Portanto, o mapeamento $\mathscr{M} : \mathbb{C} \to \mathbb{V}$ assim produzido, a menos de uma probabilidade negligenciável, será correto.

□

Os Teoremas 4.2.11 e 4.2.12 explicitam a importância e a necessidade de que os códigos de segurança r produzidos no contexto de uma eleição sejam mantidos desconhecidos, a saber, no intuito de proteger o sigilo dos votos individuais dos eleitores. Por tal motivo, e conforme mostra o Diagrama 4.2.3, os códigos de segurança r são individualmente cifrados com a chave pública dos guardiões, produzindo uma versão cifrada e que é então impressa junto aos demais parâmetros. Dessa forma evita-se que qualquer uma das partes envolvidas viole unilateralmente o sigilo dos votos, o que se tornaria uma possibilidade caso r fosse diretamente exposto na cédula impressa. O motivo para o requisito de que tal cifração seja homomórfica deve-se à necessidade de conhecimento do somatório dos códigos de segurança, como ficará claro adiante na discussão da etapa de *auditoria* da eleição. Como se há de ver, em sendo mantidos os códigos de segurança e seus respectivos registros individuais desconhecidos mesmo após a finalização da eleição, o Teorema 4.2.10 mostra que o sigilo dos votos individuais é matematicamente garantido como eterno.

Em termos dos geradores selecionados para o grupo $\mathbb{W}$ e tendo em vista seu uso na geração do recibo de uma cédula, fica agora clara a motivação para a exigência de que estes sejam independentes entre si, conforme mostra o Teorema 4.2.13:

Teorema 4.2.13. *Sejam $\mathbb{W}$ o grupo algébrico empregado no protocolo e $\alpha, \beta \in \mathbb{W}$ os geradores definidos para uma eleição $\mathscr{E}$. Se α e β são geradores dependentes, então é possível produzir, com probabilidade não negligenciável, recibos que sejam verificáveis como válidos para dois votos distintos durante o período eleitoral.*

Demonstração. Seja r o código de segurança da cédula de um eleitor cujo voto é representado por v. Por definição, o recibo c a ser gerado para compor esta cédula é dado

por $c = \alpha^r \cdot \beta^{H(v)}$. Se α e β são geradores dependentes, então, pela Definição 4.2.3, $\log_\alpha \beta$ pode ser calculado com probabilidade não negligenciável durante o período eleitoral. Assim, sendo $\log_\alpha \beta$ conhecido, o Teorema 3.4.7 provê um método direto para o cálculo de um compromisso $c' = c$ tal que $c' = \alpha^{r'} \cdot \beta^{H(v')}$ consiste em um recibo para um voto $v' \neq v$ que pode trivialmente ser mostrado como válido para r e v por conta da colisão. Para isto, conforme especifica o Teorema 3.4.7, basta gerar um novo valor r' tal que $r' = [\log_\alpha \beta \cdot (H(v) - H(v'))] + r$, em que todas as grandezas à direita da equação são conhecidas ou computáveis. □

O Teorema 4.2.13 mostra que a independência dos geradores escolhidos para uma eleição é condição necessária à garantia de integridade dos votos individuais dos eleitores. Por outro lado, o Teorema 3.4.4 mostra que, se dois recibos podem ser verificáveis como válidos para dois votos distintos, então isto implica que os geradores usados são dependentes, haja vista ser possível se computar trivialmente o logaritmo discreto entre ambos. Portanto, conclui-se que a independência dos geradores escolhidos junto à hipótese de intratabilidade do DLP no grupo algébrico considerado garantem a incapacidade polinomial de geração de recibos que possam ser validados para votos distintos durante o período eleitoral.

Voltando ao Diagrama 4.2.3, observe-se que o processo de geração da cédula se dá mediante a seguinte sequência de passos executada pelo eleitor em conjunto com a máquina de votação e sua impressora acoplada:

1. O eleitor ativa a máquina de votação utilizando-se de algum mecanismo que seja **incapaz** de pessoalmente identificá-lo ao dispositivo.

2. Imediatamente após a ativação e antes que o eleitor possa interagir com a máquina, esta gera um código de segurança $r \xleftarrow{\$} Z_q$.

3. O código r é então cifrado com a chave pública *pub* gerada durante a etapa de preparação eleitoral, sendo produzida uma versão cifrada $e = \mathsf{HOM}_{pub}(r)$ do código r.

4. O valor e é então impresso pela impressora acoplada à máquina de votação em meio físico (e.g. papel térmico) e visível ao eleitor. A impressão de e se dá em um formato prontamente identificável por máquinas (e.g. código QR) a fim de mitigar possibilidades de erros interpretativos quando de sua leitura.

5. Finda a impressão do código de segurança cifrado e, e apenas após esta, a máquina de votação disponibiliza sua interface para que o eleitor possa fazer suas escolhas, as quais são coletivamente referidas por v.

6. Manifestas as escolhas pelo eleitor, um resumo destas é mostrado em tela, sendo oferecida ao eleitor a oportunidade de descartá-las e repetir o processo de escolha ou confirmá-las e prosseguir no fluxo de votação.

7. Após confirmação, a máquina produz um recibo inicial $c' = \alpha^r \cdot \beta^{H(v)}$, o qual é então assinado com uma chave de assinatura associada à máquina de votação para produzir o recibo assinado $c = SIG_{\mathsf{machine}}(c')$.

8. A máquina então imprime v em formatação trivialmente legível e verificável pelo eleitor, a qual também permite leitura via uso de dispositivos OCR (de *Optical Character Recognition*) para reconhecimento de caracteres impressos.

9. Finalmente, a máquina de votação imprime também o recibo assinado c em um formato legível por máquinas similar ao utilizado para impressão de e.

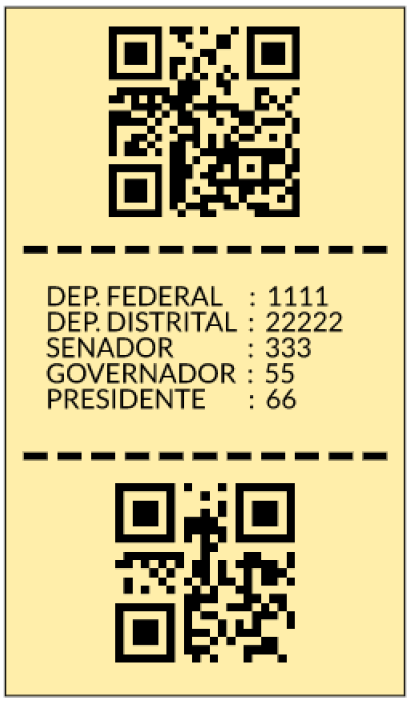
Note-se assim que a versão impressa da cédula é em realidade composta pela tripla (e, v, c), em que e corresponde a uma versão cifrada de r, v é o voto legível e c um recibo correspondente a um compromisso de Pedersen que liga matematicamente os três parâmetros. Isto, entretanto, não invalida a Definição 4.2.7: em sua essência, a cédula ainda corresponde à tripla (r, v, c), a cifração de r para produzir e apenas servindo ao propósito de ocultar r pelos motivos dispostos nos Teoremas 4.2.11 e 4.2.12. Isto posto, no decorrer do texto a tripla (e, v, c) será também denominada *cédula*.

Ademais, é importante ressaltar que o código de segurança r deve ser produzido e impresso como e antes que o eleitor expresse qualquer uma de suas opções junto à máquina. Isto ocorre para prevenir a possibilidade de que a máquina trate o parâmetro r como $r = f(v), f : \mathscr{V} \to \mathbb{Z}_q$, o que efetivamente viabilizaria o uso do código de segurança como um canal lateral de informação. Isto poderia ser utilizado, por exemplo, no sentido de violação do sigilo do voto ao especificar que para votos em uma opção X a máquina de votação produziria códigos de segurança (ou cifrações destes) que fossem terminados por uma sequência específica de bits.

Assim, a exigência operacional de que a máquina de votação deve se comprometer com um r aleatório antes que qualquer opção seja informada pelo eleitor impede que o sistema faça uso da composição de tal valor para viabilizar vazamento de informação referente ao voto (e.g. um valor r terminado em bits 10 indica voto em um candidato A). Ademais, estando devidamente definido o espaço digital dos votos e o hash eleitoral a ser empregado, bem como os demais parâmetros eleitorais públicos associados à geração do recibo c, não resta qualquer grau de liberdade na geração da cédula.

Findo o processo descrito de operação da máquina de votação, o eleitor deve ter em mãos uma cédula impressa contendo a tripla (e, v, c). Um exemplo de cédula impressa é mostrado na Figura 4.2:

Figura 4.2: Uma cédula impressa



Fonte: Elaborado pelo autor.

A Figura 4.2 mostra a clara divisão visual entre os três componentes da cédula impressa. Observe-se ainda que o voto v encontra-se claramente legível pelo eleitor. Com uma tal cédula em mãos, o eleitor então avança à etapa de depósito desta, direcionando-se à área reservada às urnas físicas. Antes de proceder à descrição de tal etapa, explora-se na subseção seguinte a possibilidade de desafio opcional de uma cédula gerada.

4.2.3.3 Desafio (Opcional)

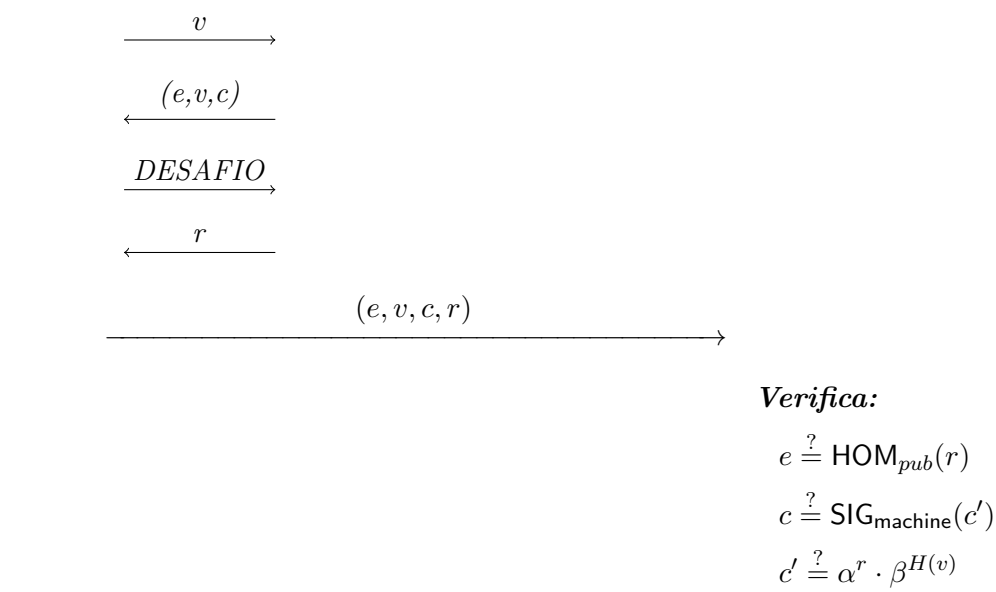
Durante a votação, um eleitor pode se questionar a respeito da correta formação de sua cédula. Afinal, embora tal eleitor possa verificar que seu voto legível impresso de fato corresponde a suas opções conforme informadas à máquina, nada lhe garante que os parâmetros r e c de fato são correspondentes com o voto v em suas mãos. Isto é, o eleitor pode questionar se a cédula a ele fornecida de fato constitui uma cédula válida. Caso r fosse explícito, o eleitor poderia fazer tal verificação matematicamente, mas os Teoremas 4.2.11 e 4.2.12 demonstram os problemas associados a um r revelado, de modo que o sistema apenas provê acesso à cifração e pelo eleitor, o que lhe impossibilita tal verificação de conformidade matemática da cédula por conta própria.

Tal possível desconfiança por parte de um eleitor, portanto, faz sentido, haja vista

que o desconhecimento de r possivelmente oferece à máquina um grau de liberdade que abre margem à adulteração do voto mantendo-se seu respectivo recibo plenamente válido perante esforços de auditoria em caráter pós-eleitoral. Uma metodologia de exploração de tal possibilidade de adulteração se utiliza desse grau de liberdade para produção de recibos inválidos, e é demonstrada na Subseção 4.4.3.2. A forma de mitigá-la passa pelo provimento de um mecanismo de auditoria *in loco* por parte do eleitor que lhe é oferecido como opção de livre escolha após o sistema ter gerado uma *suposta* cédula (e, v, c). Tal mecanismo de auditoria, aqui denominado *desafio*, força o sistema a comprovar que as cédulas por ele geradas encontram-se em conformidade com a equação matemática que lhes rege a forma, exposta na Definição 4.2.7.

Assim, conforme mencionado, o fluxo previamente mostrado no Diagrama 4.2.3 reflete a operação da máquina por parte de um eleitor que queira produzir um voto para depósito. Entretanto, e caso desejado, após geração de sua cédula impressa completa, o eleitor pode manifestar à máquina de votação sua intenção de desafiar o sistema a mostrar que a cédula produzida é válida. O processo é mostrado pelo Diagrama 4.2.4 exposto a seguir, em que as duas primeiras interações entre o eleitor e a máquina resumem todo o exposto no Diagrama 4.2.3 para geração e impressão da tripla (e, v, c) e previamente à manifestação da opção de depósito pelo eleitor:

Diagrama 4.2.4: Desafio



Note-se que o eleitor segue todo o fluxo de votação normalmente até a produção de sua cédula (e, v, c), quando então manifesta sua intenção em desafiar a cédula produzida. Ao manifestar tal intenção, a máquina deve então imprimir em texto pleno o código de segurança r utilizado na cédula, gerando uma tupla final (e, v, c, r), em que e é o código de segurança cifrado, v é o voto em texto pleno do eleitor, c é o recibo referente a um

compromisso de Pedersen assinado e r é o código de segurança em texto pleno. O eleitor então leva tais parâmetros a uma segunda máquina independente daquela de votação e dedicada à execução de procedimentos de desafio, a qual é operada junto a algum representante das autoridades eleitorais no intuito de verificar a correspondência ou não dos parâmetros da cédula, além da devida assinatura oficial da mesma. Após a verificação do desafio, a cédula desafiada deve ser publicada oficialmente pelas autoridades eleitorais, servindo como uma testemunha da devida (ou indevida) operação da máquina.

Observe-se pelo Diagrama 4.2.4 que tanto o recibo assinado c quanto o compromisso de Pedersen puro c' são verificados: o primeiro com relação à validade da assinatura e o segundo com relação à correta formação do compromisso em si. Ademais, ressalta-se que, por simplicidade, o diagrama não explicita o provável caráter probabilístico da cifra HOM utilizada. No entanto, em sendo o esquema de cifração usado probabilístico, faz-se necessária a impressão adicional do parâmetro probabilístico utilizado de modo a se atestar a devida produção do código cifrado e.

Cabe ressaltar que, tendo optado pelo desafio, o eleitor passa a ser capaz de comprovar seu voto, isto por conta da necessária revelação do código de segurança r no processo. Assim, tendo conhecimento do par (r, c), um eleitor poderia revelar, por desejo próprio ou sob coerção, seu voto v, o qual poderia ser então matematicamente verificado como correspondente a tal par utilizando-se da equação de composição da cédula. Por este motivo, é crucial que toda cédula desafiada seja considerada como imprópria para depósito, devendo ao eleitor ser dada a oportunidade de gerar uma nova cédula.

Observe-se, portanto, que o eleitor pode gerar um voto para desafio contendo qualquer seleção de opções que bem deseje, visto que lhe será assegurada a possibilidade de gerar um novo voto posteriormente. Isto é, um eleitor que pretenda desafiar o sistema pode produzir um voto em qualquer combinação de opções, não necessariamente naquelas de sua escolha. Dessa forma, o eleitor não precisa se constranger com relação à presença de um fiscal eleitoral junto ao processo de desafio: não há qualquer garantia de que o voto desafiado ali expresso corresponde ao voto que o eleitor de fato pretende depositar.

A importância da manifestação da intenção do desafio apenas após a produção da cédula deve-se à necessidade de que a máquina de votação seja incapaz de discernir se o voto em produção será desafiado ou não, sendo apenas informada de tal intenção após já ter se comprometido com uma tripla (e, v, c). Do contrário, isto é, caso a máquina saiba das intenções do eleitor anteriormente à total impressão da cédula, ela consegue produzir uma cédula inválida imune ao desafio, conforme demonstrado no Teorema 4.2.14 a seguir:

Teorema 4.2.14. *Se uma máquina de votação é capaz de discernir entre a produção de uma cédula para depósito e de uma cédula a ser desafiada antes da total impressão da tripla (e, v, c), então esta máquina é capaz de produzir cédulas inválidas imunes ao desafio.*

Demonstração. Basta observar que, sendo o último parâmetro impresso correspondente

ao recibo c, uma máquina que saiba da intenção do eleitor antes de se comprometer com um recibo consegue se utilizar do valor r já produzido para se comprometer com qualquer voto v', seja ele o voto v do eleitor ou um voto completamente diferente. O algoritmo seguinte demonstra como uma máquina desonesta que conheça previamente a intenção do eleitor de desafiar ou não a cédula gerada consegue sempre produzir o recibo desejado para a situação correta:

// Gera, cifra e imprime código de segurança
$r \xleftarrow{\$} Z_q$
$e \leftarrow \mathsf{HOM}_{\mathsf{pub}}(r)$
$\mathsf{PRINT}(e)$
// Recebe o voto pretendido pelo eleitor
$v \leftarrow \mathsf{READ\_VOTE}()$
if DESAFIO **then**
  $c = \alpha^r \cdot \beta^{H(v)}$
  $\mathsf{PRINT}(v)$
  $\mathsf{PRINT}(c)$
else
  $c' = \alpha^r \cdot \beta^{H(v')}$
  $\mathsf{PRINT}(v)$
  $\mathsf{PRINT}(c')$

Note-se, portanto, que a máquina, sabendo da intenção do desafio, produz uma cédula válida para o eleitor, de modo a garantir que o processo de desafio será realizado com sucesso. Por outro lado, ao saber que a cédula será depositada e não passará pelo processo de desafio, a máquina produz uma cédula inválida. □

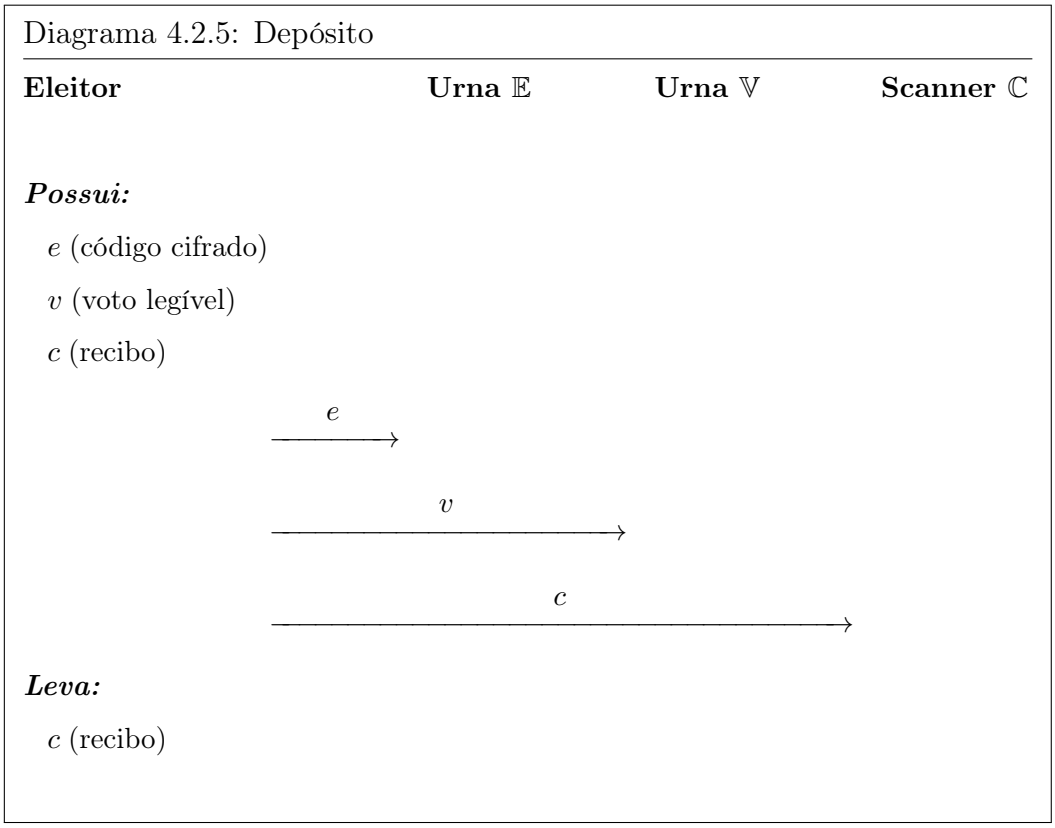
A possibilidade de produção sistemática de cédulas inválidas possibilita a construção de um ataque de adulteração eleitoral, descrito em maiores detalhes na Subseção 4.4.3.2, no qual os votos dos eleitores são substituídos por votos inconformes, mas em acordo com as cédulas inválidas produzidas durante a votação. A mesma subseção, entretanto, argumenta que é suficiente que um pequeno número de eleitores ou auditores solicite o desafio da cédula de modo a se obter uma ampla margem de confiança nos resultados.

Um último aspecto importante a ser mencionado é o de que a cifração homomórfica HOM pode ser probabilística, de modo que um valor aleatório u pode estar envolvido na produção do código cifrado e. Em tal contexto, de modo a se verificar a igualdade entre o valor e informado e a cifração $\mathsf{HOM}_{pub}(r)$, a máquina deve, adicionalmente a r, imprimir também a aleatoriedade u usada para viabilizar o desafio conforme exposto no Diagrama 4.2.4.

