Seleção PBAD/LabSEC - Etapa final

18 de outubro de 2024

1 Introdução

Esta etapa do processo de seleção consiste em implementar uma aplicação capaz de realizar algumas operações comuns à segurança da computação. O objetivo do desafio não é avaliar o seu grau de conhecimento, mas sim como você se comporta ao se deparar com um desafio.

Aqui serão considerados alguns fatores: (i) quanto você se dedicou para resolver o desafio; (ii) quanto do desafio foi resolvido; (iii) qual o comportamento que teve quando surgiu dúvidas; (iv) sua organização do código; (v) entre outros fatores.

No capítulo seguinte, serão apresentados alguns conceitos básicos, necessários para o bom entendimento da implementação do desafio.

Desde já, agradecemos o esforço do autor original deste desafio e documento, Lucas Ferraro.

2 Conceitos básicos de criptografia

Segundo Menezes *et al.* [MVvO96], criptografia é o estudo de técnicas matemáticas relacionadas a aspectos da segurança da informação, tais como confidenciabilidade, integridade, autenticação e certificação. De todos os objetivos da segurança da informação, os quatro acima formam uma base, onde os demais são derivados deles.

- **Confidencialidade** é um serviço usado para manter o conteúdo da informação para todos que tenham autorização para tal. O segredo é sinônimo de confidencialidade e privacidade. Existem várias abordagens para proporcionar confidencialidade, que vão desde a proteção física até algoritmos matemáticos que tornam os dados ininteligíveis.
- **Integridade** é um serviço que trata da alteração não autorizada de dados. Para assegurar a integridade dos dados, é preciso conseguir detectar a manipulação dos dados por terceiros não autorizados. Essa manipulação se refere a ações como inserção, exclusão e substituição.
- **Autenticação** é um serviço relacionado à identificação. Essa função se aplica a ambas as entidades e a própria informação. Duas partes que começam uma comunicação devem se identificar uma à outra.
- **Não-repúdio** é um serviço que evita uma entidade de negar autorizações ou ações anteriores. Quando surgir alguma disputa em que uma entidade não admita que autorizou que certa ação fosse tomada, é necessária uma forma de resolver o problema, então uma terceira entidade confiável é convocada para resolver a disputa.

A criptografia é caracterizada por três diferentes dimensões. A primeira é a operação usada para transformar o texto plano, ou seja, o texto original, em texto cifrado. O segundo é a quantidade de chaves usadas e o terceiro é a forma que o texto plano é processado. Para esse desafio, o ponto mais importante é o segundo, pois diferencia a criptografia simétrica da assimétrica, ou de chaves públicas. A seguir será apresentado um breve conceito de criptografia simétrica e uma forma de distribuição de chaves confiável para comunicação segura entre duas entidades distintas. Após, será visto o conceito de criptografia assimétrica e suas aplicações.

2.1 Criptografia simétrica

Segundo Paar *et al.* [PP11], a melhor maneira de definir a criptografia simétrica de forma bem simples é através do exemplo abaixo.

Tem-se dois usuários, Alice e Bob, que querem se comunicar através de um canal inseguro. O termo canal pode soar um pouco abstrato, mas é apenas uma forma geral para dizer que eles querem se comunicar, por exemplo, pela Internet. O verdadeiro problema começa quando um terceiro usuário com más intenções aparece na situação, Oscar, que deseja roubar informações da comunicação entre Alice e Bob. É claro que existem diversas maneiras de Alice e Bob se comunicarem sem serem ouvidos, entretanto, eles estão em seus escritórios e desejam enviar um ao outro documentos secretos e ninguém, além deles, pode saber do que se trata o conteúdo, como na Figura 1.

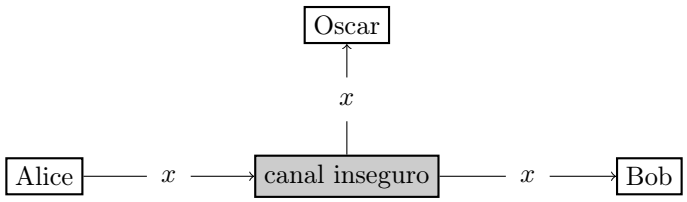


Figura 1: Comunicação insegura.

