Método de Autenticação Multi-canal Baseado em Proximidade

Ronaldo Mengato[1], Altair Santin[1], Vilmar Abreu[1], Mauro Borchardt[1]

[1]Pontifícia Universidade Católica (PUCPR)
Programa de Pós-Graduação em Informática
Curitiba – PR – Brasil

{ronaldo.mengato,santin,vilmar.abreu,mauro.borchardt}@ppgia.pucpr.br

Abstract. *Critical infrastructure (CI) systems are increasingly common today, with some of their features being exposed via the internet for remote use. However, such exposure involves risks that can cause serious damage to CI. An alternative to this is to consider the user's location as an authentication attribute, blocking remote attackers. However, the location-based authentication techniques founded in the literature are forgeable. This work proposes a method based on proximity, forcing the user to be close to reference devices (anchors). The method makes it possible to communicate authentication attributes on different channels, such as optical media and wireless networks. A prototype was built to evaluate the proposed method in terms of security and performance requirements.*

Resumo. *Sistemas de infraestrutura crítica (IC) são cada vez mais comuns atualmente, sendo que algumas de suas funcionalidades são expostas via internet para uso remoto. No entanto, essa exposição implica em riscos que podem causar danos graves a IC. Uma alternativa a isso é considerar a localização do usuário como um atributo de autenticação, bloqueando atacantes remotos. Porém, as técnicas de autenticação baseada em localização encontradas na literatura são forjáveis. Este trabalho propõe um método baseado em proximidade, obrigando o usuário a estar próximo a dispositivos de referência (âncoras). O método proposto possibilita a comunicação dos atributos de autenticação em diferentes canais, como o meio óptico e redes sem fio. Um protótipo exemplificando o método proposto foi implementado e avaliado quanto a requisitos de segurança e de performance.*

1. Introdução

Atualmente quase qualquer tipo de serviço está disponível virtualmente, oferecendo diferentes nuances de acesso a diferentes tipos de dispositivos. Alguns destes, controlam infraestruturas críticas e são acessíveis via Internet, o que pode expor vulnerabilidades que causam danos incalculáveis [Khandelwal 2018]. Um hacker pode explorar brechas de segurança e controlar o sistema remotamente, como se fosse um usuário legítimo.

A fim de mitigar este problema, é necessário implementar um mecanismo robusto de autenticação, de tal modo que mesmo explorando uma vulnerabilidade um hacker não possa controlar o sistema pela Internet. A autenticação multifator proporciona robustez combinando diferentes fatores de autenticação, assim, caso algum

destes seja comprometido isoladamente, o hacker não poderia obter acesso ao sistema. Neste trabalho um fator exigido é a prova de proximidade a um dispositivo âncora.

Na literatura, diversos autores propõem novos fatores e técnicas de autenticação, sendo que, a autenticação baseada em georreferenciamento tem sido bastante usada. Essa técnica visa autenticar um usuário apenas se este estiver fisicamente em um local ou dentro de um perímetro. Por exemplo, um médico realizando uma cirurgia por telepresença deve estar dentro de um hospital para acessar o sistema hospitalar remoto, o mesmo pode-se dizer de um funcionário em uma usina elétrica, que só pode operar equipamentos se estiver no prédio da unidade de controle. Parece fácil ter esta certeza, mas quando nos referimos a sistemas de segurança o *bypass control* é comum em casos de vulnerabilidades, então um mecanismo mais robusto precisa ser buscado.

Embora existam diversos trabalhos baseados no georreferenciamento, muitas das técnicas encontradas são forjáveis [Jansen et al. 2005, Talasila et al. 2015, Zhang et al. 2012, Xiao et al. 2013, Jaros and Kuchta 2010, Denning and MacDoran 1996], pois utilizam mecanismos frágeis para a obtenção da localização física, como por exemplo coordenadas de GPS em smartphones [Zhang et al. 2012, Jaros and Kuchta 2010], dispositivo único (canal) na autenticação [Denning and MacDoran 1996] como sinal de redes telefônicas ou então o uso de endereços IP [Jansen et al. 2005, Xiao et al. 2013].

Uma alternativa ao método georreferenciado é exigir a proximidade à um dispositivo de referência (âncora) como um requisito de autenticação. Com isso, os objetos âncoras podem ser estrategicamente posicionados em determinados locais físicos, ou então, carregados por pessoas escolhidas desde que sigam os mesmos critérios de proteção.

Como a proximidade é um fator físico que requer a presença do usuário próximo ao dispositivo âncora, é possível proteger fisicamente o dispositivo, colocando-o dentro de instalações monitoradas (salas, edifícios etc.) e observá-lo por meio dos sistemas de vigilância, como é feito em unidades de banco eletrônico, por exemplo. Uma âncora também pode ser transportada ou usada por alguém, nesse caso, a pessoa pode ter um serviço de proteção pessoal, se necessário.

É importante usar diferentes dispositivos (canais) nos fatores relacionados à autenticação por proximidade, para evitar que uma vulnerabilidade em um único canal torne o esquema todo vulnerável. Idealmente, cada dispositivo deve ser implantado em uma plataforma diferente, pois assim, um atacante terá o desafio de controlar dois ou mais dispositivos para conseguir autenticar-se. Na prática, o efeito protético dessa abordagem é o desafio da presença física e o controle de dois ou mais dispositivos diferentes.

A proposta deste trabalho é composta de dois métodos de autenticação por proximidade que usam um dispositivo âncora como referência, sendo que eles podem ser usados separadamente ou em conjunto. Os métodos propostos são: baseado em dispositivos wireless e baseado em leitura óptica. Um terceiro esquema utilizando wearable também pode ser aplicado para os cenários que exigem uma âncora móvel.