4.2.3.4 Depósito

Quando satisfeito com suas escolhas e com os desafios solicitados às cédulas, o eleitor finalmente informa à máquina de votação, após esta imprimir completamente uma nova tripla (e, v, c), que irá depositar a cédula gerada no processo, obtendo-a então da impressora conforme já mostrado no Diagrama 4.2.3. Com sua cédula para depósito em mãos, o eleitor então se dirige às urnas. O Diagrama 4.2.5 seguinte mostra o processo de depósito de um voto por um eleitor:

Diagrama 4.2.5: Depósito



Conforme expõe o Diagrama 4.2.5, o eleitor deve depositar o código de segurança cifrado e em uma urna $\mathbb{E}$ de códigos, e adicionalmente depositar seu voto em texto pleno v em uma urna $\mathbb{V}$ de votos. A totalização, como se há de ver, será realizada por meio da contagem dos votos presentes em $\mathbb{V}$. Ao final, o recibo c do eleitor, devidamente assinado pelas autoridades eleitorais, é escaneado e o eleitor o leva consigo.

Observe-se, portanto, que a única informação que o eleitor leva para o meio externo ao local de votação é o recibo c. Em particular, seu voto v em texto pleno fica depositado em uma urna física local, jamais operando como uma espécie de comprovante a terceiros da forma como o eleitor votou. O recibo c, por sua vez, em se tratando de um compromisso de Pedersen, é válido para qualquer possível voto gerável para a eleição em questão, de modo

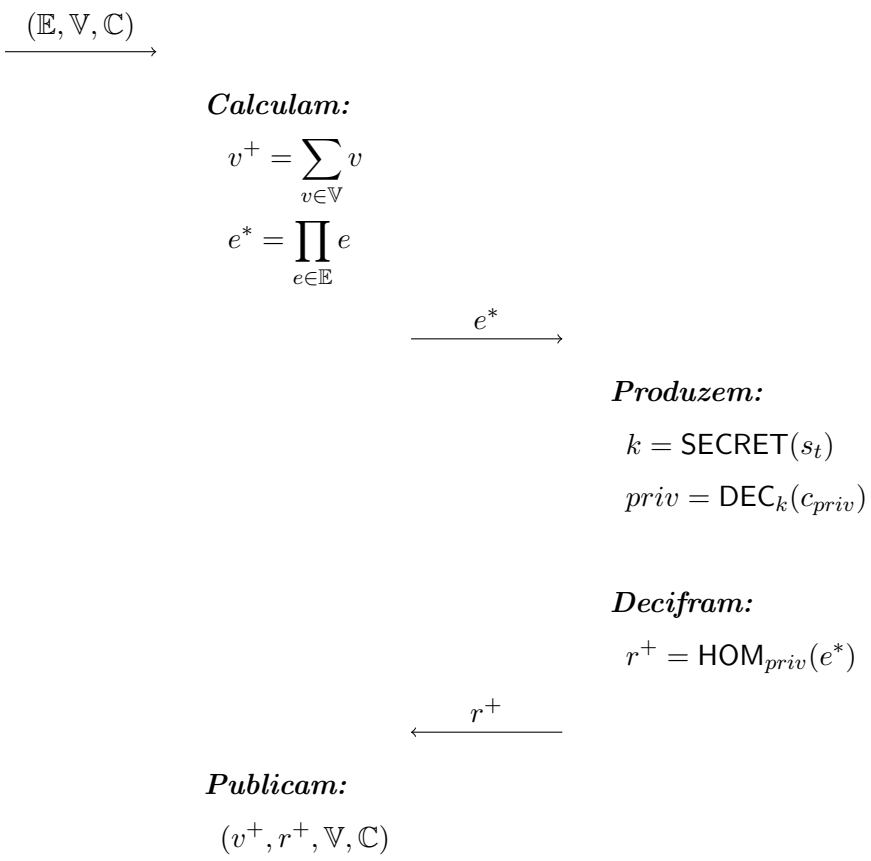
que c isoladamente é incapaz de comprovar como o eleitor votou, como já demonstrado pelo Teorema 4.2.10.

Um último aspecto importante de ser ressaltado é que, embora os recibos c dos eleitores não sejam depositados em urnas propriamente ditas, é crucial, como se haverá de ver, que haja um registro oficial de cada recibo individual. Tal registro de recibos é aqui denotado como uma espécie de *urna virtual*, representada por $\mathbb{C}$, na qual cada recibo é "depositado" via processo de escaneamento.

4.2.4 Totalização

Finda a votação, prossegue-se então à etapa de *totalização* da eleição, em que os votos depositados pelos eleitores são agregados em um resultado eleitoral correspondente à totalização dos votos depositados. O resultado eleitoral obtido é então publicado juntamente a parâmetros adicionais a este associados, conforme mostrado no Diagrama 4.2.6 seguinte, em que os componentes da tripla $(\mathbb{E}, \mathbb{V}, \mathbb{C})$ denotam respectivamente os conjuntos de códigos cifrados, votos em texto pleno e recibos eleitorais de votação:

Diagrama 4.2.6: Totalização



Conforme demonstra o Diagrama 4.2.6, o processo de totalização se inicia com a abertura oficial das urnas físicas $\mathbb{E}$ e $\mathbb{V}$, contendo respectivamente os códigos de segurança cifrados e os voto individuais de cada eleitor. Como os códigos de segurança encontram-se cifrados, essa abertura das urnas não viabiliza a produção de um mapeamento $\mathbb{C} \to V$ como demonstrado no Teorema 4.2.12. Pelo contrário, sem conhecimento dos valores r ocultos pelas cifrações em $\mathbb{E}$, note-se que $\forall v \in \mathbb{V} : \forall c \in \mathbb{C} : \exists r \in \mathbb{Z}_q : c = \alpha^r \cdot \beta^{H(v)}$.

De posse dos votos individuais presentes em $\mathbb{V}$, as autoridades, com o auxílio de dispositivos de reconhecimento de caracteres, escaneiam cada qual dos votos, produzindo uma totalização coletivamente denotada por v^+. Com relação aos códigos de segurança cifrados contidos em R, por sua vez, as autoridades produzem um produtório e^* com tais valores, produtório este que, dada a necessária natureza homomórfica aditiva da cifra HOM, corresponde a um criptograma que cifra o somatório r^+ de todos os códigos de segurança r usados na geração dos recibos.

Isto feito, o criptograma e^* é então enviado aos guardiões eleitorais, que reúnem um conjunto s_t de parcelas da chave k compartilhada quando da cerimônia de preparação tal que $|s_t| \geq t$. De posse de tal conjunto, os guardiões conseguem recuperar a chave k e com ela produzir $priv = \mathsf{DEC}_k(c_{priv})$, obtendo assim a chave privada de cifração dos códigos de segurança. Com tal chave, os guardiões decifram o valor e^*, obtendo assim o

somatório r^+ de todos os códigos de segurança usados. Tal valor r^+ é então repassado às autoridades eleitorais como resultado da cooperação dos guardiões eleitorais.

Note-se aqui que esta é uma etapa bastante sensível do protocolo, haja vista que a chave de cifração dos códigos de segurança é desvelada em texto pleno. A principal motivação para uso dessa metodologia é justamente aliviar os requisitos sobre os guardiões individuais, os quais apenas precisam guardar em segurança suas respectivas parcelas. Do contrário, em se optando por algo como um esquema de cifração de limiar, os guardiões em contrapartida seriam exigidos também quanto à geração de pares de chaves criptográficas com parâmetros devidos e seu respectivo armazenamento em conformidade com padrões suficientemente robustos de segurança. Em contextos nos quais considere-se realista esperar dos guardiões o necessário conhecimento criptográfico e de aspectos de segurança digital, o protocolo pode ser adaptado de modo a se mitigar a sensibilidade desta etapa.

Finalmente, as autoridades publicam, em meio oficial e acessível, os resultados eleitorais v^+ por ela obtidos e o somatório de códigos $r+$ fornecido pelos guardiões eleitorais. Juntamente, as autoridades publicam ainda o conjunto $\mathbb{V}$ de todos os votos contabilizados para a eleição e o conjunto $\mathbb{C}$ de todos os recibos escaneados. Observe-se que a publicação dos resultados eleitorais v^+ não é estritamente necessária visto tratar-se de mero resumo intrínseco e completamente determinado pelo conjunto $\mathbb{V}$ de votos. Não obstante, a publicação de v^+ facilita a visualização dos resultados por parte do público.

Ressalta-se que o conjunto $\mathbb{E}$ de códigos cifrados não deve ser publicado. Fazê-lo implicaria na possibilidade de sua perpetuação e armazenamento até um futuro em que a criptoanálise ou poder computacional bruto pudesse viabilizar obtenção da chave *priv* a partir da chave *pub*, permitindo assim obtenção dos códigos r individuais. Isto, juntamente aos conjuntos $\mathbb{V}$ e $\mathbb{C}$ oficialmente publicados, permitiria a construção de um mapeamento unívoco $\mathbb{C} \to V$ nos moldes do demonstrado no Teorema 4.2.12. Tal conjunto $\mathbb{E}$, portanto, deve ser devidamente destruído assim que findo o período eleitoral.

Neste ponto, os resultados da eleição já se encontram disponíveis e acessíveis ao público sem qualquer necessidade de mobilização adicional por parte dos eleitores. Àqueles que desejem uma maior garantia acerca dos resultados publicados, o protocolo viabiliza metodologias de auditoria que permitem verificar tanto se um voto individual foi devidamente contabilizado quanto se os resultados oficiais de fato correspondem aos parâmetros publicados pelas autoridades eleitorais. Tal processo é descrito em maiores detalhes na subseção seguinte.

4.2.5 Auditoria

O processo de auditoria compreende tanto a *auditoria individual* por parte do eleitor quanto a *auditoria universal* por parte de qualquer indivíduo ou observador interessado em avaliar a eleição. Cada qual dessas possibilidades de auditoria é discutida individualmente nas subseções seguintes.

4.2.5.1 Auditoria Individual

A auditoria individual exige conhecimento de um recibo válido e é empreendida por um eleitor que detenha tal recibo e deseje verificar se seu voto foi devidamente contabilizado na totalização, conforme mostrado no Diagrama 4.2.7:

Diagrama 4.2.7: Auditoria (Eleitor)

Eleitor	**Resultados**
Possui:	
c (recibo)	
	$\mathbb{C} \longleftarrow$
Verifica:	
$c \stackrel{?}{\in} \mathbb{C}$	

Note-se assim que o processo de auditoria individual é consideravelmente simples. De posse de seu recibo eleitoral c, o eleitor acessa uma interface oficial contendo o registro de todos os recibos escaneados na eleição e assegura-se de que seu recibo se encontra armazenado nesse registro. Dado o compromisso das autoridades, por meio de assinatura digital, para com cada recibo produzido, um eleitor que não encontre seu recibo na listagem oficial detém comprovação, dada pelo próprio recibo, de que seu voto foi indevidamente removido dos resultados eleitorais oficiais conforme publicados.

4.2.5.2 Auditoria Universal

A auditoria universal, por sua vez, pode ser empreendida por qualquer indivíduo interessado e exige apenas conhecimento dos parâmetros públicos e resultados associados à eleição, sem necessidade de acesso especial ou restrito a equipamentos ou dados. O processo é descrito no Diagrama 4.2.8 seguinte:

Diagrama 4.2.8: Auditoria (Universal)

Auditor **Resultados**

$\xleftarrow{v^+}$

$\xleftarrow{r^+}$

$\xleftarrow{\mathbb{V}}$

$\xleftarrow{\mathbb{C}}$

Calcula:

$$v^{+\prime} = \sum_{v \in \mathbb{V}} v$$

$$h^+ = \sum_{v \in \mathbb{V}} H(v)$$

$$c^* = \prod_{c \in \mathbb{C}} c$$

Verifica:

$$|\mathbb{V}| \stackrel{?}{=} |\mathbb{C}|$$

$$v^{+\prime} \stackrel{?}{=} v^+$$

$$c^* \stackrel{?}{=} \alpha^{r^+} \cdot \beta^{h^+}$$

De forma semelhante ao eleitor individual, o auditor interessado também obtém os parâmetros públicos e os resultados associados à eleição, todos devidamente acessíveis em caráter oficial como já mencionado. De posse de tais valores, o auditor verifica se a cardinalidade dos conjuntos $\mathbb{V}$ e $\mathbb{C}$ é idêntica. Tal restrição deve necessariamente ser verdadeira em uma eleição íntegra, haja vista que todo eleitor deve depositar uma tripla completa. Pelo mesmo motivo, tal cardinalidade deve também ser idêntica para o conjunto $\mathbb{E}$ de códigos cifrados, mas dado que este não é publicado, tal verificação por parte de um observador externo torna-se inviável.

Como outra verificação básica de integridade, com base no conjunto $\mathbb{V}$ obtido referente aos votos depositados para a eleição, o auditor computa uma totalização pessoal

$v^{+\prime}$ dos resultados eleitorais e verifica se esta de fato corresponde à totalização informada pelas autoridades.

Para além da supracitada recontagem dos votos, o auditor também recupera o valor r^+ publicado pelas autoridades e produz um somatório h^+ de todas as representações binárias $H(v)$ dos votos individuais $v \in \mathbb{V}$. Com base nesses valores r^+ e h^+, o auditor produz o chamado *recibo de integridade eleitoral*, conforme a Definição 4.2.15 seguinte:

Definição 4.2.15 (Recibo de Integridade Eleitoral)**.** *Seja $\mathscr{E}$ uma eleição com geradores independentes dados por α e β. Sejam ainda r^+ o somatório dos códigos de segurança em texto pleno conforme publicação e h^+ o somatório de todas as representações binárias $H(v)$ de votos individuais $v \in \mathbb{V}$. O* ***recibo de integridade eleitoral*** *de $\mathscr{E}$ corresponde a um compromisso de Pedersen produzido com os geradores independentes e os respectivos expoentes r^+ e h^+. Matematicamente, o recibo de integridade eleitoral corresponde ao valor C dado por:*

$$C = \alpha^{r^+} \cdot \beta^{h^+}$$

Por sua vez, com o conjunto $\mathbb{C}$ de recibos escaneados, o auditor computa o produtório c^* de todos os recibos associados à eleição. A importância do recibo de integridade eleitoral se torna clara quando o mesmo é analisado em conjunção com o produtório c^*, conforme expõe o Teorema 4.2.16 seguinte:

Teorema 4.2.16. *Seja $\mathscr{E}$ uma eleição e seja a tupla $(r^+, \mathbb{V}, \mathbb{C})$ correspondente a seus resultados íntegros. Então, denotando-se por C o recibo de integridade de $\mathscr{E}$ e por c^* o produtório de todos os recibos $c \in \mathbb{C}$, necessariamente tem-se que $C = c^*$.*

Demonstração. Seja denotado por $\mathbb{R}$ o conjunto íntegro de todos os códigos de segurança usados para os recibos válidos da eleição, isto é, os mesmos que se encontram expressos de forma cifrada em $\mathbb{E}$. Sob a hipótese de integridade da tupla $(r^+, \mathbb{V}, \mathbb{C})$ publicada, tem-se que $|\mathbb{R}| = |\mathbb{V}| = |\mathbb{C}| = S$, em que S denota a cardinalidade coletiva dos respectivos conjuntos. Ademais, é possível se construir respectivas permutações indexadas $\mathbb{R}'$, $\mathbb{V}'$ e $\mathbb{C}'$ de tais conjuntos de modo que $c_i = \alpha^{r_i} \cdot \beta^{H(v_i)}, 1 \leq i \leq S$, em que $r_i \in \mathbb{R}'$, $v_i \in \mathbb{V}'$ e $c_i \in \mathbb{C}'$. Assim sendo, seja considerado o produtório dos recibos em $\mathbb{C}'$, dado por:

$$\begin{aligned}
\prod_{i=1}^{S} c_i &= c_1 \cdot c_2 \cdot \ldots \cdot c_S \\
&= \alpha^{r_1} \cdot \beta^{H(v_1)} \cdot \alpha^{r_2} \cdot \beta^{H(v_2)} \cdot \ldots \cdot \alpha^{r_S} \cdot \beta^{H(v_S)} \\
&= \alpha^{r_1 + r_2 + \ldots + r_S} \cdot \beta^{H(v_1) + H(v_2) + \ldots + H(v_S)} \\
&= \alpha^{\sum_{i=1}^{S} r_i} \cdot \beta^{\sum_{i=1}^{S} H(v_i)}
\end{aligned}$$

Dado que $\mathbb{R}'$ e $\mathbb{V}'$ apenas consistem em permutações de $\mathbb{R}$ e $\mathbb{V}$, tem-se que tanto o somatório de códigos de segurança, dado pelo expoente $\sum_{i=1}^{S} r_i$, quanto o somatório de hashes de votos, expresso no expoente$\sum_{i=1}^{S} H(v_i)$, são insensíveis às permutações específicas consideradas. Dessa forma, denotando-se por r^+ o somatório de códigos em $\mathbb{R}$ e por h^+ o somatório dos hashes de votos aplicados aos votos em $\mathbb{V}$, a condição de integridade dos resultados publicados garante que, independente das permutações $\mathbb{R}'$ e $\mathbb{V}'$ consideradas, tem-se:

$$\sum_{i=1}^{S} r_i = r^+$$
$$\sum_{i=1}^{S} H(v_i) = h^+$$

Logo, sob a hipótese de integridade dos resultados publicados e sendo C o recibo de integridade eleitoral conforme a Definição 4.2.15, o produtório c^* necessariamente satisfaz:

$$c^* = c_1 \cdot c_2 \cdot ... \cdot c_S = \alpha^{\sum_{i=1}^{S} r_i} \cdot \beta^{\sum_{i=1}^{S} H(v_i)} = \alpha^{r^+} \cdot \beta^{h^+} = C$$

□

Assim, o auditor verifica se o produtório c^* dos recibos constantes do conjunto $\mathbb{C}$ publicado é igual ao recibo de integridade eleitoral dado por $C = \alpha^{r^+} \cdot \beta^{h^+}$, igualdade esta que, conforme mostra o Teorema 4.2.16, deve sempre se mostrar como verdadeira para uma eleição íntegra diante da publicação de resultados íntegros. Assim, com base no recibo de integridade eleitoral e no produtório de recibos, formaliza-se uma *equação de auditoria universal* que deve se mostrar válida para toda eleição íntegra. O conceito é formalizado na Definição 4.2.17 seguinte:

Definição 4.2.17 (Equação de Auditoria Universal)**.** *Seja C o recibo de integridade eleitoral de uma eleição $\mathscr{E}$ e seja c^* o produtório de todos os recibos $c \in \mathbb{C}$ dessa eleição. A* ***equação de auditoria universal*** *é dada pela igualdade:*

$$C = c^*$$

Uma observação importante acerca da etapa de totalização é que esta confere às autoridades eleitorais total controle sobre os parâmetros $\mathbb{V}$ e r^+. Conforme explicita o Diagrama 4.2.6, o conjunto $\mathbb{V}$ publicado é unicamente analisado pelas autoridades e um valor r^+ específico pode ser artificialmente construído e cifrado com facilidade, sendo enviado aos guardiões apenas para formalização de sua decifração. Tal característica viabiliza a um possível adversário a possibilidade de publicação de valores distintos daqueles que seriam originalmente obtidos por um processamento íntegro dos produtos da eleição.

Diante de tal cenário, vale o questionamento acerca de quão forte é a garantia de integridade eleitoral provida pela equação de auditoria universal mostrada. Nesse sentido, o Teorema 4.2.18 a seguir demonstra as probabilidades associadas à verificação da integridade eleitoral, sob hipóteses de integridade dos recibos e das cédulas geradas durante a votação, para diferentes cenários de compromisso dos resultados publicados:

Teorema 4.2.18. *Seja $\mathscr{E}$ uma eleição totalizada para a qual os resultados íntegros são expressos pela tupla $(r^+, \mathbb{V}, \mathbb{C})$ e suponha-se que o conjunto $\mathbb{C}$ de recibos seja obrigatoriamente publicado de forma íntegra. Denote-se por OK uma variável aleatória que recebe o valor 1 se a equação de auditoria universal para a eleição é verificada com sucesso e 0 em caso contrário. Adicionalmente, denote-se por $RESULTS$ outra variável aleatória que recebe o valor 1 quando os resultados publicados para a eleição são íntegros e 0 em caso contrário. Então, sendo N um parâmetro de segurança associado a $\mathscr{E}$ e $negl$ uma função negligenciável em tal parâmetro, tem-se, até o final do período eleitoral, que:*

- $Pr[OK = 1 | RESULTS = 1] = 1$
- $Pr[OK = 1 | RESULTS = 0] = negl(N)$

Demonstração. Denotando-se por $\mathbb{R}$ o conjunto de todos os códigos de segurança em texto pleno correspondentes àqueles cifrados em $\mathbb{E}$, o Teorema 4.2.16 previamente demonstrado expõe que, sob hipótese de integridade dos resultados publicados (i.e. $RESULTS = 1$), um auditor que verifique a equação de auditoria universal (i.e. a igualdade entre o recibo de integridade eleitoral e o produtório dos recibos publicados) necessariamente constata sua validade, de modo que:

$$Pr[OK = 1 | RESULTS = 1] = 1$$

Seja então considerada a situação em que os resultados publicados não são íntegros (i.e. RESULTS $= 0$). Isto implica que os valores r^+ ou $\mathbb{V}$ (ou ambos) conforme publicados se encontram em disparidade com aqueles que seriam verificados por uma totalização íntegra da eleição. Note-se que não são consideradas aqui alterações no conjunto $\mathbb{C}$ de recibos, por hipótese do teorema. Assim, como $\mathbb{C}$ se encontra íntegro, identifica-se um valor $c^* = \alpha^{r^+} \cdot \beta^{h^+}$ que é fixo para a eleição $\mathscr{E}$, em que r^+ corresponde ao somatório íntegro de todos os códigos de segurança e h^+ corresponde ao somatório dos hashes de todos os votos presentes no conjunto $\mathbb{V}$ de votos íntegros.

