Privacy Impact Assessment
Relatório PIA

João Sousa - up202205238

Rui Santos - up202109728

CC2009: Segurança e Privacidade

Prof.º Manuel Correia

Prof.º João Vilela

Prof.º Henrique Faria

Abril de 2024

COP-MODE

PIA

29/04/2024

Sumário Projeto: COP-MODE

Descrição

O que é COP-MODE ?

De uma forma breve, COP-MODE (COntext-aware Privacy protection for MObile DEvices) é um projeto de investigação que tem como objetivo principal o melhoramento da privicidade em dispositivos móveis. Para alcançar estes objetivos a ideia que foi proposto consiste em utilizar uma aplicação, que baseia a atribuição ou não de uma dada permissão através dados de preferencias de vários utilizadores. Com isto, atribuímos uma maior privacidade porque passamos a ter contexto associado a permissão.

Campanha de obtenção de dados

De modo a desenvolver o gestor de privacidade automático foi necessário realizar uma campanha prévia de recolha de dados. Os *PII principal/data subjects* são os utilizadores que irão participar nesta campanha. Também existem *third parties* interessados no processamento de dados pessoais.

Neste caso o nosso grupo vai agir como *PII Controller/Processor* , tendo assim responsabilidade sobre a determinação dos meios e propósitos no processamento de de dados pessoais.

Setup COP-MODE

De seguida vamos explicar passo a passo o setup do COP-MODE.

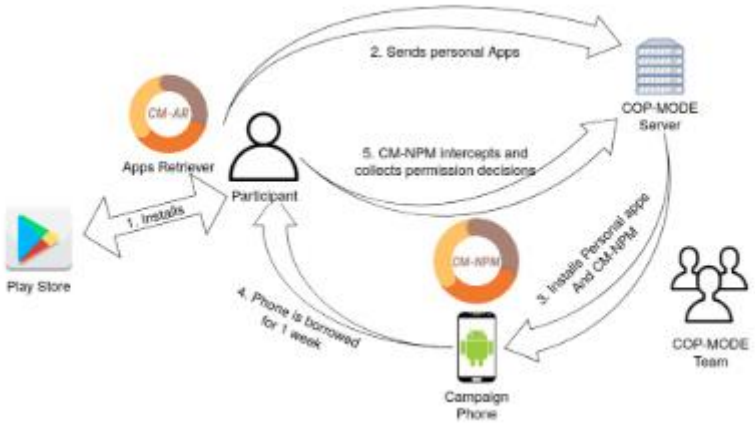

Figura 1 – Metodologia da campanha COP-MODE

1.º Passo : Recrutamento

O primeiro passo começa com o recrutamento de participantes. Um potencial participante deverá baixar e instalar a aplicação COP-MODE Apps Retriever (CM-AR) através da Play Store.

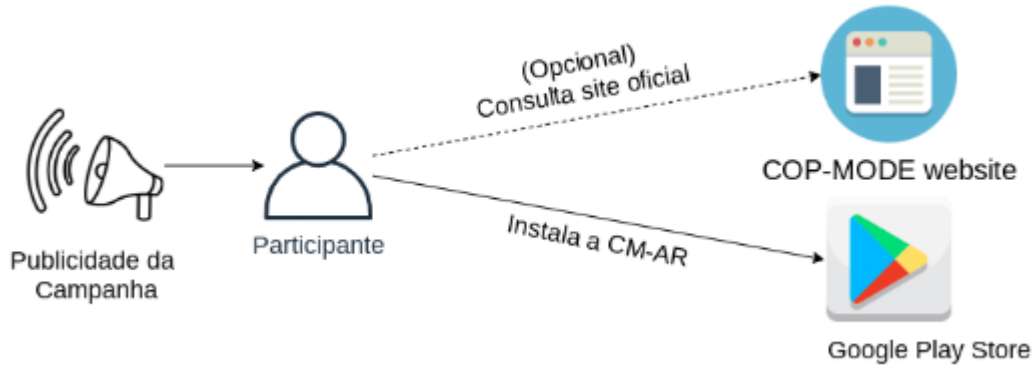

Figura 2 - Fase de recrutamento

2.º Passo : Executando o CM-AR

Após instalar o CM-AR da Google Play Store, o participante deverá executar a aplicação e seguir as instruções. Esta aplicação tratará de enviar o e-mail, data de consentimento e a lista de aplicações instaladas (e respetivas permissões) do smartphone pessoal para o nosso servidor.