O método baseado em wireless transmite um desafio em uma rede sem fio de curto alcance, como o Bluetooth por exemplo, obrigando que o dispositivo que deseja se autenticar esteja próximo ao dispositivo âncora (emissor). Um exemplo de dispositivo desse método é o Beacon, que atua como referência em locais estratégicos. O desafio de

autenticação não é facilmente falsificado, pois utiliza a criptografia de Curva Elíptica para proteger a mensagem e garantir que quem gerou o desafio seja uma âncora válida, é que está sincronizada com um sistema de Gerenciamento de Identidade (*Identity Management* – IdM) rodando na nuvem computacional.

O método baseado em leitura óptica mostra um desafio por meio de um display LCD, sendo que esta âncora usa o formato QR Code para encapsular e exibir o desafio. Este valor exibido também não é facilmente forjável, porque o dispositivo é sincronizado com o IdM e muda o desafio a cada intervalo de tempo, utilizando o conceito de *Time-based One Time Password* (TOTP). Nesse caso, a resposta do desafio é passada ao IdM como uma senha única (*One Time Password* – OTP).

O esquema baseado em wearable usa um dispositivo vestível (tal como um smartwatch) como âncora móvel. O desafio da autenticação é exibido no relógio no formato QR Code e deve ser lido por um smartphone. Deste modo, assim como no método óptico, a resposta ao desafio não é facilmente forjável, porque há uma sincronização entre o IdM e o usuário vestindo o aparelho. Este método também usa o TOTP para disponibilizar um QR Code diferente a cada intervalo. A vantagem deste esquema é que os usuários habilitados podem transportar o dispositivo âncora e realizar operações críticas, mesmo não estando fisicamente em locais previamente estabelecidos. O grau de segurança da abordagem continua sendo robusto, pois, para tornar o fator de proximidade possível, é necessário ter o dispositivo âncora móvel e estar próximo dele.

Como principal contribuição, a proposta deste trabalho evita que um hacker controle uma infraestrutura crítica remotamente explorando uma vulnerabilidade. Isso porque a presença física próxima a um dispositivo âncora é um fator que deve ser atendido para prosseguir na autenticação multifator e multicanal. Mais especificamente, o trabalho contribui com: (i) desenvolvimento de uma autenticação multicanal (óptica e wireless) para um IdM, (ii) criação de um método de autenticação robusto e multifator (proximidade e multicanais) e (iii) desenvolvimento de um método de sincronização entre cliente e IdM sem troca de chaves sensíveis pela rede.

As seguintes seções são organizadas da seguinte maneira: Seção II explica alguns dos temas usados para a construção da proposta (*background*). Seção III aponta os trabalhos relacionados existentes na literatura, enquanto na Seção IV a proposta do método de autenticação é descrita em detalhes. Em seguida, na Seção V é apresentado um estudo de caso e na Seção VI são mostrados os detalhes da construção de um protótipo. Finalmente, na Seção VII são abordadas as conclusões.

2. Background

2.1 Criptografia de Curva Elíptica

Para alcançar a conformidade com os requisitos básicos de segurança, como confidencialidade e integridade, adicionamos um modelo criptográfico que usa as chaves da curva elíptica (ECC). Esse tipo de criptografia oferece um nível de segurança equivalente a outros modelos tradicionais, mas com um tamanho menor de chaves [Menezes 2012].

Geralmente, essas chaves são usadas em conjunto com um protocolo *key agreement*, ou seja, procedimento em que duas ou mais entidades usam para obter um

valor comum e, em seguida, derivam uma ou mais chaves simétricas. O algoritmo Elliptic Curve Diffie-Hellman (ECDH) é um exemplo de um procedimento específico para curvas elípticas.

As chaves do tipo Curve25519 são exemplos de chave de curva elíptica, definida pela RFC7748 [Langley et al. 2016]. A Curve25519 é projetada para cenários de alto desempenho, especialmente nos casos em que se deseja reutilizar o mesmo par de chaves para futuras trocas. Seu nível de segurança é de 128 bits, sendo comparável ao NIST P-256 [Simon and Nir 2016]

2.2 Smart Grid

O Smart Grid (SG) é uma evolução dos sistemas tradicionais de energia, tendo como objetivo a geração, transmissão e distribuição de energia elétrica de forma mais autônoma e eficiente [Fang et al. 2012]. Dentro de um SG existem vários componentes e estágios nos quais dados e energia são transmitidos ou armazenados. Talvez um dos componentes mais importantes seja o sistema central, onde os operadores têm acesso aos dados do SG [11]. É neste módulo que operam os sistemas de *Supervisory Control e Data Acquisition* (SCADA). O trabalho [Igure et al. 2006] aponta que o controle de acesso e autenticação em redes SCADA é relevante e ainda está pendente como um desafio. Desta forma, o este trabalho tem o potencial de auxiliar na segurança de um Smart Grid.

3. Trabalhos Relacionados

Esta seção apresenta os trabalhos disponíveis na literatura relacionados à autenticação baseada em proximidade e baseada em usuário. Apesar da existência de uma quantidade significativa de trabalho que usa o local do usuário para oferecer algum serviço, poucos o usam para autenticação.

O trabalho de Jansen *et al.* (2005) usa dois PDAs (*personal digital assistant*): um fixo em um local e o outro móvel. Um PDA atua como servidor e usa uma infraestrutura de chave pública (PKI) e um certificado X.509. O modelo funciona da seguinte forma: na fase de configuração é necessário gerar os pares de chaves RSA e certificados, tanto no servidor quanto no dispositivo cliente, portanto, quando o usuário entra em um ambiente com seu PDA, procurará algum PDA fixado no local para estabelecer uma conexão segura por meio do TLS (*Transport Layer Security*). Uma vez que a conexão tenha sido estabelecida com sucesso, o PDA está no estado autenticado, habilitando as funções disponíveis para a esse PDA fixo. Apesar de trazer robustez ao modelo, o uso das chaves RSA e o algoritmo de desafio-resposta (usado para sincronizar os PDAs) traz um grande custo computacional para os dispositivos. Outra limitação é no caso de o PDA fixo ser comprometido, seja para roubo de certificado ou negação de serviço, todo o sistema fica fragilizado.