Assim sendo, denote-se por $r^{+\prime}$ o somatório de códigos conforme publicado. Da mesma forma, denote-se por $h^{+\prime}$ o somatório dos hashes de votos para o conjunto $\mathbb{V}'$ conforme publicado. Para que a equação de auditoria universal seja verificada com sucesso, os valores $r^{+\prime}$ e $h^{+\prime}$ referentes aos resultados publicados devem ser compatíveis com os recibos em $\mathbb{C}$ de modo que $\alpha^{r^{+\prime}} \cdot \beta^{h^{+\prime}} = \alpha^{r^+} \cdot \beta^{h^+}$. Sendo q a ordem prima do grupo

utilizado, há apenas duas possibilidades complementares de fazê-lo, ambas verificadas pelo já mostrado Teorema 4.2.10:

1. $r^{+\prime} \not\equiv r^{+} \pmod{q}$: nesse caso, de modo a se verificar a equação de auditoria universal com sucesso, necessariamente deve-se ter $h^{+\prime} \not\equiv h^{+} \pmod{q}$.

2. $r^{+\prime} \equiv r^{+} \pmod{q}$: nesse caso, de modo a se verificar a equação de auditoria universal com sucesso, necessariamente deve-se ter $h^{+\prime} \equiv h^{+} \pmod{q}$.

Portanto, e representando-se tais situações em termos de $r^{+\prime}$, tem-se a seguinte igualdade:

$$Pr[OK = 1|RESULTS = 0] = Pr[OK = 1|r^{+\prime} \not\equiv r^{+}] + Pr[OK = 1|r^{+\prime} \equiv r^{+}]$$

Em todos os casos, assume-se que necessariamente $\mathbb{V}' \neq \mathbb{V}$, sob a hipótese de que não faria sentido a um atacante tentar empreender um ataque sobre os resultados eleitorais nas condições do teorema que não culmine em adulteração do montante de votos.

Suponha-se então que um adversário produza um conjunto de votos $\mathbb{V}'$ tal que $h^{+\prime} = \sum_{v \in \mathbb{V}'} H(v) \neq h^{+}$. De modo a manter a correspondência com o produtório de recibos, esse adversário terá de fornecer um valor $r^{+\prime}$ tal que $\alpha^{r^{+\prime}} \cdot \beta^{h^{+\prime}} = c^{*}$, em que necessariamente, por conta de $h^{+\prime} \neq h^{+}$, deve-se ter $r^{+\prime} \neq r^{+}$. Entretanto, o Teorema 3.4.4 mostra que se um tal adversário conhece dois pares (r^{+}, h^{+}) e $(r^{+\prime}, h^{+\prime})$ distintos que satisfaçam a equação de auditoria universal, então ele consegue resolver o logaritmo discreto $\log_{\alpha} \beta$, em que α, β são os geradores independentes de $\mathscr{E}$. Portanto, pela Definição 4.2.2, um adversário apenas consegue resolver tal problema com probabilidade negligenciável durante o período eleitoral, de modo que:

$$Pr[OK = 1|r^{+\prime} \not\equiv r^{+}] = negl(N)$$

Por outro lado, mantendo-se $r^{+\prime} \equiv r^{+}$, um adversário não precisa resolver o logaritmo discreto entre os geradores, mas em contrapartida precisa encontrar um conjunto de votos $\mathbb{V}' \neq \mathbb{V}$ tal que $|\mathbb{V}'| = |\mathbb{V}|$ e $h^{+\prime} \equiv h^{+} \pmod{q}$. Seja o conjunto $\{v_1, v_2, ..., v_n\}$ representativo de todos os possíveis votos válidos distintos que podem ser produzidos para a eleição $\mathscr{E}$. Assim sendo, e utilizando-se das simplificações notacionais $h_i = H(v_i)$ e $T = h^{+}$, o adversário está interessado em encontrar um conjunto $\mathbb{V}'$ tal que:

$$h^{+\prime} = x_1 \cdot h_1 + x_2 \cdot h_2 + ... + x_n \cdot h_n \equiv T \pmod{q},$$
$$\sum_{1 \leq i \leq n} x_i = |\mathbb{V}| \text{ e } x_i \geq 0$$

Dessa forma, encontrando um tal conjunto $\mathbb{V}' \neq \mathbb{V}$ que satisfaça a equação acima, o adversário consegue publicar resultados não íntegros tais que $Pr[OK = 1] = 1$. Cabe, portanto, computar a probabilidade de que um adversário consiga produzir um conjunto $\mathbb{V}'$ com tal característica. Para tanto, e lembrando que $T = h^+$ é um valor fixo dado o conjunto $\mathbb{C}$ de recibos publicados, sejam consideradas as seguintes observações probabilísticas:

1. Seja um hash h de um voto v válido qualquer e seja considerada a probabilidade $Pr[h \equiv T \pmod{q}]$. Como o hash de votos H atribui a h $= H(v)$ um valor fixo mas uniformemente escolhido em $\mathbb{Z}_q$, a probabilidade de que tal voto tenha sido mapeado exatamente para T é dada por $Pr[h \equiv T \pmod{q}] = 1/q$.

2. Sejam agora dois votos válidos e distintos v_1 e v_2 aos quais o hash de votos H atribuiu respectivamente valores h_1 e $h2$ fixos, mas também uniformemente escolhidos em $\mathbb{Z}_q$. Seja então considerada a probabilidade $Pr[h_1 + h_2 \equiv T \pmod{q}]$. Observe-se que, para cada possível valor atribuído a h_1, existe exatamente um único valor para h_2 de modo que $h_1 + h_2 \equiv T \pmod{q}$. Como existem exatamente q possibilidades de valores atribuíveis a h_1, e observando-se que o espaço amostral de atribuições equiprováveis desses dois valores tem cardinalidade q^2, tem-se que $Pr[h_1 + h_2 \equiv T \pmod{q}] = q/q^2 = 1/q$.

3. Seguindo-se o mesmo raciocínio acima exposto é fácil demonstrar que, considerando-se a soma de n hashes de votos distintos, tem-se tal respectiva probabilidade dada por $Pr[h_1 + h_2 + ... + h_n \equiv T \pmod{q}] = q^{n-1}/q^n = 1/q$.

4. Para um conjunto de resultados eleitorais, no entanto, é plenamente possível a ocorrência de repetições de um mesmo voto. Seja considerado então um número x de repetições de um voto v de hash h fixo, mas uniformemente atribuído por H em $\mathbb{Z}_q$, em que $x \leq S$, sendo $S = |\mathbb{V}'|$ assumido tal que $S \ll q$. Como q é primo, necessariamente existe $x^{-1} \in \mathbb{Z}_q$, de modo que $x \cdot h \equiv x \cdot h' \pmod{q} \implies h \equiv h' \pmod{q} \implies h = h'$. Isto é, se dois valores de hash h e h' apresentam o mesmo valor módulo q para um mesmo múltiplo x, então esses dois valores de hash são idênticos. Isto implica que, fixo o valor x, o múltiplo $x \cdot h$, a depender do valor $h = H(v)$, pode assumir qualquer valor em $\mathbb{Z}_q$. Logo, $Pr[x \cdot h \equiv T \pmod{q}] = 1/q$.

5. A partir do acima exposto, e seguindo-se o mesmo raciocínio seguido na exposição para somatórios de votos distintos, é também de fácil demonstração a probabilidade $Pr[x_1 \cdot h_1 + x_2 \cdot h_2 + ... + x_n \cdot h_n \equiv T \pmod{q}] = 1/q$.

Portanto, a probabilidade de que um adversário consiga produzir um conjunto de votos $\mathbb{V}'$ tal que $h^{+\prime} = h^+$ é igual a $1/q$ para cada tentativa empreendida. Assumindo-se o número de tentativas empreendidas pelo adversário durante o período eleitoral como

limitado em caráter polinomial, tem-se que a probabilidade de sucesso é negligenciável em N, dado que $q = 2^N$. Portanto:

$$Pr[OK = 1 | r^{+\prime} \equiv r^{+}] = negl(N)$$

Isto posto, a probabilidade de que um adversário que publique resultados não íntegros consiga passar pela checagem da equação de auditoria universal é dada por:

$$\begin{aligned} Pr[OK = 1 | RESULTS = 0] &= Pr[OK = 1 | r^{+\prime} \not\equiv r^{+}] + Pr[OK = 1 | r^{+\prime} \equiv r^{+}] \\ &= negl(N) + negl(N) \\ &= negl(N) \end{aligned}$$

□

A hipótese de que o conjunto $\mathbb{C}$ de recibos publicado corresponde a recibos oficiais não é surreal: as assinaturas digitais oficiais sobre os recibos individuais transferem a estes, assumidas suas premissas de confiança, uma oficialidade irrefutável, de modo que um adversário não pode excluir recibos das listagens oficiais sem correr o risco de que tal exclusão seja trivialmente detectável pelo detentor do recibo. Ademais, dada a publicidade da chave de verificação de assinatura, recibos não assinados porventura incluídos podem ser trivialmente verificados por qualquer um como inválidos.

Cabe notar, entretanto, que um adversário que detenha controle das chaves de assinatura ou da máquina de votação que as utiliza é capaz de incluir triplas (r, v, c) aos resultados sem qualquer prejuízo à equação de auditoria universal. Um exemplo desse tipo de adulteração se dá na forma da popularmente conhecida *fraude de mesário*, em que um representante das autoridades eleitorais, se aproveitando da distração ou com o consentimento dos demais presentes, se utiliza de seu acesso privilegiado à máquina de votação para votar em nome e à revelia de algum dos eleitores ausentes.

Por sua vez, a hipótese de validade das cédulas depositadas também não é surreal: eleitores podem ser instruídos a se utilizar do processo de desafio de modo a angariar certeza suficiente de que sua cédula se encontra corretamente construída. Se uma ínfima amostragem de eleitores aleatórios proceder com o desafio, a validade das cédulas depositadas pode ser probabilisticamente verificada com ampla margem de certeza, conforme mostrado e discutido em maiores detalhes na Subseção 4.4.3.2. Esta mesma subseção mostra que cédulas inválidas porventura depositadas abrem margem a adulteração pontual procedida sobre os votos em tais cédulas representados.

4.3 Propriedades

Nesta seção são agregadas e brevemente discutidas as propriedades providas pelo protocolo descrito neste capítulo.

4.3.1 Independência de Software

A execução do protocolo induz a produção de registros impressos de cada voto individual, os quais passam a existir de maneira independente da máquina e em formato legível e verificável pelo eleitor. Ademais, o processo de totalização induz o uso desses registros para a geração dos resultados eleitorais. Ao final da execução do protocolo, portanto, fica criada uma trilha de evidências que é auditável e pode ser analisada para conferências ou recontagens conforme exigidas ou solicitadas.

4.3.2 Capacidades E2E-V e Auditoria Universal

Como parte do processo de votação, o eleitor recebe um recibo, correspondente a um compromisso de Pedersen, de seu voto. A possibilidade de desafio do voto permite ao eleitor atestar que seu recibo está sendo corretamente gerado em conformidade com suas opções de escolha. E com base nesse mesmo recibo o eleitor pode se assegurar, mediante devido registro deste junto aos resultados eleitorais, de que seu voto foi corretamente computado na totalização da eleição. Ademais, como o recibo é assinado com uma chave de assinatura vinculada às autoridades eleitorais, um eleitor que perceba que seu recibo não se encontra registrado detém em mãos prova de ocorrência de irregularidade. Assim, cada voto individual pode ser acompanhado por quem tenha de seu respectivo recibo em seu trâmite pelo fluxo eleitoral.

Em termos de capacidades de auditoria pública, dado o registro dos recibos e a listagem dos votos computados durante a totalização, qualquer observador interessado no processo se utiliza de tal conjunto de dados de acesso público para verificar que o recibo de integridade da eleição é válido, o que confere garantias de integridade aos resultados publicados a menos de probabilidades negligenciáveis.

4.3.3 Suporte a Cédulas Extensas

Vários protocolos criptográficos de votação se utilizam de metodologias de codificação de votos que culminam em artefatos que crescem juntamente com o número de opções. Um exemplo é o esquema descrito por [13] e que é utilizado por sistemas como o STAR-Vote. Embora tal esquema viabilize uma conveniente totalização homomórfica dos resultados, o tamanho de uma cédula preenchida é diretamente proporcional ao número de cargos e de opções por cargo da eleição, tornando sua aplicação consideravelmente custosa para eleições com grande número de opções por cédula, como é o caso brasileiro.

O presente protocolo, por sua vez, faz uso metódico de uma função de hash sobre formatações digitais bem-definidas dos votos, conferindo-lhe plena adaptabilidade a variadas composições de cédula ao manter o tamanho de sua representação digital constante. Esta estratégia abre mão da possibilidade de totalização homomórfica para viabilizar suporte à realização de eleições com amplo número de candidatos sem incorrer em sobrecarga adicional nos tamanhos dos artefatos gerados.

4.3.4 Acoplamento a Sistemas Pré-Existentes

Em seu cerne, o protocolo especifica um método de geração de recibos que atestam o depósito de artefatos por parte de indivíduos. Em um contexto eleitoral, essa operação se traduz em geração de recibos que comprovam o depósito de votos por parte dos eleitores que participam de uma eleição.

Isto posto, protocolos de votação que se utilizem de registros físicos podem se utilizar da metodologia descrita para embutir em suas eleições capacidades de verificação de ponta-a-ponta e auditoria universal. Para tanto, basta que cada voto possa ser traduzido de forma precisa para uma representação digital que possa ser empregada como expoente de um gerador do grupo algébrico utilizado na eleição.

4.3.5 Capacidades Configuráveis de Granularização

Na descrição do protocolo especifica-se uma eleição por meio da definição de múltiplos conjuntos a ela associados e posterior geração de parâmetros diversos que viabilizam

sua execução. Os eleitores, então, se utilizam de uma máquina de votação para construir cédulas e depositá-las, podendo conferir sua presença íntegra na totalização dos resultados a partir de seu recibo. Adicionalmente, quaisquer observadores interessados podem verificar a integridade dos resultados por meio da equação de auditoria universal.

Entretanto, a falha na verificação da equação de auditoria universal apenas especifica que algo de errado ocorreu, mas por si só não oferece indicativos acerca do que pode ter ocorrido ou mesmo em que contexto. Ademais, em eleições com grande número de eleitores, é plausível que múltiplas máquinas de votação alocadas a distintos locais de votação sejam empregadas de modo a viabilizar a execução eleitoral, cada qual destas possivelmente com conjuntos distintos de opções de escolha.

Esse processo de *granularização* de uma eleição em múltiplas sub-eleições, cada qual composta por possíveis configurações distintas de cédula, é prontamente suportado pelo protocolo sem necessidade de grandes modificações em sua execução, viabilizando fácil adaptação frente às necessidades da eleição corrente. O processo simplesmente consiste em interpretar cada seção eleitoral como uma execução individual de uma eleição em menor escala.

A principal vantagem da granularização, além de facilitar a execução de eleições em grande escala, é que com esta é possível aumentar o nível de precisão da verificação da equação de auditoria universal, o que, em caso de erros, permite não apenas constatar o fato de que algo deu errado, mas também onde (e.g. em que local de votação ou em que máquina) tais erros aconteceram. Isto viabiliza uma filtragem entre seções nas quais a equação foi verificada com sucesso e seções nas quais alguma possível falha de integridade ocorreu, possibilitando respostas mais precisas com base em previsões legais ou protocolos vigentes (e.g. anulação dos votos da seção ou máquina específica, execução de nova votação para seções com falhas de verificação).

4.4 Ataques

Esta seção descreve possíveis ataques ao sistema e como o protocolo base se comporta com relação a estes. Adicionalmente, e quando cabível, são discutidas também formas de adaptação do protocolo base de modo a contornar ou mitigar os ataques mostrados. Não se busca aqui prover uma listagem exaustiva de possíveis ataques ao sistema, mas sim ilustrar aqueles facilmente identificáveis ou considerados de maior interesse e relevância prática.

4.4.1 Depósito de Votos Indevidos

Da mesma forma que outros protocolos E2E-V, o presente protocolo também não busca solucionar por si só a questão do depósito de *votos indevidos* (c.f. Definição 4.1.11), prática também conhecida como *ballot stuffing* ou, em termos mais brasileiros, generalizada sob a alcunha de *fraude de mesário*. No contexto do protocolo proposto neste trabalho, dois aspectos contribuem para possibilidades de votos indevidos ou dificuldades de sua possível correção:

1. **Acesso privilegiado**: alguns atores eleitorais possuem acesso privilegiado às máquinas de votação. É o caso, por exemplo, de representantes das autoridades eleitorais, como os mesários, que podem se utilizar de seus privilégios de acesso para depositar votos em nome de outros eleitores e à sua revelia.

2. **Sigilo Incondicional**: ainda que haja possibilidade de detecção de depósito de votos em caráter indevido, a incondicionalidade do sigilo inviabiliza medidas corretivas no sentido de remoção desses votos, uma vez que torna-se impossível se estabelecer ligação entre o recibo de um eleitor e seu voto depositado.

Em sistemas que não prezam por sigilo incondicional, uma possível contramedida encontra-se precisamente no fato de que tais sistemas costumam associar a cada eleitor a versão cifrada de seus votos. Assim sendo, nestes sistemas torna-se viável que votos possam ser removidos, a pedido de eleitores e sob justificativas formais, por meio da remoção dos respectivos votos cifrados em seus nomes antes de quaisquer procedimentos de totalização. O sistema Helios [31], por exemplo, utiliza-se de tal vinculação entre voto cifrado e identidade do eleitor para viabilizar a eleitores a possibilidade de revotação sem contagem dupla de votos, ao custo da facilitação de violação futura do sigilo do voto individual.

No presente protocolo, a incondicionalidade do sigilo impede tal tipo de vinculação entre eleitores e votos sem que se incorra em semelhante violação de suas características de sigilo incondicional. A possibilidade óbvia de mitigação passa pelo estabelecimento de controle rígido de acesso à máquina de votação no sentido de impedir que acessos indevidos a esta e às urnas venham a ocorrer, ou seja, por procedimentos adequados de segurança no próprio local de votação [41, p.329-330]. O que consiste em um procedimento de segurança adequado ou em controle suficientemente rígido, no entanto, não é de simples determinação. No Brasil, por exemplo, a vinculação de eleitores a seções eleitorais únicas junto à exigência de que cada voto depositado esteja vinculado a um eleitor devidamente identificado operam naturalmente como fatores limitantes do potencial dessa adulteração em uma seção eleitoral.

4.4.2 Remoção de Votos

Entenda-se por *remoção* a tentativa de subtração de um determinado voto v da contabilização nos resultados, sem pretenções de substituição deste por qualquer outro voto (c.f. Subseção 4.4.3 para tentativas de tal natureza).

Observe-se que a tentativa de simplesmente remover um voto v do conjunto $\mathbb{V}$ de votos registrados viola a restrição de integridade básica dos resultados, que dita que $|\mathbb{E}| = |\mathbb{V}| = |\mathbb{C}|$, de tal modo tornando patente a remoção procedida. Portanto, de modo a tentar remover um voto mantendo a ilusão da integridade dos resultados, é necessário que ao menos um valor de $\mathbb{E}$ e um valor de $\mathbb{C}$ sejam também removidos. No entanto, como o montante $\mathbb{E}$ não é publicado, no referente a observadores externos, basta que valores de $\mathbb{V}$ e de $\mathbb{C}$ sejam removidos em igual cardinalidade para que a restrição básica de integridade seja publicamente mantida. Não obstante, como os recibos c são individualmente assinados digitalmente pelas autoridades, a remoção de um recibo particular torna o procedimento plenamente detectável por parte do eleitor que o detém, de modo que este pode então contestar sua remoção com lastro na assinatura oficial válida referente a este.

Suponha-se então que o adversário que empreende a tentativa de remoção consiga de alguma forma identificar recibos que ele possa ter certeza que não serão verificados (e.g. *trash attacks* [24]). Mesmo sob tal hipótese, seria necessário ao adversário conhecimento dos valores r em texto pleno para promover a remoção com garantias de manutenção da validade da equação de auditoria universal. Em tal caso, bastaria ao adversário encontrar a combinação (r, v) que leva à produção do recibo c que sabidamente não será verificado (c.f. Teorema 4.2.12) e remover cada qual de seus respectivos conjuntos. Mas, sem conhecimento dos valores r em texto pleno, e sob hipótese de uniformidade na geração destes quando da produção das cédulas, cada qual dos recibos $c \in \mathbb{C}$ é uniformemente viável de ser ligado a qualquer voto $v \in \mathbb{V}$. Desse modo, restaria ao adversário a possibilidade de tentar adivinhar aleatoriamente o par (r, v) correspondente ao recibo c que pretende remover, o que ele apenas acerta com probabilidade $1/S^2$, em que S denota a cardinalidade coletiva dos conjuntos $\mathbb{E}$, $\mathbb{V}$ e $\mathbb{C}$.

4.4.3 Adulteração de Votos

Sendo V o espaço de votos (c.f. Definição 4.1.9), entenda-se por *adulteração* a tentativa de alteração de um voto $v \in V$ para um voto $v' \in V$ em que necessariamente $v' \neq v$.