Uma das soluções para a situação de Alice e Bob é criar uma chave secreta e trocar essa chave por um canal seguro. Assim, usando um algoritmo de criptografia simétrica, Alice poderia cifrar os dados originais com essa chave secreta e enviar os dados cifrados para Bob. Este, conhecendo o algoritmo usado para cifrar, poderia usar a mesma chave já combinada anteriormente, e decifrar o dado para assim ter o texto original de forma segura. Dessa forma Oscar pode até interceptar a mensagem, porém não poderá decifrá-la, pois não tem conhecimento da chave secreta e não teria como decifrar o texto, como na Figura 2.

Se Alice e Bob não tivessem como combinar uma chave secreta por um meio seguro, essa solução não poderia ser utilizada. Em razão disso, começaram a ser elaborados protocolos para troca de chaves, sendo que um dos primeiros foi o protocolo de Diffie–Hellman, que mostra como usar a criptografia simétrica com duas chaves diferentes. Também a partir desse problema e da ideia de Diffie e Hellman, surgiu a ideia da criptografia assimétrica e da distribuição de chaves por uma entidade confiável de forma hierárquica.

2.2 Criptografia assimétrica

Segundo Stallings [Sta16], o conceito de criptografia assimétrica, ou criptografia de chaves públicas, envolve uma tentativa de resolver alguns dos problemas mais difíceis associados à criptografia

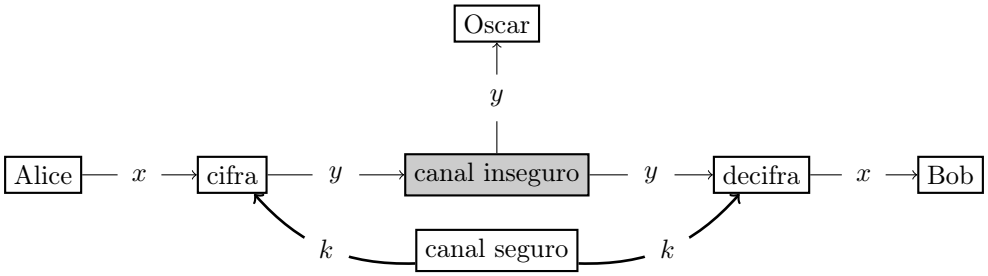


Figura 2: Criptografia simétrica.

simétrica. Um deles é a distribuição de chaves, outro a falta de proteção na comunicação entre duas entidades.

Na criptografia simétrica uma chave precisa ser estabelecida entre as duas entidades por um canal seguro, como foi demonstrado na Figura 2, onde apenas com a existência do canal seguro seria possível ocorrer uma comunicação segura. Mesmo assim, se o problema dessa distribuição de chaves fosse resolvido, o número de chaves existentes iria ser imenso, pois deveria existir uma chave diferente para cada entidade e seus pares comunicantes. Ainda existe outro fator a ser considerado, o não-repúdio.

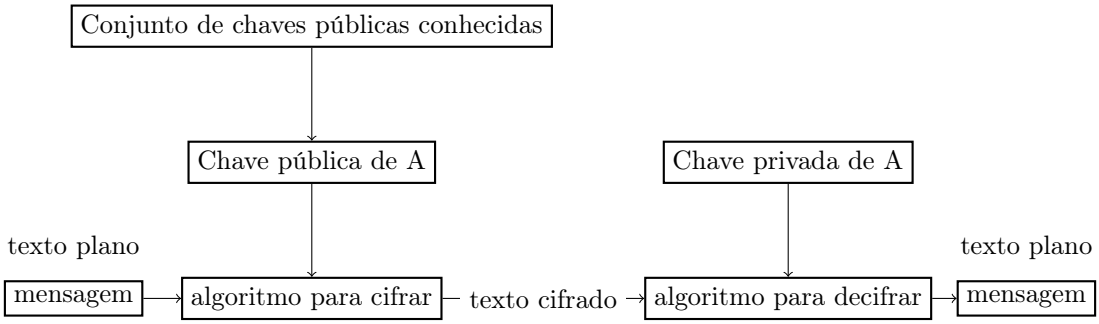
Para resolver esses problemas, Diffie e Hellman [DH76] e Merkle [Mer79] tiveram propostas revolucionárias, baseadas na seguinte ideia: a chave usada para cifrar a mensagem não necessita ser secreta. A parte essencial é que o receptor da mensagem pode decifrar apenas com uma chave privada. Para concretizar isso, o receptor teria de ter publicado uma chave anteriormente para que o emissor pudesse usá-la para cifrar a mensagem. Assim, o receptor teria duas chaves, ou um par de chaves, onde uma é a chave pública e a outra, é a privada [PP11]. Ainda de acordo com Stallings, o algoritmo de criptografia assimétrica traz uma importante característica, a de ser computacionalmente inviável determinar a chave privada pela chave pública relacionada.