A proposta de Talasila *et al.* (2015) visa um ambiente mais colaborativo. Este trabalho, chamado de LINK, visa determinar a localização de um dispositivo com base em seus vizinhos. Cada *claimer* (usuário) precisa se autenticar para utilizar alguma funcionalidade, que depende do contexto no seu entorno. Para isso, é preciso obter uma credencial de localização emitida por uma entidade chamada *Location Certification Authority* (LCA). Essa entidade controla todos os *claimers*, então, no momento da autenticação o usuário deve fazer um broadcast via Bluetooth para todos os seus

vizinhos, assim, quem recebe essa mensagem deve notificar a LCA que está próximo do host que está se autenticando. Todas as mensagens são criptografadas, pois, cada *claimer* possui seu próprio certificado, configurado em uma fase anterior. No entanto, pelo fato de utilizar *scores* dos usuários vizinhos, este trabalho precisa de uma grande quantidade de *claimers* o que limita o seu uso a sistemas grandes. Ou seja, a localização de um host se dá pela proximidade deste com os *claimers* considerados confiáveis, assim, quanto mais vizinhos maior será a taxa de acerto da LCA.

Outra técnica usada para determinar a localização dos usuários é por meio do RSSI (*Received Signal Strength Indicator*) disponível nas redes sem fio, como por exemplo, o indicador de intensidade do sinal recebido de um *Access Point* (AP). Essa abordagem é usada no trabalho de Xiao *et al.* (2013), porém, tem a desvantagem de ser facilmente forjada. Se apenas o RSSI de um AP for utilizado para determinar o local, é possível simular esse AP com o mesmo endereço MAC, até mesmo SSID (*Service Set Identifier*), e assim coletar o valor do RSSI e se autenticar de qualquer lugar no mundo.

Outro trabalho relacionado é descrito em Zhang *et al.* (2012), onde os autores propõem o uso do *Mobile Network Operator* (MNO) presente nos smartphones. A ideia é que o cliente e seu dispositivo recebam uma credencial de uma entidade. Assim, quando um cliente tenta se autenticar, o provedor de serviços da rede mobile retorna a área geográfica permitida para aquele usuário. O smartphone então consulta o MNO para descobrir qual é a área de cobertura. Em seguida, os dados recebidos que indicam a latitude e longitude da área são convertidos para o plano cartesiano e então é feita uma verificação para descobrir se o smartphone está dentro desse perímetro pré-estabelecido. Embora os autores afirmem que o modelo é preciso quanto a determinação da localização, o esquema depende do MNO e da disponibilidade do sinal da operadora.

Apesar da grande variedade de técnicas encontradas na literatura, ainda há limitações a serem superadas. Por exemplo, há casos em que a credencial de localização é facilmente forjada [Zhang et al. 2012, Denning and MacDoran 1996], outros em que todo o sistema é mantido como refém por um único ponto de falha [Jansen et al. 2005, Denning and MacDoran 1996], e ainda técnicas para cenários específicos e predeterminados. Além disso, a Tabela 1 exemplifica com uma comparação que nenhum dos trabalhos encontrados na literatura propõe uma solução integrada, com algum IdM, múltiplos fatores de autenticação e obtenção da localização não forjável.

Tabela 1. Comparação entre os trabalhos relacionados.

Trabalho	Multicanal	Multifator	IdM	Localização Forjável
Jansen et al. (2005)	Não	Não	Não	Não
Talasila et al. (2015)	Não	Não	Não	Não
Xiao et al. (2013)	Não	Não	Não	Sim
Zhang *et al.* (2012)	Não	Não	Não	Sim

4. Autenticação Baseada em Proximidade

O objetivo da nossa proposta é fornecer um método seguro e confiável (não forjável) de autenticação baseada em proximidade. Essa técnica é necessária quando for preciso garantir que um usuário esteja operando o sistema fisicamente a partir de um local estratégico. Esse fator diminui a possibilidade de um intruso da Internet acessar um

sistema que não está autorizado, pois este não terá a posse da credencial de proximidade. É importante ressaltar que o uso dessa autenticação não visa impedir o acesso remoto, mas sim, limitar os locais "remotos" de onde os usuários podem se autenticar.

Nossa proposta de autenticação é baseada na proximidade física do usuário com um dispositivo âncora, que deve estar posicionado em um determinado local (sala de operação no caso de energia elétrica, por exemplo). A comunicação entre o usuário e a âncora é multicanal para evitar que um dispositivo comprometido, afete todo o sistema de autenticação. Desta forma, propomos três métodos de autenticação baseados na proximidade, que podem ser usados individualmente ou em combinação.

A Figura 1 mostra o modelo de autenticação por proximidade proposto, que é composto pelos dispositivos âncora, sistema de Gerenciamento de Identidade (IdM) e os recursos a serem protegidos. O IdM é um serviço responsável pelo gerenciamento dos atributos de identidade do usuário, tais como e-mails, IDs e senhas. Além dos atributos convencionais, o IdM é responsável pelo gerenciamento das credenciais de proximidade do usuário. Para isso, possui um módulo responsável por validar as credenciais de proximidade emitidas por âncoras previamente registradas.