4.4.3.1 Adulteração Por Simples Troca

A maneira mais pueril de tentar um ataque de adulteração seria pela simples substituição de um voto v por um voto v' tal que $v' \neq v$, sem qualquer modificação adicional. Tal substituição, no entanto, é prontamente identificável pelo eleitor durante a votação, ou, se procedida sobre a própria urna de votos, atestável por qualquer observador, a menos de probabilidade negligenciável, via equação de auditoria universal, a qual acusaria a inconsistência dos resultados.

Esta metodologia de adulteração apenas encontraria sucesso no caso em que fosse procedida por geração de $r^{+\prime} \neq r^{+}$ capaz de manter válida a equação de auditoria universal, ou pela geração de um conjunto de votos adulterado $\mathbb{V}'$ que permitisse manutenção do somatório r^{+} original. Para ambos os casos, o Teorema 4.2.18 mostra que o adversário apenas obtém sucesso com probabilidade negligenciável.

4.4.3.2 Adulteração Por Cédulas Inválidas

Suponha-se que de alguma forma uma máquina de votação consiga identificar que um eleitor ~~e~~ não irá desafiar seu voto. Ainda assim, dada a legibilidade do voto v impresso, assume-se que a máquina seja incapaz de adulterar v sem correr um risco proibitivo de que tal adulteração seja percebida, de modo que v necessariamente deve corresponder às opções manifestas pelo eleitor.

Não obstante, a máquina pode produzir uma suposta cédula (r, v, c) de modo que $c \neq \alpha^{r} \cdot \beta^{H(v)}$ mas tal que $c = \alpha^{r} \cdot \beta^{H(v')}$ em que $v' \neq v$. Portanto, observe-se que a tripla (r, v', c) forma uma cédula válida de fato, mas a tripla impressa (r, v, c) não. Tal fato, entretanto, é mantido despercebido por conta do caráter oculto de r dada sua necessária cifração.

Assim sendo, note-se que a equação de auditoria universal, a menos de uma probabilidade negligenciável, não mais será verificada com sucesso se considerada com o voto v depositado. Entretanto, em sendo substituído este voto v pelo voto v' na etapa de totalização, a equação de auditoria passa a ser válida novamente. Mais ainda, o recibo c em posse do eleitor permanece válido, de modo que o eleitor tem a impressão de que seu voto foi corretamente contabilizado.

Este ataque respalda-se na suposição de que o eleitor não irá proceder com o desafio da cédula. Do contrário, a máquina teria de ser capaz de produzir o único valor r' tal que $c = \alpha^{r'} \cdot \beta^{H(v)}$, o que, conforme demonstra o Teorema 3.4.4, é equivalente a computar o

logaritmo discreto $\log_\alpha \beta$ entre os geradores usados, aqui supostos independentes. Além disso, múltiplas análises apontam ser suficiente que apenas um pequeno percentual de cédulas seja auditado com relação à sua devida correspondência com os votos registrados no intuito de se ter ampla margem de confiança nos resultados eleitorais produzidos [10, 20, 27].

Dessa forma, seria necessário que apenas um baixo percentual aleatório de eleitores passasse pelo processo de desafio individual da cédula de modo a se obter elevada margem de certeza dos resultados eleitorais providos. De modo a não depender estritamente da boa vontade dos eleitores, auditores devidamente identificados podem ser instados a produzir votos em seções aleatórias com o único propósito de solicitar seu desafio, de tal forma produzindo um percentual suficiente para uma margem de certeza satisfatória dos resultados.

4.4.4 Depósito de Votos Desafiados

Ao ser solicitado o desafio do voto durante a etapa de votação, a máquina utilizada imprime o código de segurança r em texto pleno. Este procedimento é necessário, visto que o código de segurança em si precisa ser utilizado para verificar se a cédula produzida é de fato válida. Entretanto, a revelação do parâmetro r permite pronta identificação do voto do eleitor com base no recibo c, conforme já demonstrado pelo Teorema 4.2.11. Assim, é crucial que um eleitor seja impedido de depositar um voto desafiado, ou ao menos que um voto desafiado seja impedido de ser considerado durante a totalização.

Sob uma hipótese de atomicidade do processo de solicitação do desafio e o desafio em si, tal questão poderia ser prontamente resolvida por meio da consulta, após totalização, da base de dados de cédulas desafiadas. Dessa forma, cédulas encontradas na totalização que porventura se encontrem também na base de dados de desafios poderiam ser automática e justificadamente removidas dos resultados. Para tanto, seria necessário que cada componente de cédula produzido contasse com algum tipo de identificador pessoal e único que permitisse tal verificação.

Tal supracitada hipótese de atomicidade da solicitação de desafio e do desafio em si, entretanto, pode ser facilmente quebrada por conivência dos representantes das autoridades eleitorais presentes no local de votação, os quais podem permitir que um eleitor, tendo desafiado seu voto, imediatamente possa gerar novo voto para depósito sem ter de passar pelo processo de desafio. Isto viabilizaria a tal eleitor a possibilidade de depósito de um voto desafiado.

Uma possível solução a este problema seria embutir também alguma forma de inter-

ligação do pedido de desafio do voto com uma marcação física na cédula (e.g. perfurações físicas visíveis no centro de cada componente impresso) que tornasse fisicamente claro e distinguível que tal cédula passou por um processo de desafio. Tal marcação, embora não impeditiva do ato físico de depósito de uma cédula desafiada, viabiliza que componentes de tal cédula porventura depositados, em sendo encontrados fisicamente marcados durante a etapa de totalização, possam ser devidamente documentados e então desconsiderados dos resultados, desmotivando tentativas de adulteração de tal natureza.

4.4.5 Ataques de Colisão

Ataques de Colisão, do inglês *Clash Attacks*, são ataques de considerável simplicidade que operam sobre as capacidades de verificação de sistemas E2E-V [23]. O ataque basicamente consiste em informar recibos idênticos para múltiplos eleitores. Procedido com sucesso, a presença de um único destes recibos no quadro de votos final da eleição é capaz de convencer todos os eleitores afetados de que seu voto se encontra devidamente incluso na totalização, haja vista que todos esses eleitores poderão verificar a presença do dito recibo na listagem. Isto abre margem para que o atacante substitua os demais votos de recibo idêntico por outros votos compostos como bem entenda, sem qualquer capacidade de detecção pelos mecanismos de integridade do sistema. No entanto, um ataque de colisão pode ser verificado se dois eleitores afetados expuserem simultaneamente seus recibos idênticos e constatarem a presença de apenas um destes no quadro de votos final.

Para que um ataque de colisão possa ser executado com sucesso, é necessário explorar a estruturação dos recibos e a ligação destes com as opções escolhidas durante a votação de modo que o recibo assim produzido possa ser verificado como válido para distintos eleitores e suas respectivas opções. A depender do sistema de votação, o ataque apenas funciona para eleitores com o mesmo conjunto de opções. Entretanto, no sistema aqui descrito, por suas capacidades de sigilo incondicional, há a possibilidade de produção de recibos iguais para opções de voto completamente distintas. Isto é, é plenamente possível se produzir $c_1 = \alpha^{r_1} \cdot \beta^{H(v_1)}$ e $c_2 = \alpha^{r_2} \cdot \beta^{H(v_2}$, com $v_1 \neq v_2$, tais que $c_1 = c_2$, embora o Teorema 3.4.4 mostre que isto equivale a computar o logaritmo discreto $\log_\beta \alpha$.

Isto posto, uma forma mais plausível de executar este ataque sem necessidade de resolver o DLOG seria controlar a aleatoriedade de r de modo a permitir a ocorrência de colisões de recibo para votos iguais. Esta possibilidade é tanto mais plausível quanto menor for $|V|$. Assim, uma máquina vulnerabilizada poderia de modo sub-reptício limitar os possíveis valores de r a um subconjunto reduzido e de aparência aleatória, permitindo colisões de recibos com probabilidades controladas.

Uma maneira de se evitar esse tipo de ataque se dá pela inclusão de algum parâmetro adicional no recibo que lhe confira uma característica de unicidade independente da sua estrita composição matemática em si. Um parâmetro de trivial inclusão nesse sentido é o próprio horário de votação, com precisão na casa das unidades de segundo. Conforme bem ilustrado por [3], o horário e a ordenação dos votantes, ao menos nos moldes das eleições presenciais correntes, não são aspectos sobre os quais deva recair qualquer tipo de premissa de sigilo. Assim sendo, sob tais moldes correntes, não deve haver possibilidade de prejuízos ao sigilo do eleitor associados ao conhecimento de seu horário de votação. Ressalta-se, entretanto, que tal marcação de horário deve estar presente unicamente no recibo c: sua associação ao registro impresso v do voto depositado, por exemplo, permitiria, pelas mesmas observações prévias, identificar o eleitor que o depositou.

Dessa forma, para que dois recibos sejam idênticos, seria necessário à máquina de votação produzi-los informando o mesmo horário com precisão de segundos, o que, dada a natureza da votação presencial em locais com uma única máquina de votação, serviria como um demonstrativo da desonestidade dessa máquina, isto porque implicaria em dois eleitores operando-a naquele mesmo instante, o que não é factível em tal contexto. Caso haja a possibilidade de mais de uma máquina de votação no mesmo local, outros parâmetros adicionais são trivialmente passíveis de inclusão, como um identificador do terminal utilizado. Esta é a metodologia descrita por [5] como sendo utilizada na composição do recibo do STAR-Vote.

Capítulo 5

Implementação

Como parte do presente trabalho, uma implementação em software do protocolo foi desenvolvida no intuito de analisar algumas de suas características práticas e operacionais, além de prover uma ferramenta ilustrativa de seu funcionamento. Este capítulo descreve em maiores detalhes as decisões particulares tomadas para viabilizar a implementação. Adicionalmente, o protocolo é revisitado em suas etapas constituintes, mas agora em contexto aplicado e à luz da implementação provida. Ao final, alguns resultados derivados do exercício da implementação são brevemente discutidos.

5.1 Decisões

Conforme explicita o Capítulo 4, o protocolo de votação proposto se utiliza de esquemas criptográficos adicionais em sua composição. Embora alguns destes sejam explicitamente determinados, a exemplo do esquema de compromisso de Pedersen, outros, como o esquema homomórfico de cifração dos códigos de segurança, são deixados em aberto no intuito de conferir maior maleabilidade à descrição do mesmo. Não obstante, tais esquemas necessitam ser determinados de modo a se viabilizar uma implementação do protocolo de votação. Para além deste aspecto, outros de cunho mais prático também necessitam ser decididos, a exemplo das linguagens de programação utilizadas, os sistemas operacionais a serem suportados e mesmo as bibliotecas de código empregadas no projeto.

Isto posto, não se busca aqui classificar decisões como corretas ou não com relação a tais aspectos, sob o credo de que tais decisões provavelmente hão de se mostrar fortemente sensíveis ao contexto particular de aplicação. Como exemplo, em contextos nos quais haja algum tipo de exigência legal de que os artefatos de software empregados em uma eleição tenham código aberto, claramente sistemas operacionais e bibliotecas de cunho proprietário mostrar-se-ão de aplicabilidade proibitiva. Em outros contextos em que tal exigência não seja imposta, tais sistemas e bibliotecas podem vir a ser considerados como decisões de projeto viáveis.

5.1.1 Linguagem de Programação: Python

Um dos principais intuitos da implementação é que esta fosse o mais didática e inteligível possível, além de prover facilidades de sua migração entre diferentes sistemas e ambientes computacionais. Em suporte a tais requisitos, a linguagem Python, por sua ênfase em uma sintaxe descomplicada e mais próxima aos construtos das linguagens naturais, oferece um ferramental propício para desenvolvimento de artefatos de software mais facilmente compreensíveis e analisáveis por quem não estritamente da área da computação. Ademais, o caráter interpretado da linguagem Python confere à implementação facilidades associadas a sua execução sob diferentes sistemas e ambientes computacionais.

Isto posto, o desempenho da linguagem Python, haja vista seu caráter interpretado, é sabidamente inferior àquele provido por linguagens mais próximas às funcionalidades do próprio hardware, a exemplo de C e C++. No entanto, a implementação padrão da linguagem Python oferece a possibilidade de programação em linguagem C para módulos que exijam maior desempenho [37]. Este ferramental foi aproveitado na implementação para fins de aplicação de bibliotecas criptográficas consolidadas e para obtenção de melhor desempenho em funções mais exigentes em termos de processamento. Este aspecto é explorado em maiores detalhes na Subseção 5.1.6.

5.1.2 Representação de Artefatos Numéricos: Códigos QR

Para fins de representação dos valores inteiros associados ao código de segurança (cifrado e em texto pleno) e respectivo recibo de uma cédula produzida pelo protocolo, foram escolhidos os *Quick Response codes*, comumente denominados de *QR codes*. Estes consistem em códigos bidimensionais padronizados que se mostram de rápida leitura por dispositivos comumente disponíveis, como *smartphones*. As principais motivações para tal escolha residem em sua natureza bidimensional e em suas capacidades configuráveis de correção de erros em caso de eventuais danos a porções dos artefatos que codificam.

O aspecto da bidimensionalidade se mostra como uma característica positiva por dotar os códigos QR de melhores capacidades de aproveitamento do espaço físico por eles ocupado quando impressos em papel. Os tradicionais códigos de barra unidimensionais tendem a crescer em uma única dimensão em caráter proporcional ao tamanho da informação que codificam, enquanto códigos QR distribuem seu crescimento tanto em âmbito horizontal quanto vertical. Essa característica os torna mais aplicáveis no contexto de produção de cédulas impressas por tecnologia térmica, em que as dimensões do papel

usado, principalmente a horizontal, são potencialmente muito mais limitadas.

Já em termos da capacidade configurável de correção, esta mostra-se de particular interesse frente à necessidade de manuseio do recibo pelos eleitores e seu eventual trânsito em condições precárias (e.g. um bolso apertado, uma mão suada ou um carro quente) entre sua produção e respectiva conferência. Uma representação que se mostre muito sensível a danos, e quando da ocorrência destes, pode vir a impossibilitar o procedimento de auditoria individual por parte do eleitor, invalidando de tal forma a utilidade primeira do recibo produzido. Cabe ressaltar que quanto maior a capacidade de correção, maior o tamanho em papel ocupado pelo código produzido. Há, portanto, uma necessidade, possivelmente de viés experimental e sensível a contexto de aplicação, de melhor se determinar as capacidades corretivas ideais. Para fins da presente implementação, optou-se pelo nível de correção médio M, que provê capacidades corretivas de até 15% do código, sendo amplamente aplicado para fins comerciais e para necessidades de armazenamento de menor duração.

Como um aspecto final, ressalta-se que códigos QR são atualmente empregados em ampla gama de contextos e para propósitos diversos, de modo que, além de terem sido comprovados como plenamente viáveis em aplicações práticas, hoje se mostram também como um popular e relativamente compreensível mecanismo de codificação de informação, potencialmente facilitando a compreensão de seu manuseio por parte dos eleitores.

5.1.3 Cifração Homomórfica: Criptossistema de Paillier

Conforme discutido no Capítulo 4, é necessário às premissas de sigilo do protocolo que os códigos de segurança utilizados sejam cifrados. No entanto, é também necessário, para fins de auditoria universal, que o somatório desses códigos seja obtido, motivo pelo qual o esquema de cifração empregado deve ser aditivamente homomórfico. Por conta de tal necessidade emprega-se na implementação o Criptossistema de Paillier [29] para cifração dos códigos de segurança, com fator $N = 2048$ bits, em que N é o produto de dois primos grandes.

Uma alternativa como ElGamal, apesar de homomórfica, apenas seria capaz de prover o produtório dos códigos de segurança (ElGamal tradicional) ou o somatório desejado dos códigos na forma de um expoente de um gerador definido (ElGamal exponencial). Este último, no entanto, e dado o aspecto uniforme da geração dos códigos de segurança, exigiria a resolução do problema do logaritmo discreto, inviabilizando sua aplicação com vistas à recuperação do somatório em si.

O Criptossistema de Paillier, em contraste, apresenta-se como aditivamente homo-

mórfico por padrão, permitindo pronta obtenção do somatório dos códigos pela decifração do produtório dos criptogramas. Ressalta-se, no entanto, que o custo do uso de Paillier se manifesta em dois principais aspectos:

1. A segurança de Paillier repousa sobre o desconhecimento do valor $\Phi(N)$ associado a um inteiro N que, conforme mencionado, é produto de dois primos grandes. Dessa forma a segurança de Paillier é semelhante à do RSA, exigindo tamanhos de chave (em bits) consideráveis.

2. Paillier opera sobre $\mathbb{Z}^*_{N^2}$ em que $N = p \cdot q$ com p e q primos grandes. Isto faz com que $|N| \approx 2048$ bits resulte em criptogramas da ordem de 4096 bits.

5.1.4 Grupo Algébrico e Assinaturas: Curvas Elípticas

Curvas elípticas são consideradas um padrão criptográfico mais moderno pela suposta dificuldade de resolução do problema do logaritmo discreto (DLP) nos grupos que definem. A grande vantagem destas sobre demais sistemas que se utilizam do DLP em outros grupos (e.g. grupos inteiros de ordem prima) reside no fato de que os algoritmos conhecidos para a resolução de tal problema no contexto de curvas elípticas são ainda menos eficientes. Isto se traduz em tamanhos menores de chaves e algoritmos mais eficientes para um mesmo nível de segurança se comparado com outros grupos algébricos. Tal fato é particularmente importante considerando-se a possibilidade de que os sistemas que venham a executar os algoritmos do protocolo descrito venham a ser particularmente limitados (e.g. computadores domésticos de baixo custo ou microcontroladoras).

Para fins da implementação, optou-se pelo uso da curva NIST P-256 principalmente por esta ser uma curva padronizada [28, p.91] amplamente documentada e implementada em diversas bibliotecas criptográficas disponíveis. Cabe ressaltar, no entanto, que há desconfiança corrente por parte do meio técnico, acadêmico e científico com relação aos valores supostamente aleatórios utilizados para definir os parâmetros dessa e de outras curvas, considerando-se via tal suposta aleatoriedade a possibilidade de introdução, em caráter proposital, de vulnerabilidades ainda indeterminadas para fins de espionagem governamental [8, 32]. Tal aspecto deve certamente ser levado em consideração quando da definição da curva elíptica a ser utilizada em possível contexto prático de aplicação do protocolo. Para as finalidades da implementação, no entanto, considera-se que a curva NIST P-256 seja plenamente suficiente.

Seguindo-se o mesmo princípio, para fins de produção das assinaturas digitais usadas para validação dos recibos produzidos optou-se pelo uso do *ECDSA* também

utilizando-se da curva NIST P-256. Pelos mesmos pretextos, a assinatura ECDSA pode ser feita em tamanho menor de bits para um nível de segurança que em outros grupos algébricos exigiria assinaturas consideravelmente maiores. Dado o aspecto de que tais assinaturas são agregadas visualmente ao recibo na forma de um código QR, isto viabiliza a produção de códigos legíveis de tamanho reduzido, ocupando menos espaço em papel e potencialmente se mostrando como uma representação mais econômica e mais facilmente legível por dispositivos de menor precisão.

5.1.5 Hash de Votos: SHA256 Truncado

Para fins de compatibilização com o grupo definido pela curva elíptica NIST P-256, optou-se pela fixação do hash de votos como uma aplicação do algoritmo SHA256 sobre representações em formato JSON dos votos em texto pleno. Ademais, tal valor é truncado pela remoção de um único bit à direita caso ultrapasse a ordem do grupo utilizado. Dessa forma garante-se que o valor assim produzido sempre se encontre dentro dos limites da ordem do grupo sem impacto significativo à necessária propriedade de resistência a colisões [16, p.9].

5.1.6 Biblioteca Criptográfica: RELIC/CREST

A implementação de algoritmos criptográficos é aspecto sensível e consideravelmente suscetível a erros muitas vezes sutis que podem render vulneráveis artefatos criptográficos produzidos a partir destes sem que tal aspecto seja necessariamente de fácil constatação. Isto posto, considerando-se que diversas bibliotecas criptográficas encontram-se já amplamente disponíveis e tendo em vista que a implementação de nova bibliteca criptográfica encontra-se fora do escopo do presente trabalho, no espírito da colaboração acadêmica e das boas práticas criptográficas, optou-se pelo emprego de uma biblioteca criptográfica já existente.

Assim, de modo a se aproveitar de esforços prévios de implementação criptográfica, optou-se pelo uso de algoritmos que fossem providos por bibliotecas correntemente mantidas e de ampla aplicação e estudo. A biblioteca escolhida para a implementação foi a RELIC por conta de sua ênfase em eficiência e flexibilidade com vistas à possibilidade de aplicação em sistemas embarcados [1]. Apesar de ser classificada como software em

padrão alfa de qualidade, trata-se de biblioteca em código aberto correntemente mantida, de aplicação documentada e estudada [34]. Outro aspecto importante para tal decisão é o fato de a biblioteca ser implementada em linguagem C, o que lhe torna prontamente adaptável à implementação Python tradicional (CPython) com seu interpretador padrão via a já mencionada capacidade de expansão modular [37].

Com isto foi desenvolvida a biblioteca CREST (Cpython Relic Envelope SuiTe), que consiste em uma série de módulos que tornam acessíveis as capacidades da RELIC via CPython. Por enquanto, principalmente dada as extensivas capacidades da biblioteca RELIC, apenas um subconjunto de suas capacidades se encontra com módulos implementados e disponíveis, a saber, aqueles estritamente necessários à implementação do protocolo eleitoral descrito por este trabalho. Adicionalmente, a biblioteca CREST provê ainda algumas implementações adicionais não providas pela RELIC, mas necessárias ao protocolo. Tais implementações adicionais foram diretamente programadas em linguagem Python utilizando-se de componentes da própria CREST em sua execução. A implementação atual não apresenta caráter estritamente voltado à retenção da eficiência provida pela RELIC: seu foco principal foi precisamente prover acesso às capacidades da RELIC via ambiente do interpretador Python, com ênfase na implementação do protocolo eleitoral descrito. Dessa forma, há ampla possibilidade corrente de melhorias na implementação com vistas ao desempenho e maior gama de aplicações.