Stallings define que um esquema de criptografia assimétrica é baseado em seis ingredientes, como visto na Figura 3.

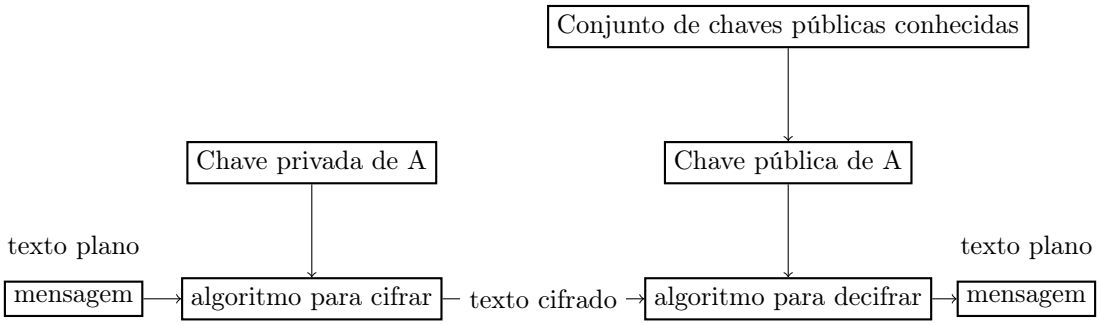
- **texto plano** é a mensagem em sua forma original e legível, também pode ser um dado qualquer. Essa é a entrada para o algoritmo responsável por cifrar.
- **algoritmo para cifrar** faz várias transformações com o texto plano para deixá-lo ilegível.
- **par de chaves** são dois ingredientes da criptografia assimétrica, a chave pública e a chave privada, sendo que uma é usada para cifrar e a outra para decifrar. O algoritmo para cifrar depende da chave usada como entrada.
- **texto cifrado** é a saída do algoritmo usado para cifrar, aqui o texto é totalmente ilegível. Esse texto, ou dado, é totalmente dependente do texto plano e da chave usada para cifrar.
- **algoritmo para decifrar** aceita um texto cifrado e sua chave correspondente para produzir o texto plano.

Existem ainda passos essenciais para esse sistema, que Stallings reporta como segue:

1. Cada usuário gera um par de chaves para cifrar e decifrar mensagens.
2. Cada usuário põe uma das chaves (note que pode ser qualquer uma delas, mas apenas uma)



Sistema de criptografia assimétrica para mensagem secreta



Sistema de criptografia assimétrica para autenticação

Figura 3: Criptografia assimétrica.

em um repositório ou qualquer local acessível ao público. Essa é chamada chave pública. A chave restante é a privada, a qual ninguém deve ter acesso, como foi visto na Figura 3. Cada usuário mantém uma coleção de chaves públicas obtida dos outros usuários.

3. Se uma entidade A desejar mandar uma mensagem confidencial para entidade B, então A precisa usar a chave pública de B para cifrar a mensagem.
4. Quando B recebe a mensagem, ele a decifra usando sua chave privada. Nenhum outro usuário pode decifrar a mensagem porque apenas B conhece sua chave privada.

Com esse procedimento, todos os participantes têm acesso a todas as chaves públicas, e as chaves privadas são geradas localmente por cada participante e não podem ser distribuídas. Enquanto a chave privada de um usuário está intacta, comunicações envolvendo esse usuário são seguras. A qualquer momento, algum dos usuários pode trocar seu par de chaves e substituir a chave pública antiga pela nova.

Stallings ainda cita que a criptografia assimétrica tem três classes de uso:

- **cifrar e decifrar**: o emissor cifra a mensagem com a chave pública do destinatário;
- **assinatura digital**: o emissor assina a mensagem com sua chave privada, visto com mais detalhes na Subseção 2.4;

- **distribuição de chaves**: existem vários protocolos e serviços de trocas de chaves, mas o que será visto nesse desafio será uma estrutura hierárquica para distribuição e controle das chaves, que será detalhada na Subseção 2.6.

2.3 Resumo criptográfico

Segundo Ferguson *et al.* [FSK11], resumo criptográfico, também conhecido por função de resumo criptográfico, é uma função que recebe uma entrada arbitrária de bits, e transforma em uma saída de tamanho fixo. Um uso típico da função de resumo criptográfico é na assinatura digital, onde, ao invés de assinar a mensagem m, que pode ter um tamanho de alguns milhões de *bytes*, pode-se aplicar primeiro uma função de resumo criptográfico e deixar a mensagem m com um tamanho fixo e muito menor, por exemplo, com 256 bits. Essa aplicação $H(m)$, pode ser chamada resumo criptográfico de m e oferece um desempenho muito maior ao processo de assinatura digital.