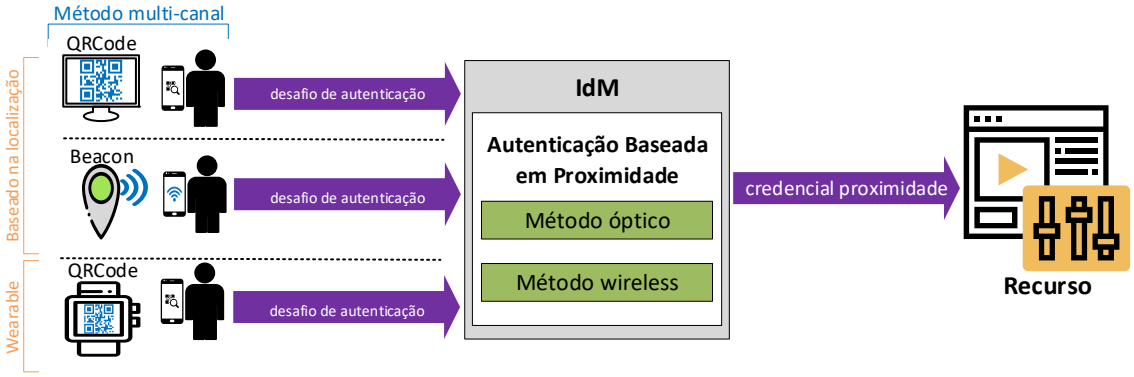

Figura 1. Modelo de autenticação baseada em proximidade

O registro dos dispositivos âncora é feito pelo administrador do sistema, que deve informar ao IdM os seguintes parâmetros: credenciais de administrador, a chave pública do dispositivo e o período em que a credencial permanecerá válida. Essa fase inicial é executada apenas uma vez, no momento de instalação da âncora no local, e garante que o algoritmo usado para gerar a credencial de proximidade seja sincronizado entre IdM e âncoras. Assim, o esquema ganha força, já que o desafio da autenticação não é transmitido pela rede, mas gerado no dispositivo âncora.

O processo para obter a credencial de proximidade e sua respectiva validação serão detalhados nas subseções seguintes, onde apresentamos os três métodos baseados em proximidade.

4.1. Método baseado em Wireless

O método baseado em wireless usa um canal sem fio de baixo alcance para transmitir credenciais de proximidade. Essas credenciais são emitidas por um dispositivo fixo (âncora), chamado Beacon. O Beacon é um dispositivo que transmite dados continuamente usando uma rede *Bluetooth Wireless Low Energy* (BLE), é um

dispositivo da Internet das Coisas (IoT) que possui hardware limitado, composto apenas de um pequeno processador, módulo de rádio e baterias [LaMarca et al. 2005].

O Beacon é colocado em um determinado local, sendo responsável pela emissão contínua da credencial de proximidade através do sinal Bluetooth, que comumente tem seu alcance máximo entre 30 e 40 metros. Entretanto, o Beacon convencional, tipicamente encontrado na literatura, tem a limitação de sempre emitir o mesmo valor como credencial, sendo facilmente imitável (forjável)).

Para mitigar a possibilidade de um invasor da Internet, imitar esse valor, adotamos o protocolo Eddystone [Eddystone s.d], desenvolvido pelo Google. Este protocolo tem um modo de comunicação chamado Eddystone-EID (*Ephemeral Identifier*), que transmite através da rede Bluetooth um identificador criptografado que se altera após um intervalo de tempo pré-estabelecido [Hassidim et al. 2016].

Para a configuração de criptografia, na fase de registro do dispositivo, é necessário adotar um protocolo *key agreement*, que define uma chave simétrica entre o Beacon e o IdM para que estes possam gerar o mesmo EID.

A criptografia é possível porque, na fase de configuração, ocorre o *key agreement*. Esse processo funciona da seguinte maneira: cada parte envolvida tem seu próprio par de chaves Curve25519 e, no momento do registro, cada uma delas disponibiliza sua chave pública. Em seguida, por meio do protocolo ECDH, cada host calcula a chave compartilhada multiplicando sua chave privada pela chave pública da outra. Dessa forma, eles estão compartilhando a mesma chave, tanto o Beacon e seu *Resolver* (IdM) podem gerar o mesmo EID.

Esse protocolo criptográfico adiciona robustez à solução, porque o valor de EID é modificado em um intervalo de tempo personalizável. Além disso, o valor de EID é baseado em uma chave secreta que nunca foi transmitida na rede.

Assim, o Beacon é responsável pela transmissão do EID, que neste caso é a credencial de proximidade. O usuário que está fisicamente próximo ao Beacon pode capturar a credencial e enviá-la ao IdM através de seu smartphone. Assim, no momento da autenticação, o IdM deve validar o EID recebido. A validação requer que o IdM gere seu próprio EID e faça a comparação.

Além das vantagens de segurança fornecidas por este método, a usabilidade também é interessante. Porque o usuário não precisa executar nenhuma ação ativa para obter a credencial emitida pelo Beacon, apenas estar próximo com seu smartphone. No entanto, a principal desvantagem deste método é o caso do ambiente do usuário apresentar ruídos eletrônicos na frequência do Bluetooth, podendo impedir a autenticação. Uma alternativa para essa ocasião será discutida na próxima seção.

4.2 Método óptico

O método baseado em leitura óptica usa o canal menos sensível ao ruído para emitir a credencial de proximidade. Esta credencial está disponível graficamente através de um QRCode exibido em um LCD posicionado em local pré-estabelecido. Este dispositivo que apresenta a credencial pode ser qualquer dispositivo com conexão à Internet e que contém um display, por exemplo: PC, smartphone, tablet etc. A conexão à Internet é obrigatória apenas no momento do registro do dispositivo no IdM.