5.2 Descrição

Nesta seção o protocolo de votação descrito no Capítulo 4 é revisitado em termos de suas etapas componentes, mas agora no contexto prático da implementação produzida. Uma versão em código Python da implementação em termos da codificação de suas etapas principais é mostrada como apêndice ao final do escrito.

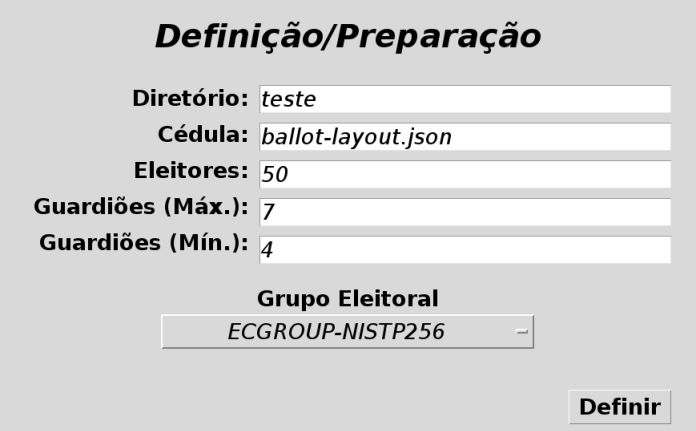
5.2.1 Definição e Preparação

Conforme mencionado quando de sua descrição no Capítulo 4, a etapa de definição mostra-se fortemente sensível ao contexto de aplicação do protocolo, de modo que as interações e decisões em seu contexto não são facilmente capturadas por meio de uma interface de usuário que faça sentido. Assim sendo, optou-se na implementação pela junção

das etapas de definição e preparação eleitoral, abstendo-se de tal forma dos processos burocráticos ou impugnatórios associados que são necessários à devida condução de uma eleição real.

Isto posto, a etapa conjunta de definição e preparação eleitoral se mostra relativamente direta, resumindo-se à definição dos diversos parâmetros eleitorais necessários à execução do protocolo, conforme ilustra a Figura 5.1:

Figura 5.1: Definição e preparação eleitoral



Fonte: Elaborado pelo autor.

Conforme mencionado, em se tratando de implementação ilustrativa do funcionamento do protocolo, as urnas físicas são substituídas por "urnas virtuais" simuladas por diretórios agregados sob um mesmo diretório raiz. No exemplo mostrado na Figura 5.1, um diretório raiz denominado *teste* é criado para a eleição, sob o qual são alocados diretórios respectivos para armazenamento dos componentes das cédulas depositadas.

Adicionalmente, em resumo da etapa de definição do conjunto $\mathbb{O}$ de opções, um arquivo devidamente formatado e contendo a disposição da cédula deve ser informado. Para fins da implementação, optou-se pela formatação em JSON tanto por sua ampla aplicação quanto por sua considerável legibilidade. Um exemplo simples de conteúdo de um arquivo de cédula é mostrado na listagem seguinte:

```
{"Deputado": {"11": "Deputado 1", "22": "Deputado 2",
              "33" : "Deputado 3", "44" : "Deputado 4",
              "0": "NULO"},
 "Senador": {"11": "Senador 1", "22": "Senador 2",
             "0": "NULO"},
 "Governador": {"11": "Governador 1", "22": "Governador 2",
                "33": "Governador 3", "0": "NULO"},
 "Presidente": {"11": "Presidente 1", "22": "Presidente 2",
                "0": "NULO"}}
```

Código 5.1: Cédula representada em formato JSON

Tal formatação de cédula estipula que a votação consistirá em escolhas para quatro cargos distintos (*Deputado, Senador, Governador e Presidente*), havendo cinco opções de escolha para *Deputado*, três opções para *Senador*, quatro para *Governador* e três para *Presidente*. Note-se que para cada cargo uma das opções consiste na possibilidade de anulação do voto, a qual, no âmbito da implementação, deve ser explicitada na formatação da cédula.

Em termos do conjunto $\mathbb{P}$ de eleitores, em um contexto de aplicação real do protocolo provavelmente é importante a especificação das identidades de cada qual destes, tanto no sentido de satisfação de exigências legais (e.g. determinar se o indivíduo está apto a votar) quanto no sentido da devida identificação do eleitor durante a etapa de votação. No contexto da implementação simplifica-se a definição de $\mathbb{P}$ para a simples definição de sua cardinalidade.

De igual maneira, a definição do conjunto $\mathbb{G}$ de guardiões muito possivelmente exige estabelecimento de identidades em contexto real de aplicação, sendo tal definição novamente simplificada no contexto da implementação também para mera determinação de sua cardinalidade. Entretanto, e conforme já mencionado, faz-se necessário também estipular o mínimo de guardiões que deve se reunir com suas parcelas para possibilitar recuperação da chave de cifração eleitoral.

Por fim, o grupo a ser empregado é selecionado. Conforme menção prévia, para fins da presente implementação o grupo escolhido consiste na curva elíptica NIST P-256. Ademais, por questões já mencionadas de compatibilização com a ordem da curva referente ao grupo, o hash de votos escolhido é um SHA256 com redução de um bit para garantir que os resultados se encontrem sempre dentro dos limites da ordem da curva.

Estando todos os supracitados parâmetros devidamente definidos, o botão *Definir* produz toda a estrutura de diretórios necessária à simulação eleitoral e adicionalmente gera os demais parâmetros eleitorais necessários à execução do protocolo, a saber, os geradores eleitorais independentes, as chaves guardiãs e suas parcelas e a chave de assinatura dos resultados. Os parâmetros públicos que definem a eleição são agregados em uma representação JSON, conforme mostrada na listagem seguinte:

```
{"BALLOT": {"Deputado": {"11": "Deputado␣1", "22": "Deputado␣2",
                         "33": "Deputado␣3", "44": "Deputado␣4",
                         "0": "NULO"},
            "Senador": {"11": "Senador␣1", "22": "Senador␣2",
                        "0": "NULO"},
            "Governador": {"11": "Governador␣1", "22": "Governador␣2",
                           "33": "Governador␣3", "0": "NULO"},
            "Presidente": {"11": "Presidente␣1", "22": "Presidente␣2",
                           "0": "NULO"}},
```

```
"GROUP": {"TYPE": "ECGROUP-NISTP256"},
"ALPHA": 58048582108789155680297047807400040692181998655678423207561376800872479019069722475233698970056530808476513973847189910931639598046635568082420618873715640,
"BETA": 60831688908398792965557824183569427450215313743756625836677160613580205791417861831436358772467216148523622675510518920440793771802429251487219213392440015,
"HASH": "SHA256-1",
"GUARDIAN": {"KEY": 1240200885107417021354474160070992782002420879389160650872662476221744167032527,
"SCHEME": "PAILLIER"},
"SIGNATURE": {"KEY": 6632244624557064369063227853304673444267349548781548615058382792020788245208257889214011861390466315675060965521862349520722031822494870763055883756356140 9,
"SCHEME": "ECDSA-NISTP256"}}
```

Código 5.2: Descrição da eleição em formato JSON

Essa listagem corresponde à descrição pública da eleição. Note-se que a descrição previamente mencionada da cédula eleitoral é agregada diretamente a essa descrição. Por sua vez, os geradores eleitorais, aqui pontos de uma curva elíptica, são normalizados para coordenadas afins e então representados como codificações inteiras de seus respectivos pontos. Já os parâmetros de chave associados aos guardiões e à assinatura representam respectivamente a chave pública usada diretamente para cifrar os códigos de segurança e a chave pública de verificação de assinaturas dos artefatos produzidos pela máquina de votação.

Acerca dos geradores eleitorais α e β, diferentes metodologias podem ser empregadas para sua produção, sempre observando a importância da necessidade de independência entre ambos (c.f. Definição 4.2.2). A implementação aqui exposta, por exemplo, e tendo em vista o caráter ilustrativo da mesma, utiliza-se simplesmente de duas leituras independentes do gerador padrão não bloqueante do sistema utilizado (e.g. */dev/urandom* em sistemas GNU/Linux). A primeira destas leituras é usada para produzir α, e a segunda destas para produzir β. Uma outra possibilidade igualmente plausível dá-se pela definição um dos geradores (α ou β) como um gerador fixo do grupo algébrico escolhido, sendo apenas o outro gerador produzido por algum método que se utilize de aleatoriedade.

Ressalta-se também que a origem de tal aleatoriedade é outro aspecto decisório crucial, sendo importante que haja uma metodologia bem definida e publicamente replicável para sua obtenção de modo a se evitar vícios ou suspeitas sobre o processo de produção dos geradores.

Não obstante, tal descrição eleitoral conforme exposta é usada como fonte autoritativa dos parâmetros eleitorais utilizados e empregada em todas as etapas subsequentes do processo eleitoral. Por tal motivo, tal descrição, em um contexto real de aplicação do protocolo, deve ser oficializada via assinaturas digitais e publicação em meios oficiais reconhecidos.

5.2.2 Votação

Estando devidamente definida e preparada a eleição, tem-se início a etapa de votação, na qual os eleitores operam uma máquina para geração de seus votos. Tal máquina inicialmente encontra-se em um estado de inativação, conforme representa a Figura 5.2:

Figura 5.2: Máquina de votação inativa

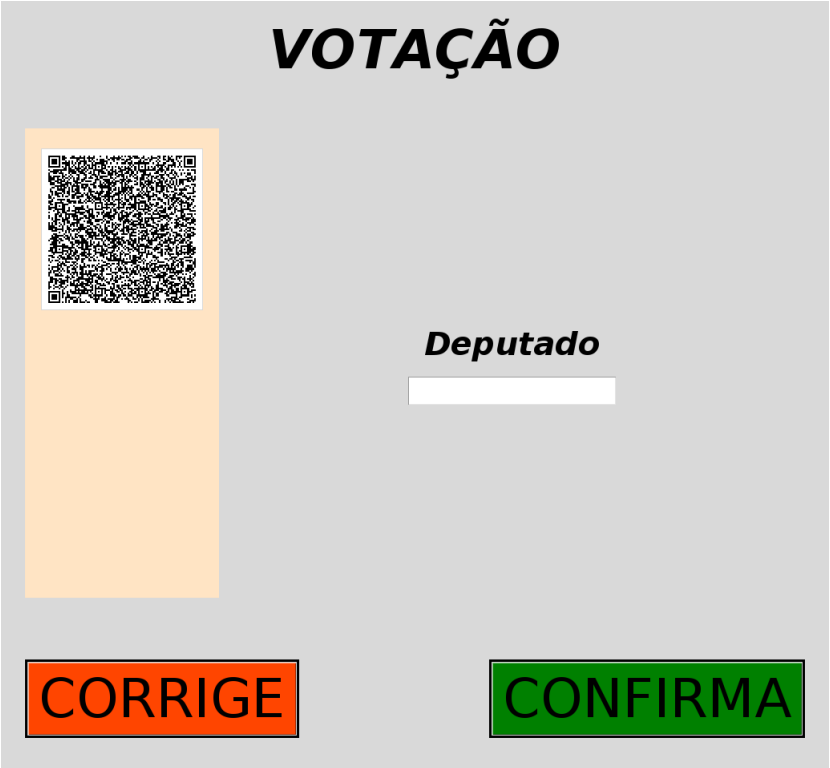
Fonte: Elaborado pelo autor.

Conforme previamente descrito, o eleitor deve se utilizar de algum tipo de interface,

como um botão físico, para ativar a máquina de votação e dar início ao processo de geração de seu voto. A Figura 5.2 demonstra este aspecto de maneira simplificada. É importante lembrar que tal ativação deve ser feita de modo que a máquina seja incapaz de determinar a identidade do eleitor.

Imediatamente após ativar a máquina e necessariamente antes de informar qualquer uma de suas opções de voto, a máquina deve gerar e imprimir o código de segurança cifrado que será usado para a geração da cédula do eleitor. Este aspecto é mostrado na Figura 5.3:

Figura 5.3: Máquina de votação ativa com código de segurança impresso



Fonte: Elaborado pelo autor.

A Figura 5.3 mostra três porções distintas da máquina de votação: à esquerda encontra-se uma visão da impressora conforme o eleitor a vê; à direita tem-se a tela da máquina com informações legíveis pelo eleitor; e abaixo os botões de correção e confirmação utilizados pelo próprio eleitor para corrigir informações ou prosseguir com suas escolhas. Adicionalmente, a máquina física de votação deve prover alguma forma de teclado para o eleitor teclar suas opções, o qual não é aqui mostrado por simplicidade da interface.

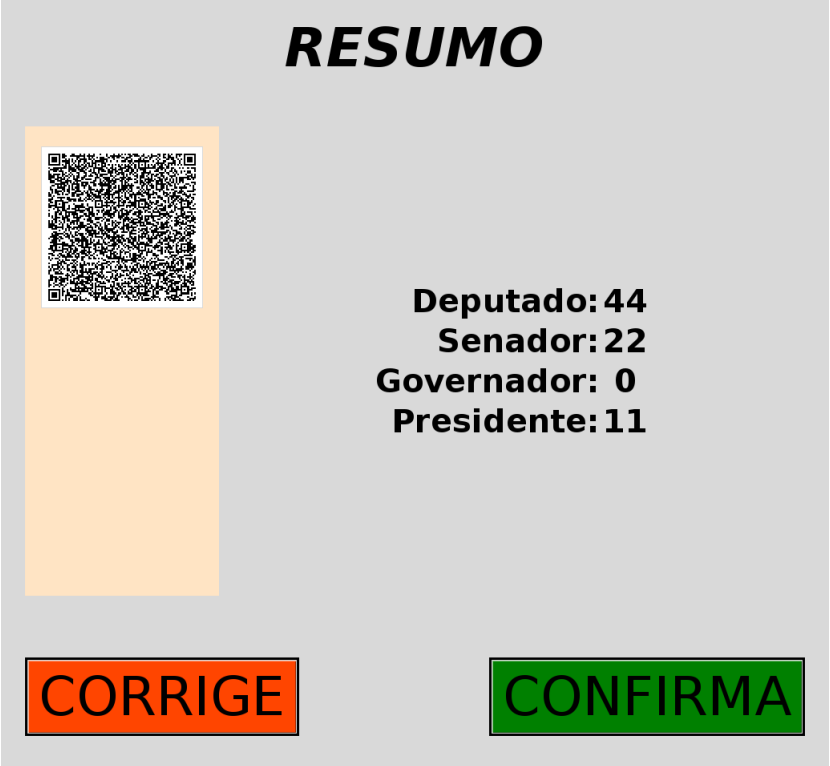
Observe-se que a máquina de pronto se compromete com um código de segurança cifrado via sua impressão em papel, o que ocorre antes que qualquer informação referente ao voto seja fornecida pelo eleitor. Isto evita que a máquina tenha qualquer grau de liberdade com relação ao código gerado em função das escolhas do eleitor.

Neste ponto a interface se encontra ativa e pronta para receber as opções de voto,

as quais são informadas pelo eleitor sequencialmente em conformidade com a ordenação estabelecida e respectivas opções disponíveis na configuração de cédula oficializada e publicada via descrição eleitoral. No exemplo descrito, o eleitor escolhe sequencialmente suas opções de voto para *Deputado, Senador, Governador* e então *Presidente*. Em cada qual, após informar o número referente a sua opção, o eleitor pressiona o botão *CONFIRMA* para prosseguir à sua escolha para o cargo seguinte. Caso confirme uma opção inválida ou nenhuma opção, o voto é considerado como nulo. Caso esteja insatisfeito com a opção informada, o eleitor pressiona o botão *CORRIGE* e informa uma nova opção.

Tendo feito suas escolhas para todos os cargos disponíveis, uma tela final com o resumo das escolhas feitas é mostrada ao eleitor para confirmação final de seu voto, conforme demonstra a Figura 5.4:

Figura 5.4: Resumo das escolhas e confirmação do voto



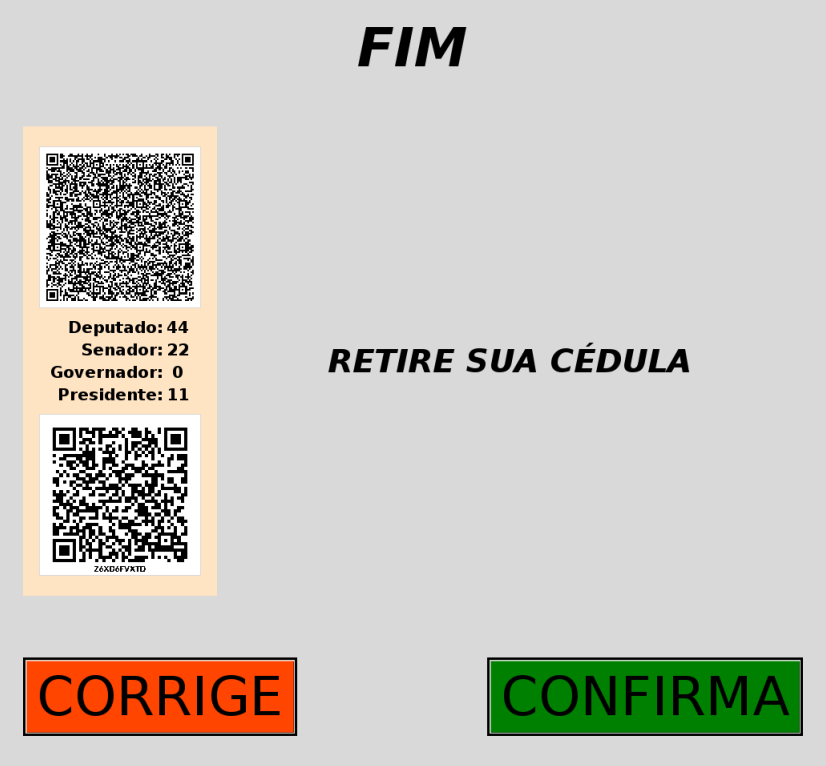
Fonte: Elaborado pelo autor.

No exemplo mostrado na Figura 5.4, o eleitor escolheu a opção 44 para *Deputado*, a opção 22 para *Senador*, anulou seu voto para *Governador* e escolheu a opção 11 para *Presidente*. É importante observar que durante toda a etapa de votação nenhum tipo de atividade ocorreu por parte da impressora: apenas o código de segurança cifrado se encontra impresso até o momento, tendo sido produzido anteriormente à manifestação das opções pelo eleitor.

Neste ponto o eleitor pode reiniciar todo o processo de escolha de suas opções apertando o botão *CORRIGE* ou confirmar suas escolhas já feitas e mostradas na tela de

resumo apertando o botão *CONFIRMA*. Caso opte por reiniciar o processo de escolha, a votação é simplesmente reiniciada e o eleitor observa o sistema no mesmo estado já mostrado na Figura 5.3. Observe-se, portanto, que o código de segurança da cédula se mantém o mesmo já produzido e impresso em papel, isto é, o reinício do processo de escolha de opções não altera o código de segurança com o qual a máquina já se comprometeu. Caso opte o eleitor pela confirmação de seu voto, a máquina então segue para o estado mostrado na Figura 5.5:

Figura 5.5: Voto confirmado



Fonte: Elaborado pelo autor.

Conforme expõe a Figura 5.5, a confirmação do voto produz a impressão em papel do restante dos componentes da cédula, a saber, o voto em formato impresso, legível e verificável pelo eleitor, e o recibo referente a um compromisso de Pedersen assinado cuja assinatura pode ser verificada como válida pela chave de verificação publicada na descrição eleitoral. Observe-se, portanto, que a confirmação do voto leva a máquina a se comprometer com uma cédula (e, v, c) completa.

O recibo c, conforme exposição prévia no Capítulo 4, é um compromisso de Pedersen produzido com base nos geradores independentes α e β, previamente definidos e publicados, em conjunto com dois outros valores aplicados como expoentes: um inteiro aleatório r temporariamente mantido em memória na máquina de votação e expresso em forma cifrada como o argumento e impresso; e um valor inteiro $H(v)$ correspondente ao resultado da aplicação do hash de votos H sobre uma representação digital bem definida

do voto em texto pleno expresso em v.

Para fins de ilustração da importância da necessidade de completa definição da representação digital dos votos de entrada, seja considerada a listagem colocada a seguir, que mostra possíveis representações digitais usadas como entrada e as respectivas saídas da aplicação do hash de votos sobre tais representações:

```
V='[{"Deputado": D}, {"Senador": S}, {"Governador": G}, {"Presidente": P}]'

V1='[{"Deputado": 44}, {"Senador": 22}, {"Governador": 0}, {"Presidente": 11}]'
H1=0x4624669EBCC78BB2C196D7D073C3F2A5CE68112EE6E1F7894C2245C49660FADA

V2='[{"Deputado": 44}, {"Senador": 22}, {"Governador":0}, {"Presidente": 11}]'
H2=0x7FA1AE329F1CF63A67A8AD056684580331623FBC80379A5950B81001983FCC1D
```

Código 5.3: Representações de votos e saídas de hash

Tendo-se em vista tal listagem, na presente implementação o hash de votos definido é computado sobre uma cadeia de caracteres (*string*) cujo conteúdo é formatado em acordo com a representação digital expressa em V, na qual, obviamente, os valores específicos da escolha particular expressa no voto substituem os campos *D, S, G* e *P*. Um exemplo de voto corretamente formatado é dado por $V1$, cuja saída do hash de votos é dada por $H1$.