A lista de aplicações é utilizada pela equipa do COP-MODE para as instalar no smartphone a emprestar ao participante (ver 3.º Passo) .

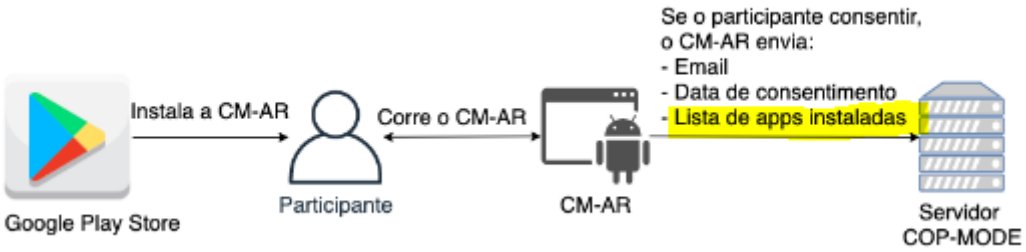

Figura 3 – Fase de execução do CM-AR

3.º Passo : Configuração e Entrega do Smartphone

Figura 4 - Fase de configuração e entrega de smartphone

Após o envio da lista de aplicações para o servidor, a equipa do COP-MODE tratará de instalar as aplicações num dos smartphones da campanha. Este smartphone virá também instalado com o COP-MODE Naive Permission Manager (CM-NPM), um gestor de permissões que tratará de pedir as permissões de acesso das apps ao participante, assim como coletar os dados que precisamos.

4.º Passo : Coleta dos Dados (1 semana)

Figura 5 - Fase de coleção de dados

Durante uma semana, o participante deverá utilizar o smartphone da campanha como sendo o seu smartphone pessoal. Durante este tempo, o gestor de privacidade CM-NPM irá notificando o participante dos acessos às permissões e perguntando se deverá dar ou negar o acesso, coletando a resposta, bem como os dados do contexto.

PII

Tipo/propósito/Categoria da informação coletada/processada

Tipo de coleção	Tipo de informação coletada/processada	Informação é considerado sensível?	Propósito Informação	Informação é categorizada como PII , NSPII ou apenas informação ?
No inicio	Endereço e-mail	✗	Forma de contacto com utilizador	NSPII
Snapshot	Nomes de app's instaladas e respetivas permissões	✔	Configuração do dispositivo que vai ser utilizado por participante na campanha de obtenção de dados	PII
	Tipo de conexão (Wifi , dados móveis, etc ...)	✗	Fornece contexto do utilizador e do próprio dispositivo que o utilizador está a usar, no momento em que o utilizador vai decidir se atribui ou não uma certa permissão.	Informação
	Contexto do dispositivo (em repouso, a ser utilizado, etc ...)	✗		Informação
	Entradas de calendário (identificador de calendário , localização e datas de inicio e fim)	✔		PII
Continua	Localização geográfica	✔		PII
	Dispositivos próximos (Bluethooth , Endereço MAC , etc ...)	✔		PII
No prompt de permissão	Localização semântica (input do utilizador)	✗		Informação
	Decisão do utilizador	✗	Dados que são guardados vão ser utilizados mais tarde para tentar inferir a decisão do utilizador face a um pedido de informação.	Informação

	Informação da aplicação (nome, versão, categoria, visibilidade)	✕	A aplicação que pede uma permissão.	Informação	

Princípios de PII afetados

Os princípios de PII que vamos analisar estão descritos em detalhe no link: Principios de privacidade EU-US . No contexto deste PIA vamos apenas colocar links , juntamente com o nome do princípio , para o artigo relacionados a cada um dos princípios , por forma a obter um PIA mais conciso. Contundo é imperativo que estes artigos e respetivas subalíneas sejam cumpridos ao pormenor.