A credencial fornecida pelo QRCode é gerada por um algoritmo TOTP (*Time-based One Time Password*) [M'Raihi et al. 2011]. O TOTP é um método comumente encontrado em mecanismos de autenticação multifator que exigem uma credencial temporária e dinâmica. O TOTP usa uma função hash, por exemplo o HMAC-SHA-512, para gerar um valor pseudoaleatório. Os parâmetros necessários para a geração são: uma chave secreta compartilhada entre as partes (dispositivo e IdM) e um número representando a variação de tempo. Esses valores são definidos no registro do dispositivo, executado pelo administrador do sistema.

Desta forma, o usuário deve scanear o QRCode exibido pelo dispositivo para obter o valor gerado pelo TOTP. Em seguida, o usuário envia ao IdM o valor TOTP a ser validado. Para realizar a validação, o IdM deve executar o algoritmo TOTP usando os mesmos parâmetros do dispositivo. A autenticação do usuário é válida se o valor gerado pelo IdM for igual ao valor gerado pela âncora. É importante notar que ambos os métodos requerem um sistema de proteção física para garantir que o dispositivo âncora esteja realmente conectado a um local previamente definido.

4.3 Método baseado em wearable

Este método é baseado no método óptico, usando QRCode e TOTP para fornecer a credencial de proximidade. A diferença é a mobilidade. Nesse cenário, o dispositivo de âncora é um wearable (um smartwatch, por exemplo) que o usuário pode usar e carregar consigo. Assim, no momento da autenticação o código deve ser lido de um wearable através do smartphone para ser enviado ao IdM.

A fase de registro desse método também é diferente, porque o wearable deve ser vinculado a um usuário e não a um local físico. Isso também garante que a credencial de proximidade emitida pelo âncora não possa ser autenticada com credenciais de outro usuário.

Com esse método, o esquema ganha flexibilidade para tratar de casos em que o usuário deve acessar o recurso, mas não pode estar fisicamente no local onde está fisicamente um âncora fixo. Por exemplo, um médico em férias deve acessar o prontuário de um paciente em caso de emergência e só o faz se estive com o wearable junto. Além disso, o modelo permanece robusto porque o usuário deve ter dois dispositivos no momento da autenticação de proximidade.

5. Estudo de caso: Sistema de Smart Grid

Esta seção apresenta um estudo de caso do modelo proposto em um ambiente de Smart Grid (SG). O objetivo principal é exemplificar a aplicabilidade da proposta em um cenário de infraestrutura crítica. O modelo proposto pode ser usado para aumentar a segurança em diferentes níveis de um SG. Apesar do fato de que neste tipo de cenário, os usuários tendem a estar em lugares remotos, porém há situações em que a proximidade é necessária. Tal como no centro de operações, de onde o sistema todo é monitorado e comandado.

Os sistemas SCADA são normalmente operados em estações centrais de um SG, fornecendo uma visão geral dos dispositivos que compõem o SG. Portanto, tais sistemas são críticos e garantem que sua segurança seja um desafio complexo [Igure et al. 2006]. A adoção de autenticação baseada em proximidade no sistema SCADA garante que o

operador esteja dentro da sala de operação, que possui controle de acesso físico mais rígido.

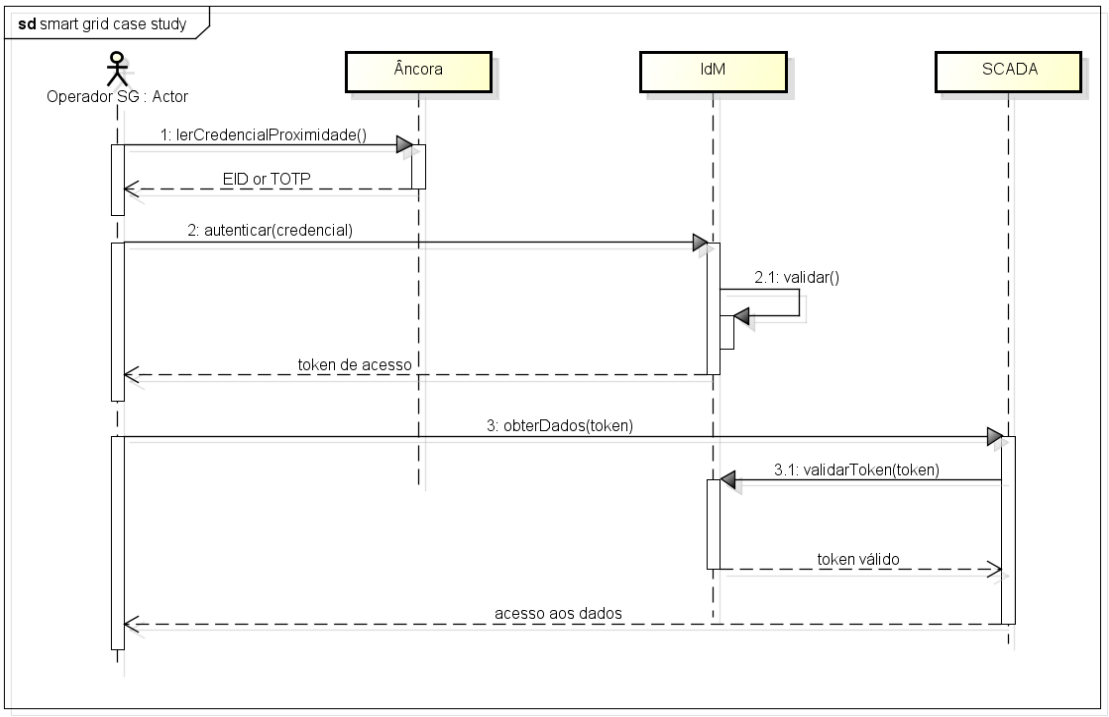

Figura 2. Fluxo da autenticação baseada em proximidade em SG

A Figura 2 mostra um diagrama com o fluxo de mensagens necessárias para o operador obter acesso aos dados da SG. Por meio de um aplicativo cliente executado em um computador ou smartphone, o operador lê a credencial de proximidade do dispositivo de âncora (evento 1). Esta credencial pode ser obtida de qualquer um dos três métodos propostos. Depois de obter a credencial, o operador a encaminha para o IdM (evento 2). Se for uma credencial válida, o operador recebe um token de acesso e o envia para o SCADA solicitando acesso a um recurso do sistema central (evento 3). O sistema SCADA verifica com o IdM se o token recebido é válido, em caso afirmativo permite o acesso ao recurso.