Note-se, no entanto, que, dada a sensibilidade esperada do hash de votos empregado, é de crucial importância que a representação digital do voto de entrada esteja plenamente definida em todos os seus aspectos, inclusive no referente à presença e posição de caracteres em branco (*whitespace*). Como exemplo, note-se que a representação $V2$ mostrada na supracitada listagem representa semanticamente o mesmo voto expresso em $V1$, mas, por conta da omissão de um único espaço em branco na representação da opção para *"Governador"*, apresenta uma saída $H2$ completamente díspare da saída $H1$.

Retornando à discussão sobre o recibo c produzido pela implementação, este corresponde a um ponto normalizado da curva elíptica NIST P-256 expresso sob uma formatação compactada correspondente à coordenada X do ponto prefixada dos bits 10, caso sua coordenada Y seja positiva, ou dos bit 11, caso sua coordenada Y seja negativa, explorando a simetria da curva. Posfixada a tal formatação compacta do ponto encontra-se a assinatura ECDSA associada ao recibo, composta por 513 bits referentes à concatenação $1|r|s$, em que o par (r, s) corresponde à assinatura ECDSA produzida sobre a supracitada representação compactada do ponto da curva elíptica.

A implementação atribui ainda a cada recibo um componente adicional referente a um rastreador de recibo, evidenciado na Figura 5.6:

Figura 5.6: Recibo c com rastreador

Fonte: Elaborado pelo autor.

O propósito desse rastreador é prover a cada recibo um identificador simples que facilite ao eleitor, como parte do processo de auditoria individual, a verificação da inclusão de seu voto na totalização eleitoral produzida posteriormente. Seu uso prático ficará mais claro quando da descrição do procedimento de auditoria individual conduzido pelo eleitor no contexto da implementação.

Tal rastreador é produzido por uma codificação específica em base 16 dos primeiros 40 bits do resultado da aplicação da função de hash *SHA256* sobre o recibo assinado, resultando em uma cadeia de 10 caracteres. O uso de um conjunto de caracteres distinto do tradicionalmente utilizado para representação de números em base 16 motiva-se principalmente para fins de evitar possíveis formações de palavras legíveis em português (e.g. *FACADA1350*), e os caracteres selecionados visam mitigar possibilidades de confusão entre caracteres a depender da fonte utilizada (e.g. entre "1", "l" e "I", ou entre "8" e "B").

Sob as hipóteses de uniformidade na geração do código de segurança r que compõe o recibo e de que o hash *SHA256* se comporta como um oráculo aleatório, todos os possíveis rastreadores com 10 dígitos são equiprováveis. Isto implica que a probabilidade de colisão de rastreadores, por exemplo, dentro de um universo de 1024 eleitores é dada por $p \approx 0.00000048$. Tal probabilidade pode ser reduzida de maneira controlável adicionando-se mais bits da saída do hash à produção do rastreador, sob pena da inconveniência de torná-lo maior e exigir do eleitor mais digitações quando da execução do processo de auditoria individual.

No contexto de uma eleição real, a mesma metodologia pode ser seguida com semelhantes resultados, bastando, por exemplo, posfixar o número da seção eleitoral na qual o recibo foi produzido a cada rastreador gerado (e.g. produzindo um rastreador da forma **Z6XD6FVXTD-251**, em que o final *251* especifica uma seção eleitoral). De tal maneira, sob a hipótese de que uma seção eleitoral possua no máximo 1024 eleitores, as probabilidades de colisão de rastreadores se mostrarão iguais ou menores àquela aqui apresentada, sob a leve sobrecarga de digitação de alguns poucos caracteres a mais pelo eleitor que queira conduzir a auditoria individual.

Voltando ao fluxo eleitoral, estando a presente cédula completa e impressa, o eleitor deve apertar o botão *CONFIRMA* para finalizar a operação da máquina e da impressora e ter sua cédula liberada para depósito, retornando assim a máquina a seu estado de inativação conforme mostrado na Figura 5.2, pronta para ser reiniciada pelo operador seguinte.

Opcionalmente, e tendo em vista o compromisso da máquina de votação com o voto impresso expresso na cédula, o eleitor pode solicitar, caso assim deseje, o desafio da cédula formada. De modo a não impactar o fluxo normal de votação esperado ou causar qualquer sorte de confusão aos eleitores, a função de desafio se encontra oculta sob o botão *CORRIGE*, podendo ser solicitada neste ponto pelos eleitores que se interessarem no processo de desafio. Ao acionar o botão *CORRIGE* nesse estágio do processo de votação, o sistema reage imprimindo um parâmetro adicional, conforme ilustra a Figura 5.7:

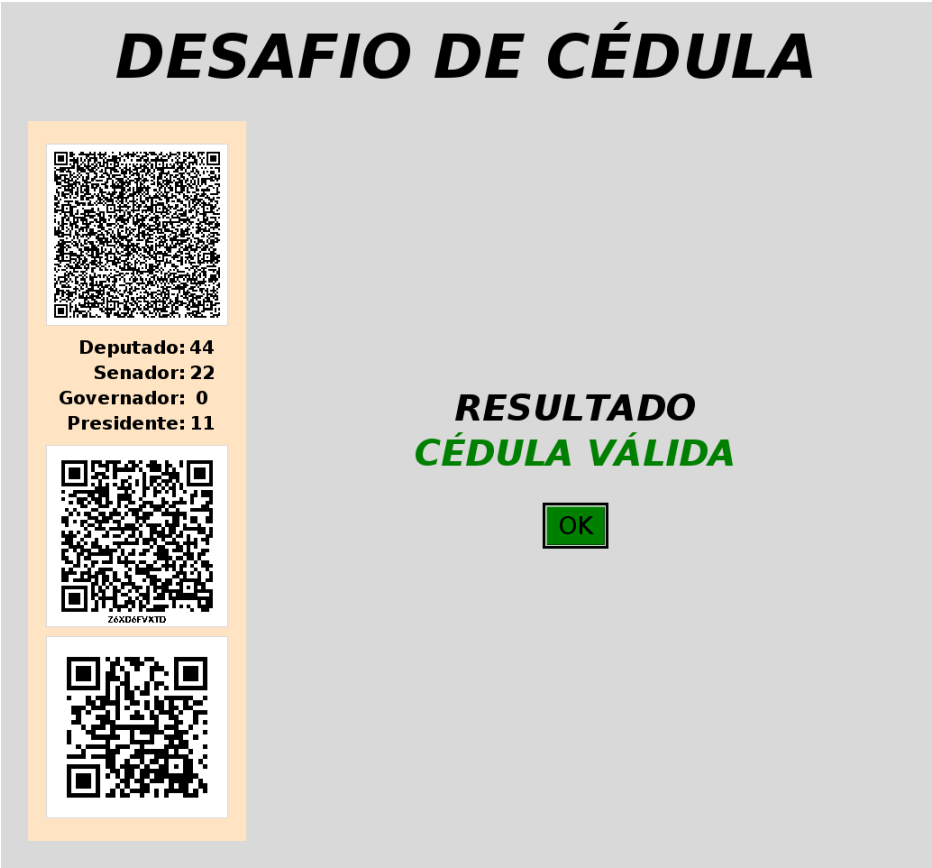
Figura 5.7: Requisição do desafio de cédula

Fonte: Elaborado pelo autor.

A Figura 5.7 mostra que o processo de desafio obriga a máquina a produzir a impressão de outro artefato na cédula, este correspondente ao código de segurança em texto pleno, de modo que a cédula passa agora a ser composta por uma tupla (e, v, c, r). Para uma cédula devidamente formada, o valor e corresponde à cifração de r e o recibo c corresponde a um compromisso de Pedersen dado por $\alpha^r \cdot \beta^{H(v)}$.

De modo a verificar tal correspondência, e tendo sido impressa e liberada a cédula (e, v, c, r) completa, o eleitor é solicitado a se direcionar a um mesário para proceder com o desafio da cédula. A máquina de votação é então desativada e retornada ao estado mostrado na Figura 5.2, pronta para ativação por seu próximo operador. Dirigindo-se o eleitor ao mesário e a ele fornecendo sua cédula, os parâmetros impressos são então escaneados e de tal modo informados a uma segunda máquina, distinta da de votação, e a supracitada correspondência entre os valores é verificada automaticamente. Do ponto de vista do eleitor, o processo de desafio se mostra de maneira consideravelmente simples, conforme expõe a Figura 5.8:

Figura 5.8: Resultado do desafio sobre cédula válida



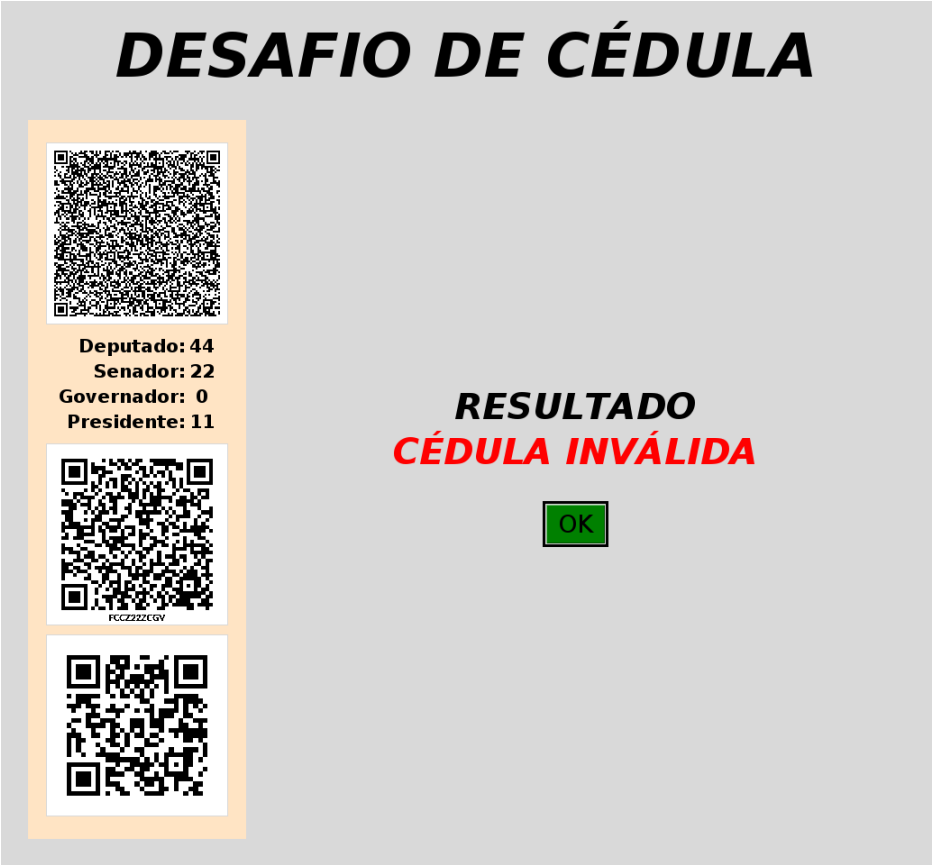
Fonte: Elaborado pelo autor.

Conforme mostra a Figura 5.8, apenas é mostrado ao eleitor o resultado do processo de desafio, que no caso da implementação foi simplificado unicamente para verificação da assinatura e análise da relação matemática entre os componentes da cédula. Acredita-se não fazer muito sentido, no contexto do fluxo de votação, se complicar a interface ou o processo de desafio em uma tentativa de explicar ao eleitor a matemática envolvida neste. Não obstante, o eleitor deve ter total direito de produzir cópia (e.g. via foto ou uso de aplicativo específico) de sua cédula desafiada para conduzir, por conta própria ou com auxílio de terceiros de sua confiança, o processo de desafio. Idealmente, as autoridades

eleitorais devem publicar posteriormente todas as cédulas desafiadas.

Obviamente, caso a cédula produzida pela máquina não seja condizente com sua formatação matemática esperada, um processo de desafio honesto produzirá um resultado claro de falha, conforme mostra a Figura 5.9:

Figura 5.9: Resultado de desafio sobre cédula inválida



Fonte: Elaborado pelo autor.

Após o processo de desafio, o eleitor deve ser reconduzido à máquina de votação para proceder com a geração de nova cédula, seguindo os mesmos passos acima descritos. Satisfeito com suas escolhas e os desafio produzidos, o eleitor então confirma seu voto e deposita os componentes de sua cédula em suas respectivas urnas, tendo seu recibo escaneado e levando-o consigo. Este processo é repetido por todos os eleitores até que a eleição seja finalizada.

5.2.3 Totalização

Finda a eleição, prossegue-se à etapa de totalização dos resultados. Em uma eleição física real, nessa etapa as urnas são abertas, possivelmente após um procedimento de embaralhamento físico de seus conteúdos (e.g. via movimentação física aleatória de cada urna) e os componentes individuais de cada qual são escaneados com o auxílio de algum dispositivo eletrônico. São assim produzidos um conjunto $\mathbb{E}$ de votos cifrados e um conjunto $\mathbb{V}$ de votos depositados. Adicionalmente, há também um conjunto $\mathbb{C}$ de recibos escaneados, este continuamente construído durante a própria etapa de votação conforme os recibos dos eleitores vão sendo registrados.

Na simulação provida pela implementação construída, conforme já mencionado, alguns desses aspectos físicos são simplificados: as urnas são representadas por diretórios que agregam os respectivos componentes (códigos, votos ou recibos) e o embaralhamento físico, cujo propósito é impossibilitar um mapeamento entre eleitores e respectivos artefatos depositados, é substituído por uma nomeação aleatória dos arquivos que representam os componentes. Como exemplo, considere-se a Figura 5.10, que ilustra a urna de códigos cifrados, no exemplo composta por registros de dez eleitores:

Figura 5.10: Urna de códigos de segurança cifrados

Fonte: Elaborado pelo autor.

A Figura 5.10 explicita ainda a decisão de implementação de que os códigos de segurança cifrados sejam registrados da forma mais próxima possível a uma aplicação real do protocolo, isto é, como imagens de códigos QR, tendo em vista que tais imagens serão o que de fato compõe o conteúdo de uma urna utilizada em um contexto real de aplicação.

Em termos da urna de votos correspondente, esta é ilustrada pela Figura 5.11:

Figura 5.11: Urna de votos


```
[{"Deputado": 22}, {"Senador": 22}, {"Governador": 22}, {"Presidente": 22}]
[{"Deputado": 44}, {"Senador": 22}, {"Governador": 0}, {"Presidente": 0}]
[{"Deputado": 44}, {"Senador": 11}, {"Governador": 11}, {"Presidente": 22}]
[{"Deputado": 33}, {"Senador": 11}, {"Governador": 11}, {"Presidente": 11}]
[{"Deputado": 0}, {"Senador": 11}, {"Governador": 11}, {"Presidente": 22}]
[{"Deputado": 11}, {"Senador": 11}, {"Governador": 11}, {"Presidente": 22}]
[{"Deputado": 0}, {"Senador": 0}, {"Governador": 0}, {"Presidente": 11}]
[{"Deputado": 0}, {"Senador": 22}, {"Governador": 11}, {"Presidente": 22}]
[{"Deputado": 44}, {"Senador": 11}, {"Governador": 33}, {"Presidente": 22}]
[{"Deputado": 0}, {"Senador": 11}, {"Governador": 33}, {"Presidente": 11}]
```


Fonte: Elaborado pelo autor.

Conforme demonstra a Figura 5.11, optou-se aqui pelo registro dos votos sob uma formatação JSON conforme já demonstrada previamente. Apesar de distinta da formatação mostrada pelo protocolo padrão, tal forma de registro facilita o tratamento dos votos na simulação provida pela implementação e evita possíveis complexidades associadas a metodologias OCR, que não compõem o foco do presente trabalho.

Finalmente, em termos do registro de recibos válidos produzidos, este é mostrado na Figura 5.12:

Figura 5.12: Registro de recibos

Fonte: Elaborado pelo autor.

De modo semelhante ao caso da urna de códigos, optou-se aqui pelo registro dos recibos também sob uma formatação de imagem. O intuito de tal decisão é meramente o de facilitar a visualização de recibos e reduzir a necessidade de processamentos posteriores, evitando-se conversões entre representações binárias dos recibos e respectivas imagens. Ressalta-se, entretanto, que os recibos podem facilmente ser registrados sob formas mais econômicas em termos de persistência (e.g. uma cadeia binária, um arquivo JSON simples), de tal modo reduzindo requisitos relacionados a armazenamento.

Um aspecto final importante referente aos recibos respalda-se em seus nomes de armazenamento: estes correspondem precisamente aos rastreadores presentes em cada

recibo, viabilizando de tal modo rápida forma de consulta quando necessária a título de execução de procedimentos de auditoria individual por parte dos eleitores. Tal metodologia pode ser facilmente traduzida para um contexto de uso de bancos de dados relacional, bastando a uma tabela de recibos a adição de uma coluna que relacione seu respectivo rastreador.

Em termos da totalização propriamente dita, o procedimento, conforme indica seu próprio nome, se inicia com a agregação dos votos presentes no conjunto $\mathbb{V}$, de tal forma produzindo somatórios que indicam os totais de votos obtidos para cada opção disponível no pleito. Uma tal totalização é mostrada na Figura 5.13:

Figura 5.13: Quadro de totalização dos votos

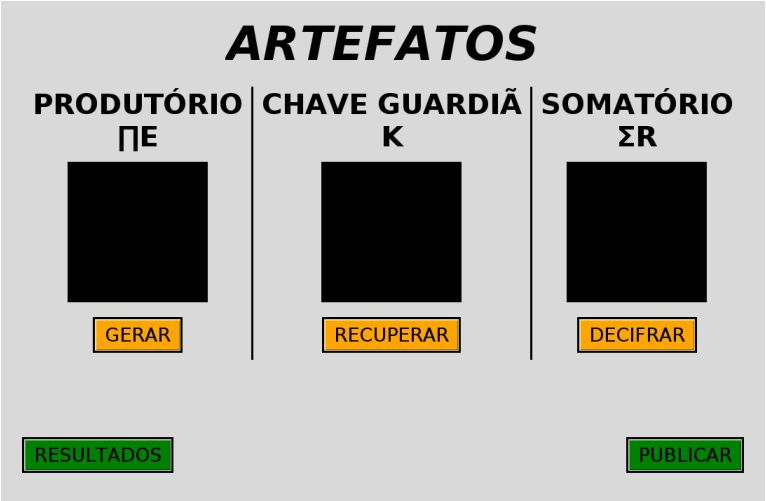
Fonte: Elaborado pelo autor.

A Figura 5.13 mostra um quadro-resumo que explicita os resultados da totalização em termos de cargos e suas respectivas opções disponíveis, ordenados em conformidade com as especificações estabelecidas pela descrição da cédula. Este quadro de resultados resume aquilo mais comumente compreendido pelo público geral como os resultados propriamente ditos da eleição, constituindo a informação em que provavelmente a maior parte dos eleitores estará de fato interessada.

Entretanto, do ponto de vista do protocolo e das garantias por este providas, os resultados advindos da totalização eleitoral constituem apenas uma parte dos resultados. Conforme já mostrado no Capítulo 4, o processo de totalização, para fins de produção de um lastro matemático e criptográfico de auditoria e verificação de integridade, requer ainda a produção do somatório em texto pleno r^{+} de todos os códigos armazenados

cifrados em $\mathbb{E}$ e a respectiva publicação de tal somatório r^+ e dos conjuntos $\mathbb{V}$ de votos e $\mathbb{C}$ de recibos. As funcionalidades de geração do somatório r^+ e respectiva publicação de resultados são acessadas pelo botão de rótulo *ARTEFATOS* mostrado na Figura 5.13, levando à interface exposta na Figura 5.14:

Figura 5.14: Geração de artefatos eleitorais



Fonte: Elaborado pelo autor.

Conforme observável pela Figura 5.14, a interface da implementação ilustra, em ordem, os artefatos necessários para a produção do somatório r^+ de códigos: primeiro é necessário produzir o produtório de códigos cifrados e^*, representado no quadro mais à esquerda. Em seguida, os guardiões devem se reunir para recuperar a chave guardiã K, representada ao centro, a qual corresponde à chave privada de Paillier que será usada para decifrar o supracitado produtório de códigos. Finalmente, dada a natureza homomórfica da cifra de Paillier usada para cifrar cada código de segurança, a decifração desse produtório produz o somatório em texto pleno r^+ dos códigos de segurança utilizados, representado no quadro mais à direita. A produção desses artefatos é ilustrada na Figura 5.15, que também demonstra a decisão de projeto de representar os artefatos sob a forma de códigos QR de modo a torná-los de mais fácil visualização:

Figura 5.15: Artefatos eleitorais gerados

Fonte: Elaborado pelo autor.

É importante também observar que, em uma implementação real, a chave guardiã privada K não deve ser obtida da forma replicável como aqui mostrada, sob risco de possibilitar sua utilização para fins de decifração dos códigos de segurança individuais e consequente violação do sigilo dos votos dos eleitores. Na prática, as parcelas dos guardiões podem ser fornecidas a uma controladora de baixo custo, a qual é posteriormente alimentada com o produtório dos códigos cifrados, retornado puramente, e de maneira verificável, apenas o somatório decifrado dos códigos. A controladora posteriormente pode ser fisicamente destruída de modo a se evitar a possibilidade de recuperação da chave privada decifrada.