Uma função de resumo criptográfico, ou simplesmente, uma função *hash* H precisa ter as seguintes propriedades [Sta16]:

- H tem que ser aplicável a um bloco de dados x de qualquer tamanho;
- H precisa produzir uma saída de tamanho fixo;
- $H(x)$ é relativamente fácil de ser computada para qualquer x, fazendo com que as implementações de *hardware* e *software* sejam práticas;
- Para qualquer valor de resumo criptográfico h, é impossível descobrir o bloco de dados x que o gerou;
- Para qualquer x, deverá ser computacionalmente inviável encontrar um bloco $y \neq x$ tal que $H(y) = H(x)$.

O resumo criptográfico, por ser uma forma canônica de representar um bloco de dados de tamanho arbitrário, pode ser usado como uma forma de garantir a integridade de uma mensagem, pois é possível enviar a mensagem por completo a um destinatário e com ela enviar um resumo criptográfico dessa mensagem. Assim, antes de ler a mensagem recebida, o destinatário, conhecendo a função de resumo criptográfico usada, poderia aplicar essa função à mensagem e assim conferir se o resumo criptográfico recém-gerado é o mesmo que o enviado.

Agora, ele poderia ler a mensagem e ter certeza sobre sua integridade. Essa forma de uso do resumo criptográfico acrescentada à criptografia assimétrica é o que fundamenta a assinatura digital. Dessa forma o emissor usa sua chave privada para cifrar o resumo criptográfico da mensagem a ser enviada e o receptor confere a mensagem enviada usando a chave pública do emissor para recriar o resumo criptográfico, e assim ter a certeza que a mensagem é autêntica e íntegra.

2.4 Assinatura digital

Em situações em que não existe confiança mútua entre o emissor e o receptor de mensagens, algo além da autenticação é necessário. A maneira mais atrativa de solucionar isso é com assinatura digital, análoga à assinatura manuscrita [Sta16]. Para tanto, a assinatura digital deve conter as seguintes propriedades:

- deve verificar o autor da assinatura;
- tem que autenticar o conteúdo no momento da assinatura;

- precisa ser verificável por terceiros, para resolver disputas.

Com isso, a assinatura digital inclui a função de autenticação. Stallings diz que com essas propriedades é possível formular os seguintes requisitos para a assinatura digital:

- a assinatura precisa ser um padrão de bits, que depende do conteúdo que está sendo assinado;
- a assinatura precisa usar alguma informação única para o emissor, para prevenir a falsificação do conteúdo e da autoria;
- precisa ser relativamente fácil gerar uma assinatura digital;
- precisa ser relativamente fácil reconhecer e verificar uma assinatura digital;
- precisa ser computacionalmente inviável forjar a assinatura digital, tanto construindo uma nova mensagem para uma assinatura já existente, quanto construindo uma assinatura fraudulenta para uma mensagem qualquer;
- precisa ser prático guardar uma cópia da assinatura digital.

Um método básico para criar uma assinatura digital seria primeiro gerar o resumo criptográfico dos dados a serem assinados e depois o assinante teria de aplicar sua chave privada a esse resumo. Dessa forma seria possível verificar a autenticidade da assinatura e a integridade dos dados assinados. Para tal verificação, basta aplicar a mesma função de resumo criptográfico H aos dados assinados e obter h', que deve ser igual ao resumo criptográfico cifrado pelo assinante.

Para decifrar tal resumo, é necessário conhecer a chave pública do assinante e o algoritmo usado para cifrar. Feito isso, o resultado deve ser o resumo criptográfico h dos dados assinados. Caso $h = h'$ está comprovada a integridade e autenticidade da assinatura digital; caso contrário, essa assinatura foi falsificada e não deve ser confiável. As assinaturas digitais têm muitas aplicações na segurança da informação, incluindo autenticação, integridade de dados e não-repúdio.

Uma das mais importantes aplicações das assinaturas digitais é a certificação de chaves públicas em grandes grupos. Certificação é um meio para uma terceira parte confiável ligar a identidade de um usuário à sua chave pública. Assim, algum tempo depois, outras entidades podem verificar a autenticidade da chave pública sem a necessidade de uma terceira parte [MVvO96]. Isso será abordado com mais detalhes na Subseção 2.5.