A abordagem fornece robustez e usabilidade ao operador, especialmente usando smartwatch no momento da autenticação. No entanto, é importante enfatizar o risco de ataques internos de funcionários ou pessoas com acesso físico autorizado. Para esses casos, o uso de auditorias e a renovação periódica de registros de dispositivos de âncoras se tornam ações obrigatórias.

6. Protótipo

Esta seção detalha a implementação e a avaliação do protótipo do modelo proposto.

6.1 Implementação

A Figura 3 mostra as tecnologias utilizadas na implementação do protótipo. A implementação do IdM foi dividida em dois componentes: *Authenticator* e *Resolver*.

Para a construção do componente *Authenticator*, foi utilizada a linguagem Java e o framework Spring Security, com o pacote Spring OAuth2 e o MITREid Connect (uma

implementação Java do padrão OpenID Connect). O *Resolver* também foi construído em Java e disponibilizados APIs no formato REST para comunicação. Além disso, uma base de dados no PostgreSQL mantém os dados dos usuários e suas identidades, assim como os registros de âncoras ativas e suas respectivas localizações. Estes dois componentes compõe o IdM do modelo proposto.

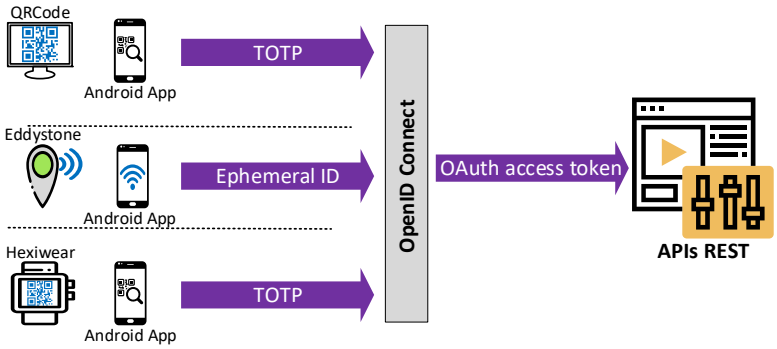

Figura 3. Componentes do protótipo

Outro componente desenvolvido no protótipo foi um aplicativo móvel que o usuário usa para se autenticar, e depois acessar algum recurso ou informação privilegiada. O aplicativo é responsável por coletar uma credencial de identidade do usuário (e-mail, por exemplo) e fazer uma requisição ao IdM para autenticar a credencial. Em seguida, a credencial de proximidade é coletada por meio dos métodos de autenticação para serem autenticados também no IdM. No final deste procedimento, se as duas credenciais forem válidas o aplicativo receberá uma resposta contendo o token de acesso, que é usado para acessar os recursos solicitados pelo usuário.

Para implementar o método baseado em wireless, um simulador de beacon foi desenvolvido para a plataforma Android com o protocolo Eddystone-EID. Para manipular as chaves Curve25519 foi utilizada a biblioteca Signal [Curve25519 s.d], uma vez que esta é compatível com a plataforma Java. Assim, qualquer dispositivo compatível com Android, conexão com a Internet e Bluetooth 4.0+, pode funcionar como um Beacon.

Para cadastrar o Beacon o administrador deve informar ao IdM: a) o ID do local onde está sendo implantado o dispositivo; b) o expoente de rotação (K) indicando o tempo em que o EID estará ativo (2^K segundos); c) um contador inicial para o beacon; d) um apelido de identificação; e) a credencial do administrador (seu login e senha). Em seguida, o aplicativo gera seu par de chaves e envia sua chave pública junto com os parâmetros informados ao servidor. Depois que o simulador é registrado começará a transmitir o EID via BLE, quando o período de rotação expira, o aplicativo gera automaticamente um novo valor sem precisar se comunicar com o servidor.

Para implementar o método baseado em leitura óptica, foi criado um gerador QRCode em Java, podendo ser instalado em diversos dispositivos, incluindo dispositivos mobile. Também usamos a biblioteca de Signal para as chaves, e a AeroGear [AeroGear s.d] como implementação da especificação TOTP. Além disso, a biblioteca QRCodeGen [QRCodeGen s.d] foi usada para gerar a matriz de pontos encapsulando o valor TOTP.

O registro do dispositivo gerador de QRCode é semelhante ao beacon, o administrador deve enviar os parâmetros: a) identificação do local; b) chave pública; e c) sua credencial. O IdM então salva os dados no banco de dados, gera um número aleatório de 90 a 120 (segundos que o TOTP estará ativo) e retorna este valor junto com sua chave pública para o dispositivo, assim ambos os lados usam a chave pública – um do outro para gerar a chave compartilhada. Essa chave é usada como semente para gerar o valor TOTP que será lido pelo aplicativo do usuário.