Estando todos os artefatos devidamente gerados, o botão *PUBLICAR* é então usado para gerar uma descrição dos resultados em formato JSON. Uma tal descrição é mostrada na listagem seguinte:

```
{"TALLY": {"Deputado": {"11": ["Deputado␣1", 1], "22": ["Deputado␣2", 1],
                        "33": ["Deputado␣3", 1], "44": ["Deputado␣4", 3],
                        "0": ["NULO", 4]},
           "Senador": {"11": ["Senador␣1", 6], "22": ["Senador␣2", 3],
                       "0": ["NULO", 1]},
           "Governador": {"11": ["Governador␣1", 5], "22": ["Governador␣2", 1],
                          "33": ["Governador␣3", 2], "0": ["NULO", 2]},
           "Presidente": {"11": ["Presidente␣1", 3], "22": ["Presidente␣2", 6],
                          "0": ["NULO", 1]}},
 "CODESUM": 28201044131882148823615841740417135623707337510646231894540483692045450246809
 "V": "votes",
 "C": "receipts"}
```

Código 5.4: Resultados publicados em formato JSON

Observe-se, portanto, que a descrição JSON dos resultados eleitorais, por simplicidade, segue uma formatação muito semelhante àquela da descrição da cédula eleitoral, mas aqui o valor associado a cada identificador de opção sob cada cargo é substituído por uma listagem do nome associado à opção junto ao montante de votos computado para esta.

Adicionalmente, a descrição dos resultados agrega o somatório dos códigos conforme obtido via decifração pelos guardiões eleitorais, e também indicadores de onde estão publicados o conjunto $\mathbb{V}$ de votos (V) e o conjunto $\mathbb{C}$ de recibos (C). Em uma implementação real, os valores associados a tais chaves possivelmente indicarão algum website específico. Na implementação, tais valores apenas apontam para diretórios locais relativos ao diretório de publicação da descrição JSON de resultados eleitorais.

Por fim, novamente ressalta-se que o conjunto $\mathbb{E}$ referente aos códigos de segurança cifrados não faz parte dos resultados nem é publicado, visto que tal procedimento possibilitaria sua perpetuação e potencial decifração dos códigos individuais futuramente. Assim, apenas o somatório r^+ é disponibilizado publicamente.

5.2.4 Auditoria

Estando os resultados eleitorais publicados, eleitores individuais e mesmo observadores independentes interessados no processo podem, em ambientes computacionais e condições de sua escolha e controle, se utilizar dos supracitados artefatos eleitorais produzidos para conduzir verificações independentes acerca da integridade dos resultados eleitorais informados.

Em termos de um eleitor que participou da votação, este pode se utilizar do procedimento de auditoria individual para averiguar que o recibo que possui em mãos de fato se encontra registrado e publicado como elemento do conjunto $\mathbb{C}$ de recibos oficiais. No sistema implementado, tal aspecto é facilitado pela sequência de rastreio impressa logo abaixo do código QR referente ao recibo eleitoral. Para conduzir tal verificação, o eleitor se utiliza de uma interface semelhante à mostrada na Figura 5.16:

Figura 5.16: Interface de auditoria individual

Fonte: Elaborado pelo autor.

Conforme ilustra a Figura 5.16, basta ao eleitor interessado na verificação da inclusão de seu voto na totalização informar o número de rastreio constante em sua cédula, o que, em termos da presente implementação, consiste em digitar apenas 10 dígitos simples (i.e. números e letras, sem diferenciação entre minúsculas e maiúsculas). A Figura 5.17 mostra o resultado esperado da busca por um recibo devidamente gerado no contexto da eleição:

Figura 5.17: Recibo eleitoral encontrado

Fonte: Elaborado pelo autor.

Ao verificar que seu recibo se encontra devidamente registrado, um eleitor obtém garantias de que, a menos de probabilidades negligenciáveis, seu voto se encontra totali-

zado conforme pretendido nos resultados eleitorais finais.

Por outro lado, caso seu recibo não seja encontrado, o eleitor pode contestar os resultados com base na assinatura digital embutida. Em tal caso, o eleitor possui em mãos uma prova publicamente verificável de que um recibo devidamente assinado não se encontra registrado junto aos resultados eleitorais publicados.

Em termos da auditoria universal, esta pode ser conduzida por qualquer observador interessado, sem necessidade de informações adicionais para além daquelas oficialmente publicadas. O procedimento se inicia com a reprodução dos resultados eleitorais, conforme demonstra a Figura 5.18:

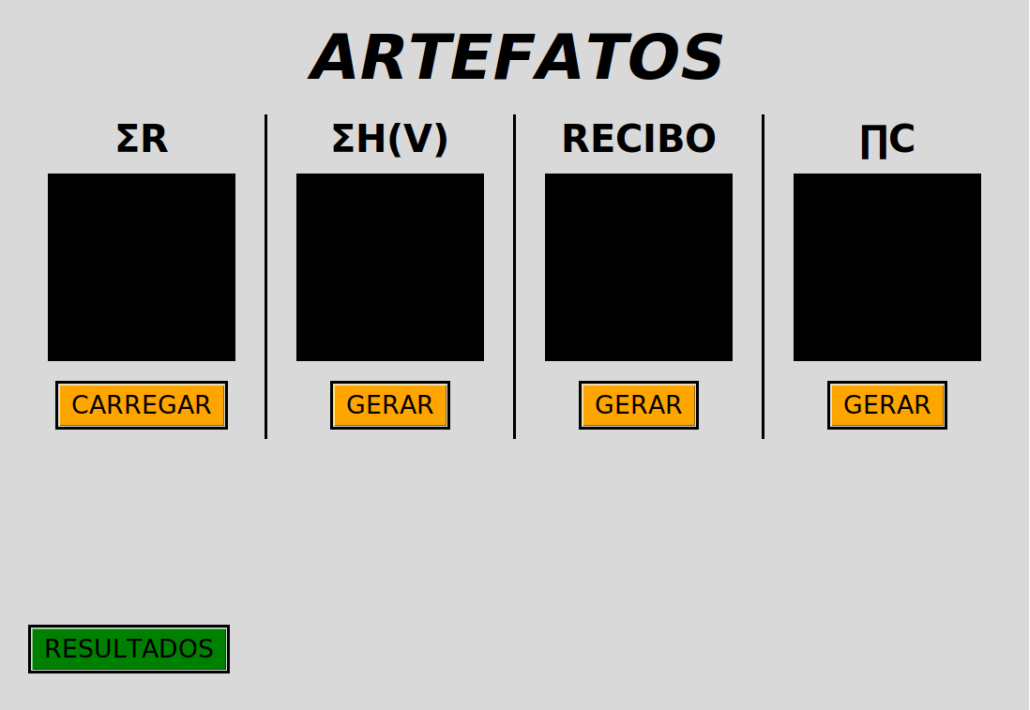
Figura 5.18: Reprodução dos resultados eleitorais

Fonte: Elaborado pelo autor.

Em se tratando de uma reprodução honesta de um resultado íntegro e devidamente publicado, note-se que os resultados eleitorais expostos pela Figura 5.18 são idênticos aos revelados pela Figura 5.13. A diferença básica entre ambos reside nos conjuntos de votos usados para produção de cada qual: enquanto aqueles da Figura 5.13 são produzidos pelas autoridades eleitorais (ou representantes destas) por contagem direta dos artefatos depositados nas urnas físicas, os resultados expostos na Figura 5.18 são produzidos por qualquer indivíduo realizando uma recontagem dos votos com base no conjunto $\mathbb{V}$ publicado, devendo ambos ser de fato idênticos por premissa básica de integridade.

De forma semelhante ao procedimento de totalização, a auditoria universal não é finalizada com a mera recontagem dos votos, devendo o auditor também verificar a consistência dos resultados com base nos demais artefatos de totalização publicados. O auditor universal, portanto, deve se utilizar dos valores e conjuntos publicados para verificar a equação universal de integridade eleitoral exposta no Capítulo 4. Tais artefatos são expostos na Figura 5.19:

Figura 5.19: Artefatos de auditoria universal

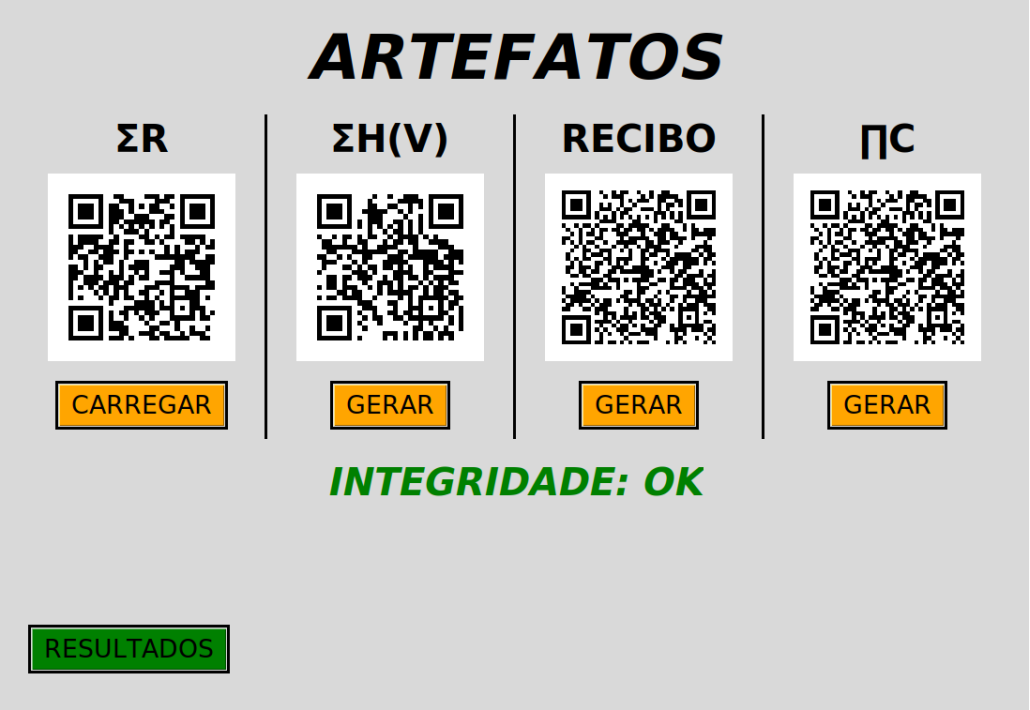


Fonte: Elaborado pelo autor.

Conforme mostra a Figura 5.19, o auditor deve obter o somatório r^+ (primeiro quadro) presente na descrição dos resultados publicados e se utilizar dos votos registrados conforme constam no conjunto $\mathbb{V}$ também publicado para produzir o somatório $h^+ = \sum H(v)$ referente à soma das saídas do hash de votos para cada $v \in \mathbb{V}$ (segundo quadro). De posse de tais valores, o auditor pode então computar o recibo de integridade eleitoral (terceiro quadro), o qual corresponde a um compromisso de Pedersen utilizando-se dos valores r^+ e h^+ produzidos. Conforme demonstrado pelo Teorema 4.2.16, o recibo de integridade eleitoral produzido deve ser idêntico ao produtório de todos os recibos $c \in \mathbb{C}$ publicados (quarto quadro).

A Figura 5.20 mostra os resultados da geração de tais artefatos para uma eleição íntegra:

Figura 5.20: Artefatos de auditoria universal gerados



Fonte: Elaborado pelo autor.

Assim, conforme demonstra a Figura 5.20, observe-se que o recibo de integridade eleitoral e o produtório de recibos (os dois últimos quadros) apresentam o mesmo valor, na implementação expressos sob a forma de códigos QR para fins de facilidade visual. Conforme expresso pelo Teorema 4.2.18, sob a hipótese da dificuldade de resolução do DLP, é portanto negligenciável a probabilidade de que os resultados publicados não sejam íntegros.

5.3 Resultados

A implementação mostrada provê uma ilustração simplificada da aplicação do protocolo em um contexto prático. Tal implementação foi exercitada em um sistema Slackware GNU/Linux (Kernel 5.4.80) executado por um processador Celeron N4000 de 1.10GHz para obtenção de alguns aspectos quantitativos associados à execução do protocolo.

Em termos de desempenho, como haveria de se esperar, a principal sobrecarga da implementação deve-se ao uso do criptossistema de Paillier, o qual requer uma quantidade de bits considerável dado seu aspecto de segurança semelhante ao do RSA (fatoração). Ainda assim, a geração de uma cédula completa utilizando-se de uma chave Paillier de

2048 bits com um algoritmo simples simulando um eleitor que escolhe suas opções aleatoriamente leva em média aproximados 385 milisegundos mesmo no processador utilizado, consideravelmente barato e de baixo consumo energético. Isto evidencia que, mediante uso de dispositivo de capacidades computacionais semelhantes, o maior tempo gasto durante a geração de uma cédula na prática fica por conta da operação da máquina pelo eleitor para manifestação de suas opções individuais. Para aqueles eleitores que não desejem fazer o desafio da cédula, a única demanda adicional de tempo fica por conta do processo de depósito dos componentes de cédula em suas respectivas urnas e o escaneamento do recibo, os quais requerem uma experimentação em caráter prático para levantamento de estimativas de tempo realistas.

Traçando-se um paralelo com a realidade brasileira e utilizando-se das estatísticas eleitorais de 2014 (que apresentam publicação muito mais detalhada se comparada com as publicações referentes a eleições mais recentes), tem-se um total de 142.822.046 eleitores distribuídos entre 451.501 seções eleitorais [47]. Majorando-se grosseiramente tal valor para 150.000.000 e supondo-se a manutenção do número de seções eleitorais, tem-se uma média de aproximadamente 333 eleitores por seção. Assim, para fins de ilustração, fazendo uma estimativa de 10 segundos para manipulação de cada registro individual em papel (i.e. valores e e v) para escaneá-los, cada seção com a média de eleitores apresentaria $2 \cdot 333 = 666$ registros escaneáveis (i.e. $|\mathbb{E}| + |\mathbb{V}|$) e demoraria aproximadas 2 horas para gerar seus resultados. Supondo-se eleições finalizadas às 17 horas, tal margem, mesmo considerando-se atrasos razoáveis associados à transmissão dos resultados a uma central de totalização, provê ainda possibilidade de declaração de resultados eleitorais no mesmo dia da votação, algo que se consolidou quase que como requisito na realidade eleitoral brasileira.

No entanto, dada a não-uniformidade da distribuição populacional brasileira junto à possibilidade de mobilidade entre seções a pedido dos eleitores em caráter prévio, a hipótese da igualdade distributiva de eleitores em cada seção dificilmente se mostraria verdadeira na prática, de modo que é esperado que os números de eleitores variem entre seções distintas. Assim sendo, seções que apresentem um maior número de eleitores naturalmente demorariam mais para enviar seus resultados para totalização e publicação. Supondo-se, por exemplo, uma seção com 1024 eleitores, e sob as mesmas estimativas prévias de tempo, tal seção demoraria aproximadamente 5 horas e 40 minutos para produzir seus resultados. Cabe observar, entretanto, que a totalização apresenta um caráter paralelizável intrínseco que permite que os montantes parciais produzidos por cada seção sejam agregados de forma independente. Desse modo, os resultados produzidos por seções menos demandosas podem ser recebidos e agregados mais rapidamente, não havendo necessidade de aguardar a totalização parcial das seções com mais eleitores para se iniciar a construção do montante referente à totalização final em si. Tal aspecto de paralelização já é amplamente explorado na atual metodologia de totalização eleitoral brasileira. Ainda

assim, o prazo para tais seções demandosas ainda possibilitaria a publicação do resultado final no mesmo dia da votação.

No referente a estimativas de espaço de armazenamento necessário para publicação dos resultados, o registro digital de cada voto em texto pleno ocupa em média 74 bytes para a eleição ilustrada com 4 cargos (Deputado, Senador, Governador e Presidente). Cabe ressaltar, entretanto, dois aspectos importantes: primeiro, que é esperado que o registro do voto cresça em conformidade com o número de *cargos* disponíveis para escolha pelo eleitor (mas não com o número de opções por cargo); e segundo, que tal registro pode ser reduzido para fins de armazenamento utilizando-se das simples opções numéricas em caráter sequencial, ao custo de prejuízos à pronta legibilidade do registro. Assim, o valor de 74 bytes é fortemente intrínseco à eleição mostrada em conjunção com a escolha específica de formatação do registro dos votos. Em termos dos recibos, dada a escolha pelo registro em imagem, cada qual ocupa em média aproximados 1414 bytes. Utilizando-se de tais valores médios, uma seção eleitoral com 1024 eleitores exigiria um armazenamento de 1.523.712 bytes, ou aproximados 1,5 megabytes. Considerando-se uma eleição nacional com aproximados 160 milhões de eleitores, seria necessário um espaço de armazenamento aproximado de 221.8 gigabytes para os artefatos eleitorais produzidos, espaço este encontrado mesmo em dispositivos de comercializados para uso doméstico.

Os códigos de segurança individuais, conforme mencionado, não são publicados individualmente, sendo necessária apenas publicação do somatório destes ante a configuração de granularização escolhida. Cada somatório, entretanto, corresponde a um expoente do grupo algébrico escolhido, que no contexto da implementação pode ser codificado em no máximo 256 bits. Assim, considerando-se granularização por seção eleitoral, e utilizando-se da estimativa de 451.501 seções eleitorais [47, p.15], seria necessário um espaço aproximado de 13 megabytes para armazenamento de todos os somatórios de códigos de segurança na granularidade de seção eleitoral.

Tais resultados são fortes indicativos de que o protocolo descrito apresenta potencial de aplicação prática em eleições presenciais, conferindo independência de software e capacidades E2E-V a pleitos realizados sem incorrer em significativas sobrecargas ao eleitor em termos de tempo demandado pelo fluxo normal de votação. Ademais, os requisitos de armazenamento exigidos pelos artefatos centrais ao protocolo se mostram de plena satisfação mesmo por aparato computacional simples e comercialmente disponível.

Capítulo 6

Conclusão

O desenvolvimento de protocolos eleitorais é certamente uma tarefa de complexo empreendimento. A dicotomia entre integridade e sigilo resulta no desenvolvimento de sistemas com propriedades não triviais e com grandes potenciais para introdução de erros e vulnerabilidades, em caráter acidental ou não. Tal aspecto se mostra particularmente forte no contexto de protocolos eleitorais criptográficos, em que se faz aparentemente necessário um compromisso entre integridade incondicional dos resultados ou sigilo incondicional dos votos, com um claro favoritismo pela incondicionalidade da integridade por parte dos protocolos mais proeminentes.

Nesse sentido, e em contraste, o presente trabalho descreve em detalhes um protocolo eleitoral que preza pela incondicionalidade do sigilo, sob a justificativa de que não deve haver um prazo, ainda que possivelmente indeterminado, para a privacidade do eleitor no referente a suas escolhas eleitorais. Em termos de integridade, esta é garantida computacionalmente por um período eleitoral bem-definido ante premissas criptográficas e esquemas bem-estabelecidos e amplamente estudados. Tal opção justifica-se por se acreditar ser de limitado impacto a descoberta acerca de como adulterar a integridade dos resultados eleitorais apenas após findo o período para quaisquer contestações legais referentes ao pleito em questão.

Ademais, o protocolo descrito respeita o princípio da independência de software, provê capacidades de verificação de ponta-a-ponta (E2E-V) para votos individuais e permite aferição de integridade por parte de qualquer observador interessado no processo. Adicionalmente, é oferecido ainda suporte a cédulas extensas e potencial de granularização configurável para maior precisão de análises de integridade junto a possíveis correções quando da ocorrência de erros.

Em termos de aplicabilidade, o protocolo implica em leve sobrecarga ao eleitor no referente à finalização do processo de votação, que passa a consistir na inserção, em urnas físicas, de artefatos impressos referentes a um código de segurança e um voto em texto pleno, além do escaneamento do recibo provido ao eleitor. Estimativas realistas quanto ao tempo real demandado para tal processo provavelmente requerem aferição prática para maior precisão. Ademais, há certamente sobrecarga de tempo para aqueles eleitores que desafiarem suas cédulas produzidas, mas, tendo-se em vista que tal etapa é optativa, tal

sobrecarga não se mostra como um incômodo obrigatório. Por sua vez, no referente à operação da máquina de votação para produção de uma cédula, a implementação produzida demonstra que o uso de criptografia não incorre em sobrecargas perceptíveis ao eleitor. Com relação às necessidades de armazenamento, também a implementação provida leva a crer ser plausível esperar que estas não sejam superiores a capacidades providas por dispositivos comerciais acessíveis e facilmente encontrados à venda em dias atuais.

Assim, o protocolo proposto neste trabalho demonstra potencial de aplicação prática mesmo em contextos demandosos como eleições nacionais. Neste sentido, experiências como as do Prêt à Voter, Scantegrity e STAR-Vote mostram que governos e instituições democráticas estão dispostos a revisar seus sistemas eleitorais para acomodar as demandas contemporâneas com lastro nos desenvolvimentos científicos e tecnológicos recentes, embora haja ainda resistência e dificuldades de ordens diversas (e.g. históricas, culturais, políticas ou legais). No Brasil, há forte resistência, principalmente por parte das autoridades, em evoluir o ultrapassado projeto eleitoral eletrônico do país, ao mesmo tempo em que inexiste um diálogo saudável acerca de tais sistemas junto à academia.

No entanto, os resultados de projetos como o Você Fiscal [4] mostram que esforços de conscientização e engajamento junto à sociedade civil apresentam potencial no sentido de conseguir vencer a letargia evolutiva já característica das autoridades eleitorais brasileiras. Isto provê indícios de que projetos de melhoria de sistemas eleitorais não devem necessariamente esperar pela boa vontade das autoridades responsáveis, mas que podem ser conduzidos pela academia com base em ciência e divulgação junto à sociedade civil, de tal modo fomentando debate e conscientização devidamente lastreados acerca da necessidade de evolução dos sistemas empregados.