Princípios de PII afetados (estes princípios aplicam-se as PII presentes neste projeto)
• *Notificação* • *Escolha* • *Responsabilidade pela transferência de dados a terceiros* • *Segurança* • *Integridade dos dados e limitação de fins* • *Acesso* • *Responsabilidade e reforço sobre recursos*

Sistemas e processos envolvidos na gestão da PII

Nos tópicos que se seguem vamos explicitar, de forma detalhada como a PII vai ser:

- Transferida
- Recolhida
- Tratada
- Armazenada

Nota: estamos a supor que nenhuma medida de segurança foi implementada neste momento.

Transferência

A transferência de PII ocorre sobre um canal não protegido http , que permite que informação "circule" entre Participante e servidor COP-MODE . É importante referir que o protocolo http executa a transferência de dados no formato *raw* , ou seja, os dados que passam por este "canal" não tem qualquer tipo encriptação associada.

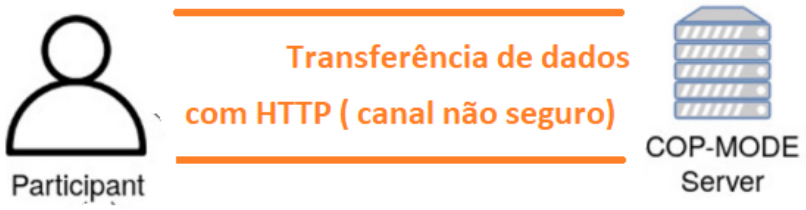

Figura 6 - Transferência de dados entre participante e servidor COP-MODE

Recolhimento

O recolhimento de dados é feito através de ficheiros no formato *JSON e ocorre sempre que o gestor de privacidade CM-NPM notifica os utilizadores dos acessos às permissões (o servidor recebe a informação de contexto e da resposta ao pedido de permissão por parte do utilizador). Com* este formato é possível obter um array de objetos com chaves e valores. Mais uma vez , para evitar sobrecarregar este PIA , decidimos colocar um link com uma descrição detalhada sobre o tipo de informação presentes nestes ficheiros que pode ser acedida aqui.

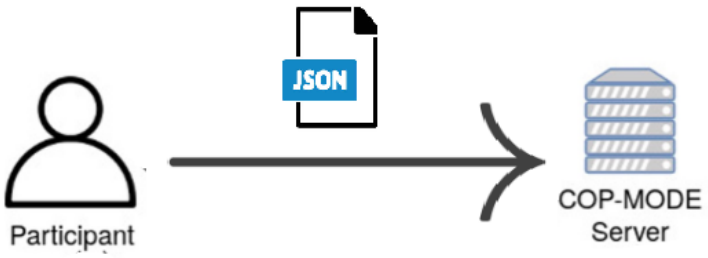

Figura 7 - Recolhimento de dados do participante

Tratamento

Nesta fase inicial os dados não tem qualquer tipo de tratamento associado, ou seja, toda a informação encontra-se no formato *raw* .

Armazenamento

A informação é armazenada no servidor do COP-MODE. Este servidor está configurado para permitir acesso remoto (inicialmente sem proteção qualquer).É importante referir que os dados são guardados no formato *raw* .

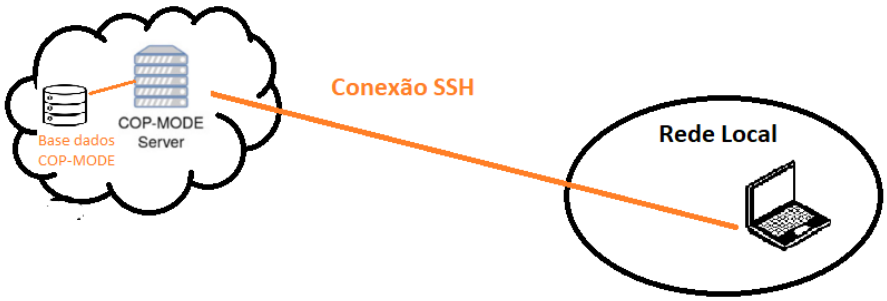

Figura 8 - Representação do acesso ao servidor COP-MODE via SSH (sem qualquer tipo de medida de segurança

