Problema atual - usuários não estão conscientes de quem pode acessar seus dados em mídias sociais e como esses dados podem ser utilizados.

Foco em dois cenários: violações de privacidade e anonimização de dados.

Violação de privacidade: alguém quer descobrir informações privadas de um indivíduo utilizando dados da rede social possivelmente anônimos.

Anonimização de dados: preservação da privacidade dos usuários numa mídia social.

Dados devem, então, serem providos com um mecanismo que possibilite o acesso aos mesmos sem comprometer a privacidade dos usuários em geral.

Dois tipos principais de networks: user-user e user-group.

Uma mídia social pode ser representada por um grafo G = (V, Ev, H, Eh) onde:

* V é um conjunto de n nodos que representam os perfis dos usuários.
* Ev é um conjunto de arestas que representam uma relação entre dois nodos (amizade). (social link).
* H representa preferências do usuário (grupos, fotos, vídeos, etc).
* Eh representa a afiliação do novo à um grupo. (affiliation link).

**Violações de privacidade em banco de dados:**

Uma violação de privacidade ocorre quando informações privadas de um indivíduo são expostas para alguém que, normalmente, não tenha acesso à mesma, comprometendo a privacidade do usuário. Tradicionalmente, dois tipos de violação são estudados: divulgação de identidade e de atributos. Além desses, o artigo apresenta mais dois: affiliation link e disclosure.

**Identity disclosure:**

Ocorre quando alguém é capaz de determinar o mapeamento do perfil v em uma mídia social de um usuário qualquer p. O adversário geralmente quer uma ou mais das definições:

* Mapping query: verificar o perfil de um usuário, ou seja, numa dada mídia social, achar um perfil que mapeia para um indivíduo em particular.
* Existence query: verificar se a pessoa tem um perifl numa mídia social.
* Co-reference resolution query: para dois perfis quaisquer, verificar se os mesmos se referem a mesma pessoa

A violação de identidade é dada quando o adversário consegue realizar o mapping query com 100% de certeza, o que é complicado pois o adversário deve ter um conjunto de atributos que levam aquele perfil a determinado usuário. Geralmente é trabalhado cálculos probabilísticos.

**Attribute disclosure**

Três tipos de sobreposições de atributos pessoais:

* Atributos de identificação - identificando alguem unicamente.
* Atributos do tipo *quasi-identifying -* combinação que identificam uma pessoal unicamente, como nome e endereço. (juntos).
* Atributos sensíveis - atributos aos quais os usuários podem querer esconder do público, como afiliação política, etc.

A violação de atributos ocorre quando o adversário é capaz de determinar o valor de um atributo sensível, ao qual o usuário tinha a intenção de manter privado.

Se o adversário pode identificar atributos numa rede social, então a resposta para a definição de *mapping query* vista acima é trivial, pois o adversário tem os atributos necessários para verificação de perfil de usuário. (atributos de identificação única).

Atributos do tipo quasi-identifying tem o mesmo comportamento.

**Violação de links sociais:**

Ocorre quando o adversário é capaz de descobrir relações entre dois usuários, a qual deveria ser privada.

Exemplos desses tipos de relacionamentos “sensíveis” podem ser descobertos em mídias sociais, dados de doenças e outros. Ou seja, dependendo do meio que o indivíduo está, podemos ou não inferir “coisas” sobre esse indivíduo. Por exemplo, inferir partido plítico a partir dos amigos do facebook.

**Violação de links de filiação:**

Dois usuários podem ser filiados ao mesmo grupo, o que pode, em alguns casos, levar a uma violação de atributos, de links sociais ou até mesmo de identidade. Portanto, cabe esconder as afiliações para preservar a privacidade dos indivíduos.

Violação de links de filiação podem ser problemáticos, pois, com pequenas buscas num buscador qualquer, pode-se ligar essas buscas, seus resultados, etc à identidade de uma pessoa.

Ex: nos servidores do bittorrent, adversários descobriram padrões de grupos na comunidade baseados em links sociais, e monitorando somente um membro por grupo eles conseguiram inferir nos interesses de todos daquele grupo.

**Definições de privacidade para publicação de dados:**

Preservação de privacidade é considerada uma área muito nova no campo de pesquisa.

Muitas vezes, os dados contêm informações sensíveis que precisam ser “limpas” antes de serem publicadas a fim de preservar esses dados confidenciais.

A sanitização de dados tem o objetivo de remover os atributos de dados que ajudam o adversário a inferir informações sensíveis. Num grafo de dados, como são costumeiramente apresentados, a dificuldade é entender as dependências do grafo e remover informações sensíveis que podem inferir direta ou indiretamente informações sensíveis.

Uma das saídas é realizar a anonimização de dados. Neste processo, as informações de identificação são removidas e outros atributos são perturbados.

Outro tipo de sanitização de dados é prover um mecanismo privado de acesso aos dados, como o uso de algoritmos com preservação de privacidade.

Segundo Chawla, “our privacy is protected to the extent we blend in the crowd”.

Privacidade diferencial:

Conceito que diz que “O risco da privacidade de um indivíduo não deve crescer substancialmente a menina que esse indivíduo participa de um banco de dados”.

**Mecanismos de preservação de privacidade**

Como a preservação de privacidade pode ser alcançada considerando a estrutura da internet. Cada mecanismo que será mencionado foi desenvolvido para contra-atacar um ataque específico.

Mecanismos de preservação de privacidade para mídias sociais:

Considera a anonimização que retira todos os atributos pessoais dos perfis mas mantem as estruturas dos links sociais entre usuários.

Estrutura de anonimização:

Um jeito ingenuo de ver a anonimização é retirando todos os atributos do perfil de usuário e somente manter a estrutura dos links sociais. Isso cria um grafo isomorfo ao original onde não há como o usuário sofrer ataques de violação de atributos e de identidade, preservando a privacidade.

O problema com essa abordagem é o volume de informações importantes perdidas e que, na realidade, isso não garante totalmente a privacidade dos usuários.

As estratégias de anonimização para estruturas de redes sociais estão dispostas em quatro categorias:

* Modificação de borda: aplicada devido a desconectividade que um grafo teria se fossem removidos os vínculos da estrutura. O objetivo do algoritmo é minimizar o número de adição e deleção de bordas.
* Randomização: mecanismo que altera a estrutura do grafo removendo e adicionando bordas de forma randômica e preservando o número total de bordas. Para que o grafo não perca métricas importantes das relações de seus nodos, é geralmente utilizada a técnica de spectrum-preserving randomization, onde o espectro do grafo (conjunto de valores aos quais as propriedades do grafo são relacionadas) é preservado. Preservando esse espectro fica mais fácil termos um guia de randomização na deleção e adição de novas edges.
* Generalização da network: publicando informações agregadas sobre a estrutura das propriedades dos nodos. Os nodos então tem a mesma estrutura de propriedades, então o adversário nao pode distinguir entre os nodos. Se pegarmos subgrafos de um grafo geral, esse subgrafo terá propriedades do grafo generalizado.
* Mecanismos diferencialmente privados: série de algoritmos que garantem que os indivíduos são protegidos sob a definição de privacidade diferenciada. (maximizar a acurácia da query sem identificação de seus registros.)