Assim sendo, o projeto de sistemas eleitorais aptos à aplicação em âmbitos demandosos se mostra como importante lastro na promoção de tais esforços de conscientização por prover alternativas plausíveis aos sistemas atualmente empregados por grandes democracias. E, seguindo-se a necessária contextualização de tais sistemas conforme preconizada por Dahl, a realidade eleitoral brasileira deve certamente considerar o histórico de quedas e recuperações democráticas do país, naquelas sempre havendo perseguição política a grupos específicos, e nestas sempre se identificando tentativas de indução política por parte de atores poderosos sobre membros da população. Sob tais prospectos, um projeto que preze pelo sigilo incondicional dos votos se mostra como valioso instrumento tanto para coibir influências indevidas sobre eleitores quanto para inviabilizar possíveis perseguições futuras motivadas por opções políticas manifestas em âmbitos eleitorais prévios.

6.1 Trabalhos Futuros

Conforme observado, o principal fator limitante associado ao protocolo em termos de desempenho encontra-se na aplicação do criptossistema de Paillier para cifração dos códigos de segurança, que acaba por exigir tamanhos de chave consideráveis ($N \approx 2048$, o que no caso de Paillier implica em trabalhar com $N^2 \approx 4096$). Isto leva a um crescimento considerável dos códigos QR produzidos, principalmente aqueles referentes ao artefato e de uma cédula, o que pode se mostrar inconveniente tanto em termos de espaço ocupado quanto no que se refere à velocidade de leitura, a depender do dispositivo empregado. Assim sendo, uma melhoria óbvia deixada para trabalhos futuros seria a substituição do criptossistema de Paillier por um esquema de cifração possivelmente mais leve em termos de exigências de tamanho do parâmetro de segurança.

Outro trabalho futuro de relevante consiste no desenvolvimento de protótipos funcionais de máquinas para as etapas de definição, votação e desafio, além de sistema de totalização e publicação, preferencialmente seguindo-se as ideias do STAR-Vote acerca de uso de equipamentos de aplicação comercial (*COTS*, do inglês *commercial off-the-shelf*) em lugar de "equipamento eleitoral especializado", mais caro e provavelmente sujeito a restrições proprietárias que podem minar a transparência do processo. O desenvolvimento de tal protótipo viabilizaria também a produção de estimativas de custo e, mediante aplicação em experimentação prática, também aferições mais precisas acerca das demandas de tempo nas várias etapas de execução do protocolo, principalmente aquelas associadas à participação do eleitor.

Apêndice A

Código

Lista-se aqui uma versão simplificada do código, em Python, referente às etapas consecutivas de execução do protocolo no contexto da implementação mostrada. É importante notar que em aplicações reais os artefatos privados, a exemplo das parcelas dos guardiões e chaves de assinatura, não devem ficar armazenados em memória para livre acesso, devendo ser devidamente armazenados e protegidos por seus guardiões responsáveis.

```
election = None # election specs
guardian_tuple = None # (public key, shares, share modulus)
sign_pubkey, sign_privkey = None # used to produce voting machine signatures

def SETUP(election_name,
          total_voters,
          total_guardians,
          minimum_guardians,
          ballot_file):

  # Initialize CREST
  init()

  guardian_tuple =  GENERATE_GUARDIAN_SHARES(2048,
                                             total_voters,
                                             total_guardians,
                                             minimum_guardians)

  sign_pubkey, sign_privkey, signature_size = GENERATE_SIGNING_KEY()

  election = GENERATE_ELECTION(election_name,
                               ballot_file,
                               guardian_tuple[0],
                               sign_pubkey,
                               sign_privkey)
```

Código A.1: Definição e preparação da eleição

```
def VOTE():
    challenge, ballot = GENERATE_BALLOT_TRIPLET(election)
    if challenge:
        CHALLENGE(ballot, election)
    else:
        DEPOSIT(ballot, election)
```

Código A.2: Votação

```
def TALLY():
    e_prod = GENERATE_CODE_PRODUCTORY(election)
    r_sum = GUARDIAN_DECRYPT(e_prod, guardian_tuple)
    tally = GENERATE_TALLY(election)
    GENERATE_RESULTS(election, tally, r_sum)
```

Código A.3: Totalização

```
def VOTER_AUDIT(receipt):
    CHECK_RECEIPT_IN_RESULTS(receipt, election)
```

Código A.4: Auditoria Individual

```
def UNIVERSAL_AUDIT():
    r_sum = LOAD_CODE_SUM(election)
    vh_sum = GENERATE_VOTE_HASH_SUM(election)
    c_prod = GENERATE_RECEIPT_PRODUCTORY(election)
    INTEGRITY_CHECK(election, r_sum, vh_sum, c_prod)
```

Código A.5: Auditoria Universal

Referências

[1] D. F. Aranha, C. P. L. Gouvêa, T. Markmann, R. S. Wahby, and K. Liao. RELIC is an Efficient LIbrary for Cryptography. https://github.com/relic-toolkit/relic, 2021.

[2] Diego F. Aranha, Pedro Y.S. Barbosa, Thiago N.C. Cardoso, Caio Lüders Araújo, and Paulo Matias. The return of software vulnerabilities in the brazilian voting machine. *Computers & Security*, 86:335–349, September 2019.

[3] Diego F. Aranha, Marcelo M. Karam, André de Miranda, and Felipe B. Scarel. Software vulnerabilities in the brazilian voting machine. In *Design, Development and Use of Secure Electronic Voting Systems*, pages 149–175. IGI Global, 2014.

[4] Diego F. Aranha, Helder Ribeiro, and André Luis Ogando Paraense. Crowdsourced integrity verification of election results. *Annals of Telecommunications*, 71(7-8):287–297, April 2016.

[5] Susan Bell, Josh Benaloh, Michael D. Byrne, Dana DeBeauvoir, Bryce Eakin, Gail Fisher, Philip Kortum, Neal McBurnett, Julian Montoya, Michelle Parker, Olivier Pereira, Philip B. Stark, Dan S. Wallach, and Michael Winn. STAR-Vote: A Secure, Transparent, Auditable and Reliable Voting System. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 375–403, 2017.

[6] Mihir Bellare, Joe Kilian, and Phillip Rogaway. The Security of the Cipher Block Chaining Message Authentication Code. *Journal of Computer and System Sciences*, 61(3):362–399, December 2000.

[7] Jonathan Ben-Nun, Niko Fahri, Morgan Llewellyn, Ben Riva, Alon Rosen, Amnon Ta-Shma, and Douglas Wikström. A new implementation of a dual (paper and cryptographic) voting system. In Manuel J. Kripp, Melanie Volkamer, and Rüdiger Grimm, editors, *5th International Conference on Electronic Voting 2012 (EVOTE2012)*, pages 315–329, Bonn, 2012. Gesellschaft für Informatik e.V.

[8] Daniel J. Bernstein and Tanja Lange. SafeCurves: choosing safe curves for elliptic-curve cryptography., 2017. Disponível em: https://safecurves.cr.yp.to (Acesso: 30.07.2021).

[9] Brasil. *Lei Nº 10.740*. Brasil, 2014. Dispõe sobre a introdução do registro digital do voto. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2003/

l10.740.htm
(Acesso: 20.05.2021).

[10] Joseph A. Calandrino, J. Alex Halderman, and Edward W. Felten. Machine-Assisted Election Auditing. In *2007 USENIX/ACCURATE Electronic Voting Technology Workshop (EVT 07)*, Boston, MA, August 2007. USENIX Association. Disponível em: http://electionmathematics.org/em-audits/US/Neff-ElectionConfidence.pdf
(Acesso: 27.06.2021).

[11] Richard T. Carback, David Chaum, Jeremy Clark, Alexsander Essex, Travis Mayberry, Stefan Popoveniuc, Ronald L. Rivest, Emily Shen, Alan T. Sherman, Poorvi L. Vora, John Wittrock, and Filip Zagórski. The Scantegrity Voting System and Its Use in the Takoma Park Elections. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 237–276, 2017.

[12] D. Chaum. Secret-ballot receipts: true voter-verifiable elections. *IEEE Security & Privacy Magazine*, 2(1):38–47, January 2004.

[13] Ronald Cramer, Rosario Gennaro, and Berry Schoenmakers. A Secure and Optimally Efficient Multi-Authority Election Scheme. In *Advances in Cryptology — EUROCRYPT '97*, pages 103–118. Springer Berlin Heidelberg, 1997.

[14] Édouard Cuvelier, Olivier Pereira, and Thomas Peters. Election verifiability or ballot privacy: Do we need to choose? In *Computer Security – ESORICS 2013*, pages 481–498. Springer Berlin Heidelberg, 2013.

[15] Robert Alan Dahl. *On Democracy*. Yale University Press, 1998.

[16] Quynh Dang. Recommendation for Applications Using Approved Hash Algorithms. Technical report, National Institute of Standards and Technology, 2012-08-24 2012. Disponível em: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=911479
(Acesso: 29.11.2021).

[17] Márcio Falcão and Fernanda Vivas. Hackers tiveram acesso a dados deste ano de servidores do TSE, apontam Polícia Federal e tribunal. *TV Globo*, 2020. Disponível em: https://g1.globo.com/politica/noticia/2020/11/19/policia-federal-aponta-que-hackers-acessaram-dados-de-2020-de-funcionarios-do-tse.ghtml
(Acesso: 20.05.2021).

[18] Oded Goldreich. *Foundations of Cryptography - Basic Tools*. Cambridge University Press, 2001. The Foundations of Cryptography - Volume 1.

[19] Daniel M. Gordon. A survey of fast exponentiation methods. *Journal of Algorithms*, 27(1):129–146, April 1998.

[20] Kenneth C. Johnson. Election Certification by Statistical Audit of Voter-Verified Paper Ballots. *SSRN Electronic Journal*, 2004.

[21] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography*. Chapman and Hall/CRC, 2 edition, 2014.

[22] Neal Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48(177):203–203, January 1987.

[23] Ralf Kusters, Tomasz Truderung, and Andreas Vogt. Clash Attacks on the Verifiability of E-Voting Systems. In *2012 IEEE Symposium on Security and Privacy*. IEEE, May 2012.

[24] Peter Hyun-Jeen Lee and Siamak F. Shahandashti. Theoretical Attacks on E2E Voting Systems. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 219–235, 2017.

[25] Mark Lindeman and Philip B. Stark. A Gentle Introduction to Risk-Limiting Audits. *IEEE Security & Privacy*, 10(5):42–49, September 2012.

[26] Eduardo Maretti. Auditoria no TSE 'mostra que PSDB é oposição patética e destrutiva', diz Damous. *Rede Brasil Atual (RBA)*, 2015. Disponível em: https://www.redebrasilatual.com.br/politica/2015/11/auditoria-no-tse-mostra-que-psdb-e-oposicao-patetica-e-destrutiva (Acesso: 16.05.2021).

[27] C. Andrew Neff. Election Confidence: A Comparison of Methodologies and Their Relative Effectiveness at Achieving It. *VoteHere Inc.*, 2003. Disponível em: http://electionmathematics.org/em-audits/US/Neff-ElectionConfidence.pdf (Acesso: 27.06.2021).

[28] NIST. Digital Signature Standard (DSS). Documento de Padronização, National Institute of Standards and Technology, July 2013. FIPS PUB 186-4.

[29] Pascal Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In *Advances in Cryptology — EUROCRYPT '99*, pages 223–238. Springer Berlin Heidelberg, 1999.

[30] Torben Pryds Pedersen. Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In *Advances in Cryptology — CRYPTO '91*, pages 129–140. Springer Berlin Heidelberg, 1992.

[31] Olivier Pereira. Internet Voting with Helios. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 277–308, 2017.

[32] Nicole Perlroth, Jeff Larson, and Scott Shane. N.S.A. Able to Foil Basic Safeguards of Privacy on Web. *The New York Times*, 2013. Disponível em: https://www.nytimes.com/2013/09/06/us/nsa-foils-much-internet-encryption.html (Acesso: 30.07.2021).

[33] Fabien A. P. Petitcolas. Kerckhoffs' principle. In *Encyclopedia of Cryptography and Security*, pages 675–675. Springer US, 2011.

[34] Daniel Fernando Pigatto, Natássya Barlate Floro da Silva, and Kalinka Regina Lucas Jaquie Castelo Branco. Performance Evaluation and Comparison of Algorithms for Elliptic Curve Cryptography with El-Gamal based on MIRACL and RELIC Libraries. *Journal of Applied Computing Research*, 1(2), February 2012.

[35] Caleb Pritchard. STAR-Vote collapses. *Austin Monitor*, 2017. Disponível em: https://www.austinmonitor.com/stories/2017/10/star-vote-collapses/ (Acesso: 05.08.2021).

[36] PSDB. Auditoria Especial no Sistema Eleitoral 2014. Relatório de Auditoria, Partido da Social Democracia Brasileira - PSDB, 2014. Disponível em: http://www.brunazo.eng.br/voto-e/arquivos/RelatorioAuditoriaEleicao2014-PSDB.pdf (Acesso: 16.05.2021).

[37] Python Software Foundation. *Extending and Embedding the Python Interpreter*. Python Software Foundation, 2021. Disponível em: https://docs.python.org/3/extending/index.html (Acesso: 30.06.2021).

[38] Ronald L. Rivest and Madars Virza. Software Independence Revisited. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 3–17, 2017.

[39] Phillip Rogaway and Thomas Shrimpton. Cryptographic Hash-Function Basics: Definitions, Implications, and Separations for Preimage Resistance, Second-Preimage Resistance, and Collision Resistance. In *Fast Software Encryption*, pages 371–388. Springer Berlin Heidelberg, 2004.

[40] Matthew Rosenberg, Nicholas Confessore, and Carole Cadwalladr. How Trump Consultants Exploited the Facebook Data of Millions. *The New York Times*, 2018. Disponível em: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html (Acesso: 20.05.2021).

[41] Peter Y. A. Ryan, Steve Schneider, and Vanessa Teague. Prêt à Voter - The Evolution of the Species. In *Real-World Electronic Voting: Design, Analysis and Deployment*, pages 309–341, 2017.

[42] Adi Shamir. How to Share a Secret. *Communications of the ACM*, 22(11):612–613, November 1979.

[43] Douglas R. Stinson. *Cryptography: Theory and Practice*. Chapman & Hall/CRC, 3 edition, 2006.

[44] TSE. Testes Públicos de Segurança do Sistema Eletrônico de Votação. https://www.tse.jus.br/eleicoes/eleicoes-anteriores/eleicoes-2010/testes-publicos-de-seguranca-do-sistema-eletronico-de-votacao, 2010. (Acesso: 17.05.2021).

[45] TSE. Série urna eletrônica: da máquina de votar ao voto informatizado. https://www.tse.jus.br/imprensa/noticias-tse/2013/Setembro/serie-urna-eletronica-da-maquina-de-votar-ao-voto-informatizado, 2013. (Acesso: 12.05.2021).

[46] TSE. Série urna eletrônica: RDV permite recontagem dos votos e amplia transparência. https://www.tse.jus.br/imprensa/noticias-tse/2013/Setembro/serie-urna-eletronica-rdv-permite-recontagem-dos-votos-e-amplia-transparencia, 2013. (Acesso: 20.05.2021).

[47] TSE. *Informações e Dados Estatísticos Sobre as Eleições 2014*. Tribunal Superior Eleitoral (TSE), 2014. Disponível em: https://www.tse.jus.br/o-tse/catalogo-de-publicacoes/lista-do-catalogo-de-publicacoes?publicacoes=informacoes-dados-estatisticos-eleicoes-2014 (Acesso: 17.05.2021).

[48] TSE. *Parecer da STI sobre a Petição TSE 23.891*. Tribunal Superior Eleitoral, 2014. Disponível em: https://cic.unb.br/~rezende/trabs/eleicoes2014/PeticaoTSE23891-parecerSTI.html (Acesso: 18.05.2021).

[49] TSE. *Resolução Nº 23.603*. Tribunal Superior Eleitoral, 2014. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2019/resolucao-no-23-603-de-12-de-dezembro-de-2019 (Acesso: 19.05.2021).

[50] TSE. Plenário do TSE: PSDB não encontra fraude nas Eleições 2014. https://www.tse.jus.br/imprensa/noticias-tse/2015/Novembro/plenario-

do-tse-psdb-nao-encontra-fraude-nas-eleicoes-2014, 2015.
(Acesso: 16.05.2021).

[51] TSE. *Resolução - Testes de Segurança das Urnas*. Tribunal Superior Eleitoral, 2015. Disponível em: https://www.justicaeleitoral.jus.br/arquivos/tse-pa-18862-teste-publicos-de-seguranca/rybena_pdf?file=https://www.justicaeleitoral.jus.br/arquivos/tse-pa-18862-teste-publicos-de-seguranca/at_download/file
(Acesso: 17.05.2021).

[52] TSE. Biometria - Planejamento 2017-2022. https://www.tse.jus.br/eleitor/biometria/informacoes-sobre-o-planejamento-da-biometria-2017-2022, 2017.
(Acesso: 12.05.2021).

[53] TSE. *Termo de confidencialidade do TPS tem objetivo de manter a segurança dos sistemas eleitorais*. Tribunal Superio Eleitoral, 2017. Disponível em: https://tse.jusbrasil.com.br/noticias/509938477/termo-de-confidencialidade-do-tps-tem-objetivo-de-manter-a-seguranca-dos-sistemas-eleitorais
(Acesso: 17.05.2021).

[54] TSE. *Edital do Teste Público de Segurança*. Tribunal Superior Eleitoral, 2019. Disponível em: https://www.justicaeleitoral.jus.br/tps/arquivos/TPS-Edital-Testes-Seguranca-2019.pdf
(Acesso: 17.05.2021).

[55] TSE. TSE celebra aniversário de 10 anos de uso do Linux na urna eletrônica. https://www.tse.jus.br/imprensa/noticias-tse/2019/Maio/tse-celebra-aniversario-de-10-anos-de-uso-do-linux-na-urna-eletronica, 2019.
(Acesso: 12.05.2021).

[56] TSE. *Informações e Dados Estatísticos Sobre as Eleições 2020*. Tribunal Superior Eleitoral (TSE), 2020. Disponível em: https://www.tse.jus.br/imprensa/noticias-tse/arquivos/catalogo-dados-estatisticos-eleicoes-2020/rybena_pdf?file=https://www.tse.jus.br/imprensa/noticias-tse/arquivos/catalogo-dados-estatisticos-eleicoes-2020/at_download/file
(Acesso: 24.05.2021).

[57] TSE. Manual do Mesário. https://www.justicaeleitoral.jus.br/eleicoes/mesario/assets/arquivos/manuais/Manual_do_Mesario_2020_final_WEB.pdf, 2020.
(Acesso: 12.05.2021).

[58] TSE. O caminho do voto: o que acontece depois que você aperta o botão "Confirma"? https://www.tse.jus.br/imprensa/noticias-tse/2020/Novembro/o-caminho-do-voto-o-que-acontece-depois-que-voce-aperta-o-botao-201cconfirma201d, 2020.
(Acesso: 13.05.2021).

[59] TSE. Série Desvendando a Urna: o equipamento não é auditável? https://www.tse.jus.br/imprensa/noticias-tse/2020/Outubro/serie-desvendando-a-urna-o-equipamento-nao-e-auditavel, 2020.
(Acesso: 16.05.2021).

[60] TSE. TSE confirma empresa Positivo como vencedora para aquisição de novas urnas eletrônicas. https://www.tse.jus.br/imprensa/noticias-tse/2020/Julho/tse-confirma-empresa-positivo-como-vencedora-para-aquisicao-de-novas-urnas-eletronicas, 2020.
(Acesso: 24.05.2021).

[61] TSE. Ecossistema da urna. https://www.tse.jus.br/eleicoes/processo-eleitoral-brasileiro/logistica-e-preparacao/ecossistema-da-urna, 2021.
(Acesso: 12.05.2021).

[62] TSE. Glossário Eleitoral - Boletim de Urna. https://www.tse.jus.br/eleitor/glossario/termos-iniciados-com-a-letra-b#boletim-de-urna, 2021.
(Acesso: 13.05.2021).

[63] TSE. Glossário Eleitoral - Registro Digital do Voto. https://www.tse.jus.br/eleitor/glossario/termos-iniciados-com-a-letra-r#registro-digital-do-voto, 2021.
(Acesso: 13.05.2021).

[64] TSE. Glossário Eleitoral - Zerésima. https://www.tse.jus.br/eleitor/glossario/termos-iniciados-com-a-letra-z#zeresima, 2021.
(Acesso: 13.05.2021).

[65] TSE. Preparação das urnas. https://www.tse.jus.br/eleicoes/processo-eleitoral-brasileiro/logistica-e-preparacao/preparacao-das-urnas, 2021.
(Acesso: 12.05.2021).

[66] TSE. Urna eletrônica. https://www.tse.jus.br/eleicoes/urna-eletronica, 2021.
(Acesso: 12.05.2021).

[67] TSE. Urna eletrônica 25 anos: lançado em 1996, equipamento é o protagonista da maior eleição informatizada do mundo. https://www.tse.jus.br/imprensa/noticias-tse/2021/Maio/urna-eletronica-25-anos-lancado-em-1996-equipamento-e-o-protagonista-da-maior-eleicao-informatizada-do-mundo, 2021.
(Acesso: 04.08.2021).

[68] Jeroen van de Graaf. *O Mito da Urna: Desvendando a (in)segurança da urna eletrônica*. Sem Editora, 2017. Disponível online em: https://inscrypt.dcc.ufmg.br/pt/o-mito-da-urna/.

[69] Alexander Yakobson. Secret ballot and its effects in the late roman republic. *Hermes*, 123(4):426–442, 1995.