Análise de risco e probabilidade de ocorrência

Risco	LIkelihood	Severity
Eavesdroping	Very High	High
Unauthorized server/data access	Very High	High
Data at rest linkage	Very High	High
Sensitive information leakage	Very High	High
Not Signed Email	High	Moderate

Eavesdroping

Eavesdroping - É o risco de segurança que corresponde à interseção da comunicação entre o telemóvel e o servidor, como não existe nenhuma medida de segurança implementada um possível atacante pode intercetar a mensagem, tendo assim acesso a dados sensíveis do utilizador (e.g. localização geografica , dispositivos na proximidade , etc ...), este ataque pode executado por qualquer pessoa não sendo necessário conhecimentos ou equipamentos avançados, logo a probabilidade de ocorrência é muito elevada. Como toda a informação coletada seria exposta este ataque tem severidade alta.

Unauthorized server/data access

Unauthorized server/data access - É um risco correspondente a aceder ao servidor onde os dados dos utilizadores estão guardados, que pela falta de medidas de segurança (principalmente o facto de estarem guardados no formato *raw)* leva a que o risco e a probabilidade de ocorrência associados sejam muito elevados para um ataque deste tipo.

Data at rest linkage

Data at rest linkage - É um risco associado ao facto de terceiros terem o acesso aos dados e com eles serem capazes identificar a pessoa que corresponde aos dados. Como os dados encontram-se sobre o formato *raw* é muito provável isto ser possível sendo um elevado risco à segurança do utilizador.

Sensitive information leakage

Sensitive information leakage – Por todos os motivos acima referidos é muito provável que haja um vazamento dos dados sensíveis do utilizador representado um risco bastante elevado no que consta à segurança do utilizador.

Not Signed Email

Not Signed Email – Consiste no risco associado a uma individuo personificar um utilizador. Como na fase de setup do COP-MODE apenas é pedido o email, nada garante que um e-mail vindo de um dado utilizador é na realidade enviado por esse utilizador. Tendo isto em conta a probabilidade de ocorrer é grande e a severidade é moderada (supondo que o e-mail apenas serve como meio de comunicação de informação não sensível entre utilizador e membros do COP-MODE).

Conclusão da primeira análise risco e probabilidade de ocorrência

Após uma primeira análise , concluímos que existem ameaças significativas à segurança dos dados dos utilizador, que devem ser abordadas, por forma a garantir privacidade aos utilizadores.

Medidas corretivas

Risco	Proposta de solução
Eavesdroping	Encriptação da ligação utilizando o protocolo HTTPS, que permite realizar uma ligação iniciada por cifras assimétrica para estabelecer ligação seguida de cifra simetrica com hash para garantir integridade dos dados e autenticidade
Unauthorized server/data access	• **Acesso ao servidor:** configurar SSH para ter autenticação baseada em chaves. Esta autenticação teria de ser configurada no servidor do COP-MODE e nos dispositivos dos membros autorizados a aceder o mesmo (ver o seguinte link , que contem um exemplo de configuração deste tipo de autenticação). • **Acesso aos dados:** Uma possível solução seria a utilizar a encriptação simétrica, visto que apenas é necessária uma única chave para encriptar e desencriptar. Devemos ter uma chave para cada linha (utilizador) na base de dados. Para encriptar as linhas na base de dados devemos utilizar a respetiva chave. Logicamente é necessário guardar de forma segura todas as chaves (um forma possível seria através de um dispositivo fisico , e.g. PEN com autenticação configurada) . Seria necessário ter mais do que uma PEN com as resptivas chaves, por questões de backup.Por fim estes dispositivos fiscos devem ser atribuidos apenas aos membros do COP-MODE que tenha a respetiva autorização de visualizar os dados originais.
Data at rest linkage	Com a chave de encriptação simétrica associada a cada utilizador, encriptamos toda a informação que consideramos sensível e o respetivo Email. Desta forma apenas os membros que possuam a PEN com as respetivas chaves , tem a capacidade de desencriptar estes dados (os restantes dados permaneceriam no formato *raw*).
Sensitive information leakage	
Not Signed Email	Criar um par de chaves OpenPGP (pública e privada). Na primeira comunicação entre utilizador e servidor COP-MODE, o utilizador

	deve enviar a respetiva chave publica, que é devidamente guardada no respetivo servidor. Sempre que o utilizador quiser enviar e-mail, o mesmo deve assiná-lo utilizando a respetiva chave privada.