2.4.1 Atributos de assinatura

Um atributo é uma estrutura que caracteriza a assinatura, trazendo mais informações para sua validação, contextualização e longevidade. Atributos são definidos de acordo com documentos técnicos, como as RFCs, os ETSIs e o IETF W3C.

Segundo o ETSI EN 319 122-1, atributos são caracterizados como assinados e não assinados. Os primeiros são atributos abrangidos pelo valor da assinatura digital produzida pelo signatário utilizando a sua chave privada, o que implica que o signatário processou esses atributos antes de criar a assinatura. Os atributos não assinados são adicionados pelo signatário, pelo verificador ou por outras partes após a produção da assinatura. Assim, eles não são protegidos pela assinatura, no entanto, podem ser efetivamente cobertos por atributos subsequentes de carimbo de data/hora.

2.4.2 Carimbos de tempo

Um carimbo de tempo é um atributo não assinado, que carrega uma assinatura realizada por uma carimbadora, tendo a função de registrar autenticidade a longo prazo de uma assinatura. Ele indica um momento específico do tempo onde a assinatura encontra-se válida.

O dicionário de Cambridge [Dic] define um carimbo de tempo como um registro de forma impressa ou digital que mostra em qual marco temporal algo aconteceu ou foi feito.

2.5 Certificado digital

Segundo Menezes *et al.*, um certificado digital é uma estrutura de dados que consiste em uma parte de dados e uma parte de assinatura. A parte de dados contém um texto legível, incluindo, no mínimo, a chave pública e um texto identificando a entidade associada a essa chave pública. Nota-se que essa entidade pode ser uma pessoa. A parte da assinatura consiste em uma assinatura digital feita por uma Autoridade Certificadora (AC, explicada na Subseção 2.6.1) sobre a parte de dados da entidade requerente, assim unindo a identidade da entidade com sua respectiva chave pública.

Para gerar um certificado digital, a princípio é necessário seguir os seguintes passos:

1. escolher qual será a AC emissora do certificado;
2. solicitar a essa AC a emissão de um certificado digital, informando os dados do solicitante, tal como seu nome e sobrenome (os dados aqui dependem muito da política usada para a certificação), e esses formarão a parte de dados do certificado digital;
3. em alguns casos, a AC pode solicitar que o usuário solicitante vá até uma autoridade de registro, a qual tem a incumbência de validar os dados informados na solicitação do certificado;
4. caso todos os dados do solicitante estiverem de acordo, a autoridade de registro envia a requisição para a AC, que emite o certificado do solicitante.

A verificação de um certificado digital é feita da mesma maneira que a verificação de uma assinatura digital. Nota-se que a chave pública usada nesse processo será a chave pública da AC emissora do certificado digital. Essa verificação citada pode implicar em um resultado negativo, caso o certificado tenha sido emitido pela AC, mas seja invalidado por alguma razão. Um dessas razões é o certificado ter expirado, ou seja, ter chegado ao fim do tempo previsto de validade. Outra maneira do certificado estar inválido é caso ele tenha sido revogado. Segundo Stallings, existem três possibilidades para a revogação do certificado, que são:

- o usuário avisa que sua chave privada foi comprometida;
- o usuário não é mais certificado pela AC emissora;
- o certificado da AC emissora foi comprometido.

Para manter um registro que algum certificado foi revogado, as ACs guardam essa informação de revogação em uma estrutura chamada lista de certificados revogados (LCR), que está descrita na Subseção 2.6.2.

2.6 Infraestrutura de chaves públicas

Citando Ferguson *et al.*, uma infraestrutura de chaves públicas (ICP) é uma infraestrutura que permite a um usuário reconhecer a quem uma chave pública pertence.

A estrutura clássica de uma ICP é baseada em hierarquia de autoridades, parecida com um grafo no formato de árvore. A ideia aqui é ter um ponto de confiança, que normalmente é uma autoridade certificadora, que por ser de confiança e ser o nó inicial dessa estrutura, é chamada AC-Raiz. Além da AC-Raiz, a autoridade máxima dessa estrutura, existem as ACs intermediárias e ACs finais, essas últimas responsáveis pela emissão dos certificados digitais dos usuários finais dessa infraestrutura.

Quando alguma entidade deseja validar um certificado digital de outra entidade, ela precisa conhecer a AC emissora do certificado de tal entidade. Conhecido isso, a entidade pode verificar a assinatura aposta no certificado usando a chave pública da autoridade emissora. Esse mesmo processo tem que ser realizado para a AC emissora e a AC superior a essa emissora, para verificar se a tal AC emissora é válida.