O desenvolvimento do método baseado em wearable foi implementado com o smartwatch Hexiwear [Hexiwear s.d], uma solução de hardware e software para protótipos de dispositivos Wearable. Foi necessário usar uma estação integradora (*docking station*) com um módulo de conexão WiFi para comunicar o smartwatch com o IdM. Sua operação é semelhante à do QRCode, sendo que a diferença é apenas o registro pois não precisa do ID do local. O registro precisa do id do usuário vinculado ao relógio na verdade.

6.2 Avaliação de Segurança

Para apresentar os possíveis riscos dentro do modelo proposto, a Tabela 2 apresenta o modelo do adversário, que mostra quais os possíveis impactos dentro da arquitetura quando um componente for comprometido por algum atacante. Possíveis contramedidas que poderiam ser exploradas em trabalhos futuros também são apresentadas.

Outros componentes tradicionais dentro do IdM, como os módulos de controle de acesso e autorização, devem ter os mesmos cuidados que em qualquer outra implementação. Por exemplo, sempre usar sistemas de detecção de intrusão, modelos preditivos e políticas de resposta a incidentes [Steven and Peterson 2006].

Tabela 2. Modelo do Adversário

Componente	Impacto	Contramedida
Âncora baseada em wireless (Beacon)	Se o invasor tiver acesso aos arquivos do dispositivo, saberá os parâmetros usados para gerar um EID válido. O impacto de um invasor escutar a rede Bluetooth é mínimo, já que além do alcance baixo que limita sua posição física, este deve enviar o EID de um aplicativo válido para o IdM. Ainda o invasor precisaria saber um login e a senha válidos.	• Fechar a conexão com a internet logo após a fase de registro; • Adicionar data de validade para o registro do dispositivo; • Posicionar o Beacon em local fechado, limitando seu alcance; • Invalidar o registro de uma âncora no caso de vários usuários errarem suas senhas.
Âncora baseada em leitura óptica (QR Code)	Um código malicioso na câmera do cliente poderia ler o valor e enviá-lo para um atacante remoto, que, de posse da senha, poderia ser	• As mesmas contramedidas do Beacon; e • O valor TOTP só

	autenticado remotamente. O controle do dispositivo também daria energia ao atacante para reproduzir o valor do TOTP.	permanece ativo por no máximo 120 segundos; • Desligar o display LCD quando este não estiver em uso.
IdM	O invasor pode registrar uma nova âncora em regiões não permitidas. No entanto, seria necessário conhecer as credenciais de identidade de algum usuário válido com privilégios de administrador.	• Desativar imediatamente o segundo fator de autenticação e notificar o administrador.
Aplicativo mobile	O invasor pode tomar posse apenas da credencial de um usuário específico. No entanto, só seria capaz de se autenticar caso estiver perto fisicamente de alguma âncora válida.	• Notificar o usuário sobre as tentativas de acesso em outros canais, assim, este teria ciência do ataque e mudaria sua credencial.

6.3 Avaliação de performance

Para avaliar o desempenho do servidor IdM, vários testes foram realizados com solicitações de autenticação simultâneas. O servidor foi instalado em uma máquina física com 8GB de RAM, processador de 8 núcleos com clock de 3.2GHz e sistema operacional Ubuntu 14.04 LTS. A fase de coleta da credencial de proximidade foi desconsiderada dos testes, pois para cada cenário as etapas são diferentes e relacionadas ao comportamento humano. O cliente é um aplicativo Java projetado para gerar a credencial (EID ou TOTP) e enviar a solicitação para o IdM, salvando o tempo necessário para executar esta ação. É possível executar o cliente passando dois parâmetros: qual cenário será testado (Beacon ou QRCode) e o número de threads simultâneas executando a requisição. O cliente é executado no sistema operacional Windows 10, 16GB de RAM e processador Intel 8-core i7 com 8 threads e clock de 3,6GHz.

A Figura 4 mostra o gráfico com os resultados das execuções dos testes, o eixo X representa o número de solicitações simultâneas e o Y o tempo médio que o IdM levou para atender uma única requisição. A escalabilidade do servidor mostra bom comportamento à medida que o número de solicitações aumenta. É possível observar que o cenário com QRCode leva vantagem em relação ao Beacon. Isso ocorre devido à maneira como o EID é validado, o IdM deve ler cada beacon na sua relação do banco de dados, gerar seu próprio EID e comparar com o EID do cliente, enquanto no QRCode o código do dispositivo e o código de localização são anexados ao valor TOTP, facilitando a busca do IdM no momento da autenticação. Embora seja um pouco mais lenta, essa abordagem do Beacon foi desenvolvida para garantir a interoperabilidade com outros *Resolvers*, inclusive com o Google Resolver.

Por fim, os testes concluem que a implementação do IdM proposta com validação de proximidade é viável mesmo para hardwares modestos. Além disso, a

proposta é facilmente aplicável em cenários com centenas de usuários simultâneos. Usar o QRCode pode ser menos intuitivo para o usuário final. No entanto, é mais escalável do que o Beacon, tornando-o preferível para cenários com milhares de usuários.

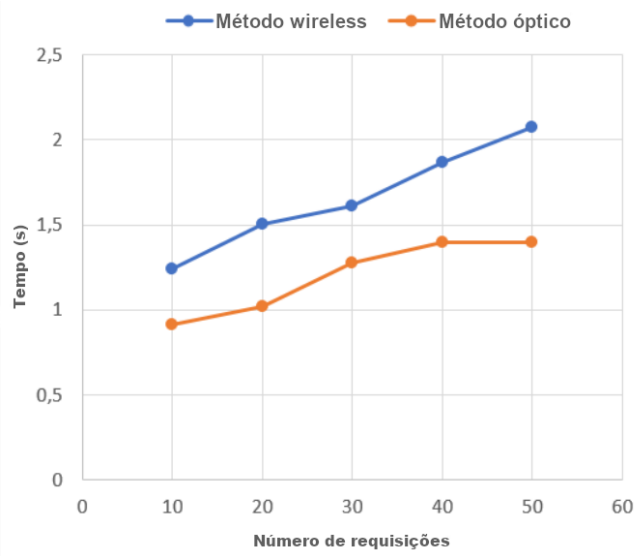

Figura 4. Resultados da avaliação de performance

7. Conclusão

Este trabalho propos um sistema de autenticação multifator baseado na proximidade do usuário dos dispositivos âncoras, devido aos sistemas de autenticação fraca, ao grande número de invasores distribuídos geograficamente e aos requisitos de segurança de sistemas de infraestrutura crítica (CI).