Análise de risco (com novas medidas de mitigação)

Risco	LIkelihood	Severity
Eavesdroping	Very High	Very Low
Unauthorized server/data access	Very Low	Low
Data at rest linkage	Very Low	Low
Sensitive information leakage	Very Low	Low
Not Signed Email	Very Low	Very Low

Justificação de análise

Eavesdroping

Com as medidas implementadas, os dados durante a circulação passaram a estar encriptados sem haver a possibilidade de obter alguma informação dos dados tendo por isso uma severidade baixa, mas quando a informação se encontra em transito continua sujeita a evesdroping pois esse é um risco sempre presente havendo uma probabilidade muito alta que aconteça, mas sem que nenhuma informação seja obtida.

Unauthorized server/data access

Com as medidas implementadas, o acesso ao servidor passa a ser feito com autenticação SSH baseada em chaves, sendo que um atacante para ter acesso aos dados teria de adivinhar a respetiva chave SSH , que é muito difícil de adivinhar , e como os dados sensíveis se encontram encriptados o atacante não obteria nenhuma informação sensível. É importante referir que os membros de COP-MODE que vão ter o acesso SSH ao servidor devem ter a respetiva chave guardada localmente no seu computador de uma forma segura.

Data at rest linkage

Com as medidas de encriptação implementadas sobre os dados que permitem identificar um utilizador , a severidade e a probabilidade de ocorrência ficam substancialmente baixos.

Sensitive information leakage

Com as medidas implementadas, é muito pouco provável que informações sensíveis sejam libertadas por estarem encriptadas, mas caso aconteça como não à um grande ganho de informação pelo facto dos pii estarem encriptados e não ser possível saber a qual utilizador eles pertencem.

Not Signed Email

Com a utilização de assinatura digital de e-mail , a probabilidade de ocorrer um ataque deste tipo é muito baixa . Em termos de severidade , esta é muito baixa , visto que os membros do COP-MODE só devem aceitar e-mail's que esteja devidamente assinado por respetivo participante. É importante referir que mesmo que as chaves OpenPGP sejam capturadas por um atacante, o mesmo não tem

qualquer ganho com as mesmas. Contudo cada participante deve guardar a respetiva chave privada de forma segura.