Isso se repete até que a AC-Raiz seja alcançada, e quando isso ocorre, a verificação termina, visto que a chave pública da AC-Raiz é publicada por outros meios; por exemplo, a AC-Raiz da Infraestrutura de Chaves Públicas Brasileira (ICP-Brasil) é publicada no Diário Oficial da União. Nota-se que esta é uma forma simplificada para esse processo de validação, pois o mesmo envolve outros passos que vão além da verificação das assinaturas nos certificados.

2.6.1 Autoridade certificadora

Segundo Menezes *et al.*, a autoridade certificadora (AC) é uma entidade confiável cuja assinatura sobre o certificado digital comprova que a chave pública pertence à entidade identificada no certificado.

2.6.2 Lista de certificados revogados

As listas de certificados revogados são estruturas em forma de lista emitidas por ACs, de todos os níveis, onde são mantidas informações sobre qual certificado emitido, em algum momento, está revogado. Essas listas são publicadas periodicamente e assinadas pela própria AC emissora.

De acordo com Stallings, a única forma que as ACs tem para identificar um certificado presente na LCR é o *serial number* do certificado, um identificador numérico. Esse é emitido pela AC junto ao certificado de um usuário, e as buscas por certificados revogados dentro da LCR são feitas através desse identificador numérico.

2.7 Políticas de assinatura

As políticas de assinatura são um conjunto de regras e diretrizes designado para caracterizar assinaturas durante todo o seu ciclo de vida. Um regulamento costuma trazer uma definição dos termos utilizados durante sua elaboração, os tipos de assinaturas possíveis entre outras caracterizações.

A identificação de uma política de assinatura é dada por um OID (object identifier) condizente com sua caracterização, como por exemplo, para a política de assinatura CAdES AD-RB em sua versão 2.3 seu OID é **2.16.76.1.7.1.1.2.3**. A presença de uma política de assinatura é dada através do atributo SignaturePolicyID.

De acordo com o ETSI EN 319 122-1, uma política de assinatura é um conjunto de regras para criação e validação de uma assinatura eletrônica, sobre qual assinatura pode ser determinada válida

2.8 Perfis de assinatura

Perfis de assinatura são um conjunto de caracterizações para assinaturas digitais criado e regulamentado pela união europeia. Eles possuem flexibilidade na definição, não necessitando políticas de assinatura e sendo definidos implicitamente pelos atributos contidos na estrutura da assinatura.

Cada assinatura é caracterizada de acordo com a presença de atributos obrigatórios e opcionais, além da não presença de atributos proibidos.

De acordo com o ETSI TS 103 173, um perfil proporciona os recursos necessários básicos para um amplo alcance de uso governamental e empresarial de casos de uso para procedimentos eletrônicos e comunicações quando a uma clara necessidade de interoperabilidade para assinaturas digitais avançadas utilizadas em documentos trocados entre fronteiras.

2.9 Listas confiáveis

Listas confiáveis são documentos que regulamentam quais organizações podem emitir serviços reconhecíveis por empresas, governos ou instituições. Um serviço de confiança é um serviço que aprimora a confidência em transações eletrônicas, como por exemplo, as autoridades certificadoras, que carregam seus certificados nas listas de confiança.

De acordo com o ETSI, listas de confiança são listas que provém informações sobre o status e o histórico de status de serviços de confiança dos provedores de serviços de confiança.

Referências

[DH76] W. Diffie and M. Hellman. New Directions in Cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, September 1976.

[Dic] Cambridge Dictionary. Timestamp. `https://dictionary.cambridge.org/us/dictionary/english/timestamp`. [Accessed 15-10-2024].

[FSK11] N. Ferguson, B. Schneier, and T. Kohno. *Cryptography Engineering: Design Principles and Practical Applications*. 1st edition, 2011.

[Mer79] R. C. Merkle. *Secrecy, Authentication, and Public Key Systems*. PhD thesis, Stanford University, June 1979.

[MVvO96] A. J. Menezes, S. A. Vanstone, and P. C. van Oorschot. *Handbook of Applied Cryptography*. 1st edition, 1996.

[PP11] C. Paar and J. Pelzl. *Understanding Cryptography: A Textbook for Students and Practitioners*. 1st edition, 2011.

[Sta16] W. Stallings. *Cryptography and Network Security: Principles and Practice*. 7th edition, 2016.