O uso de dispositivos como Beacon e QRCode, juntamente com chaves baseadas em curva elíptica, protocolo ECDH e um sistema IdM personalizado tornam o modelo proposto uma solução robusta e difícil de forjar. O estudo de caso fornece um exemplo real de aplicação do modelo, demonstrando sua viabilidade.

Com a implementação do protótipo foi possível instanciar os componentes propostos no modelo e analisar seu comportamento, principalmente como interagem e se integram com tecnologias da indústria como Spring e Eddystone-EID.

A validação de segurança mostra que um invasor deve conseguir comprometer o dispositivo de âncora e o aplicativo do usuário para se autenticar. Mesmo assim, possíveis contramedidas são apresentadas e podem ser exploradas em trabalhos futuros.

No entanto, apesar da necessidade de mais alguns testes serem executados, especialmente com âncoras reais, o modelo mostrou grande potencial de aplicação em ambientes onde é necessária a autenticação forte, fornecendo um método não invasivo e de fácil implantação e manutenção. A fim de melhorar e refinar a proposta, o trabalho futuro pode explorar o uso de novos dispositivos de ancoragem ou mesmo implantações em um ambiente real.

Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), concessão 310671 / 2012-4.

References

AeroGear. AeroGear Security OTP Specification. In https://aerogear.org/docs/specs/aerogear-security-otp/

Cao, Y. and Yang, L. (2010). “A survey of Identity Management technology,” 2010 IEEE Int. Conf. Inf. Theory Inf. Secur., pp. 287–293, 2010.

Curve25519, https://github.com/signalapp/curve25519-java

Denning, D. E. and MacDoran, P. F. (1996). “Location-based authentication: Grounding cyberspace for better security,” Comput. Fraud Secur., vol. 1996, no. 2, pp. 12–16, 1996.

Eddystone, https://developers.google.com/beacons/eddystone.

Fang, X., Misra, S., Xue, G. and Yang, D. (2012). “Smart Grid — The New and Improved Power Grid: A Survey,” IEEE Commun. Surv. Tutorials, vol. 14, no. 4, pp. 944–980, 2012.

Hassidim, A., Matias, Y., Yung, M. and Ziv, A. (2016). “Ephemeral Identifiers: Mitigating Tracking & Spoofing Threats to BLE Beacons,” pp. 1–11, 2016.

Hexiwear, https://www.hexiwear.com/.

Igure, V. M., Laughter, S. A., Williams, R. D. (2006) “Security issues in SCADA networks,” Computers & Security, Volume 25, Issue 7,2006.

Jansen, W., Korolev, V. and Hamilton, B. (2005) “Proximity-based Authentication for Mobile Devices,” 2005.

Jaros, D. and Kuchta, R. (2010). “New location-based authentication techniques in the access management,” Proc. - 6th Int. Conf. Wirel. Mob. Commun. ICWMC 2010, no. 1, pp. 426–430, 2010.

Khandelwal, S. (2018). "Bank Servers Hacked to Trick ATMs into Spitting Out Millions in Cash". In https://thehackernews.com/2018/10/bank-atm-hacking.html.

LaMarca, A., Chawathe, Y., Consolvo, S., Hightower, J. and Smith, I. (2005). “Place Lab: Device Positioning Using Radio Beacons in the Wild,” in International Conference on Pervasive Computing, 2005, pp. 116–133.

Langley, A., Hamburg, M. and Turner, S. (2016). “Elliptic curves for security,” No. RFC 7748. 2016.

Menezes, A. (2012). “Elliptic Curve Public Key Cryptosystems”, Springer Science & Business Media, 2012.

M'Raihi, D., Machani, S., Pei, M. and Rydell, J. (2011). "TOTP: Time-Based One-Time Password Algorithm", RFC 6238, DOI 10.17487/RFC6238, May 2011, <https://www.rfc-editor.org/info/rfc6238>.

OpenID, https://openid.net/

QRCodeGen, https://github.com/kenglxn/QRGen.

Simon, J. and Nir, Y. (2016). "Curve25519 and Curve448 for the Internet Key Exchange Protocol Version 2 (IKEv2) Key Agreement," 2016.

Steven, E. J. and Peterson, G. (2006). “Introduction to Identity Management Risk Metrics.”

Talasila, M., Curtmola, R. and Borcea, C. (2015). “Collaborative Bluetooth-based location authentication on smart phones,” Pervasive Mob. Comput., vol. 17, no. PA, pp. 43–62, 2015.

Xiao, L., Yan, Q., Lou, W., Chen, G. and Hou, Y. T. (2013). “Proximity-based security techniques for mobile users in wireless networks,” IEEE Trans. Inf. Forensics Secur., vol. 8, no. 12, pp. 2089–2100, 2013.

Zhang, F., Kondoro, A. and Muftic, S. (2012). “Location-based authentication and authorization using smart phones,” Proc. 11th IEEE Int. Conf. Ubiquitous Comput. Commun. IUCC-2012, pp. 1285–1292, 2012.