Matriz de risco de privacidade

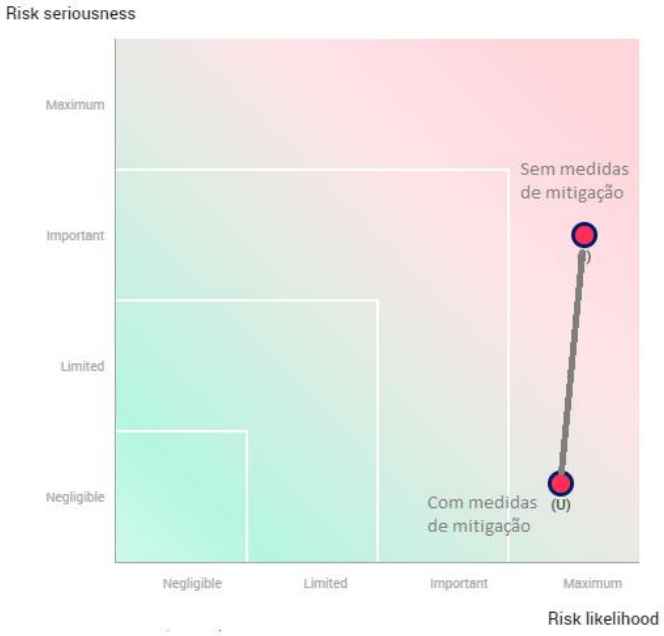

Figura 9 – Matriz de risco para Eavesdroping

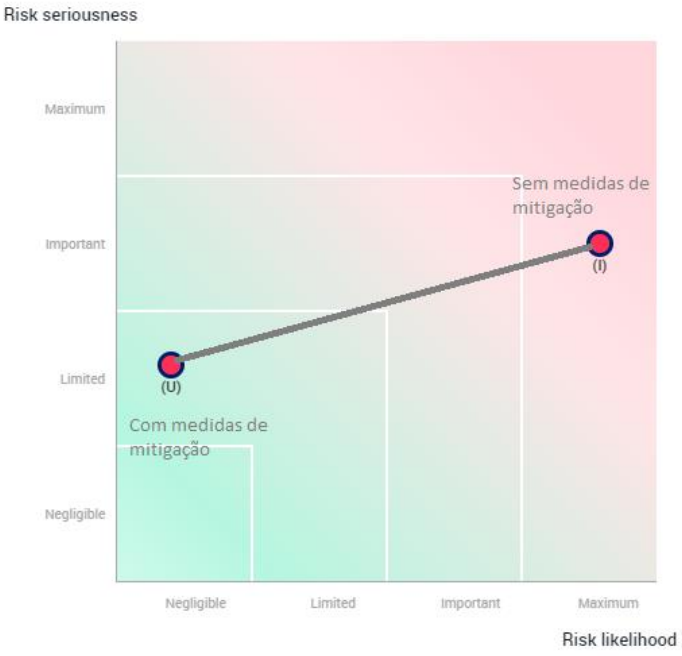

Figura 10 – Matriz de risco para: Unauthorized server/data access , Data at rest linkage e Sensitive information leakage

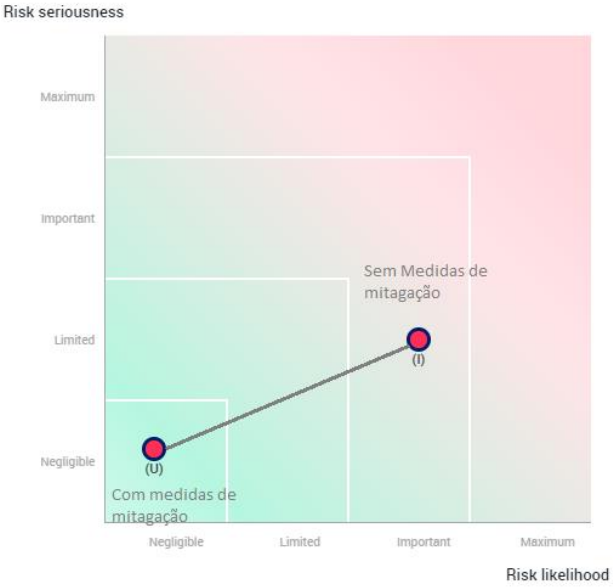

Figura 11 – Matriz de risco para Not Signed Email

Conclusão

As medidas implementadas são as que a nosso ver garantir um melhor *tradeof* de privacidade e usabilidade dos dados sendo que com as medidas implementadas foi possível reduzir significativamente os ricos a privacidade inerentes da realização desta coleta de dados , sendo que se no futuro for descoberta alguma outra vulnerabilidade serão tomadas medidas de imediato para mitigar esses riscos, também será mantido um olho atento a alguma alteração da probabilidade dos ricos conhecidos e mediante alterações significativas novas medidas serão tomadas.

Bibliografia

- https://cop-mode.dei.uc.pt/
- https://energy.ec.europa.eu/document/download/eee93bb8-1bda-4bdc-ac64-7edd6d0e60bc_en?filename=dpia_for_publication_2018.pdf
- https://www.stats.govt.nz/assets/Uploads/Retirement-of-archive-website-project-files/Privacy-Impact-Assessment/Privacy-impact-assessment-for-the-Integrated-Data-Infrastructure/idi-privacy-impact-assessment.pdf
- https://www.privacy.org.nz/publications/guidance-resources/privacy-impact-assessment-toolkit/
- https://www.dataprivacyframework.gov/program-articles/Participation-Requirements-Data-Privacy-Framework-(DPF)-Principles
- https://www.digitalocean.com/community/tutorials/how-to-configure-ssh-key-based-authentication-on-a-linux-server