Privacitat i anonimització de dades

PID_00247389

Jordi Casas Roma

Temps mínim de dedicació recomanat: 4 hores

Universitat Oberta de Catalunya

Índex

Introducció

En aquest mòdul didàctic tractarem del paper que té la preservació de la privacitat en el procés de publicació de dades. En concret, veurem els principals models de preservació de la privacitat, entre els quals destaquem els mètodes d'aleatorització o pertorbació de dades, el model de *k*-anonimitat i el model de la privacitat diferencial.

A continuació, aprofundirem en els mètodes d'emmascarament de dades i el model de *k*-anonimitat, que representa un dels models més coneguts i emprats en els processos d'anonimització. Veurem la teoria subjacent a cada model, a més d'exemples d'aplicació en dades estructurades en format de taules i, també, en dades semiestructurades en format de xarxa o de graf.

L'objectiu d'aquests processos és assegurar la privacitat de les dades dels individus quan es publica informació que conté dades personals. D'una banda, la publicació d'aquestes dades és molt útil per a la recerca que duen a terme institucions, universitats i empreses; però d'altra banda, s'ha d'evitar la violació de privacitat que podrien patir els individus que apareixen en aquests conjunts de dades.

Objectius

En els materials didàctics d'aquest mòdul trobarem les eines indispensables per assimilar els objectius següents:

1. Conèixer les principals amenaces relacionades amb la privacitat dels usuaris que pot comportar la publicació de dades.
2. Conèixer els principals models teòrics que permeten preservar la privacitat dels usuaris en processos de publicació de la informació.
3. Conèixer quins mecanismes d'emmascarament de dades poden evitar o dificultar la tasca d'identificació d'usuaris per part d'un atacant.
4. Identificar quin és el mètode més apropiat per a cada entorn de publicació de dades, considerant l'objectiu del posterior anàlisi per realitzar i el tipus de dades per emprar.
5. Conèixer els principals models d'anonimització en dades estructurades en format de taules.
6. Conèixer els principals models d'anonimització en dades semiestructurades en format de xarxa o graf.

1. Introducció

La mineria de dades (*data mining*, en anglès) és el procés d'extreure informació útil, interessant i desconeguda fins al moment de conjunts de dades. L'èxit de la mineria de dades es basa en la disponibilitat de dades de qualitat sobre les quals executar aquests processos. En aquest sentit, la recopilació de la informació digital per part de governs, corporacions i individus ha de facilitar l'intercanvi i la disponibilitat de dades a gran escala per a la seva anàlisi posterior. Hi ha una demanda d'intercanvi de dades entre diversos actors impulsada, d'una banda, pels beneficis mutus i, d'altra banda, per les regulacions que requereixen que certes dades siguin publicades. Per exemple, als Estats Units es requereix que els hospitals de Califòrnia publiquin les dades demogràfiques específiques de tots els pacients donats d'alta en les seves instal·lacions. Generalment, la publicació de dades en obert (*open data*, en anglès) que incloguin dades personals pot induir a una violació de la privacitat, si no es tracten de forma adequada.

Lectures complementàries

D. M. Carlisle; M. L. Rodrian; C. L. Diamond. (2007, juliol) *California inpatient data reporting manual, medical information reporting for California, 5th edition*. Informe tècnic. Oficina de Desenvolupament i Planificació Sanitària Estatal.

1.1. Antecedents i contextualització

Al juny de 2004, el Comitè Consultiu de Tecnologies de la Informació (*Information Technology Advisory Committee*) dels Estats Units va publicar un informe titulat «Revolucionant l'atenció sanitària a través de les tecnologies de la informació». Un punt clau va ser establir un sistema nacional de registres mèdics electrònics que fomentés l'intercanvi de coneixements mèdics. Exemples similars es poden trobar en pràcticament tots els dominis. Per exemple, Netflix*, un popular servei de lloguer de pel·lícules en línia va publicar un conjunt de dades que conté qualificacions de les seves pel·lícules de cinc-cents mil subscriptors, en un intent per millorar la precisió de les recomanacions de les pel·lícules basades en les preferències personals. De manera similar, AOL**, un conegut cercador de pàgines web que opera als Estats Units, va publicar un conjunt de registres de consultes, però ràpidament es va veure obligat a retirar les dades a causa de la identificació d'un usuari en les dades.

***http://www.netflix.com**
****http://www.aol.com/**

Lectures complementàries

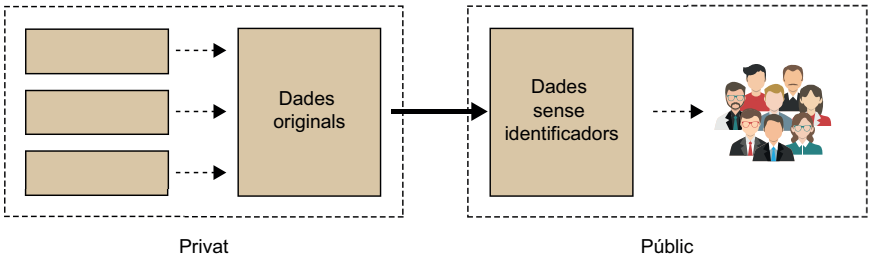
M. Barbaro; T. Zeller (2006, 9 d'agost). «A face is exposed for AOL searcher no. 4417749». *New York Times*.

Una tasca de gran importància és el desenvolupament de mètodes i eines que permetin la publicació de dades, de manera que les dades publicades mantinguin la seva utilitat a la vegada que preserven la privacitat dels usuaris que hi apareixen. Aquest procés s'anomena preservació de la privacitat en la publicació de dades (*privacy-preserving data publishing* o PPDP, en anglès), que pot ser vist com una resposta de caràcter tècnic per a complementar les polítiques de privacitat que implementa cada país o regió.

1.2. Publicació de dades

L'escenari típic de la recopilació i publicació de dades es descriu a la figura 1. En la fase de recollida de dades, el titular de les dades recopila informació dels diferents usuaris. A continuació, el propietari de les dades recopilades —«dades originals» a la figura— ha de protegir i assegurar la privacitat dels usuaris que hi apareixen abans de fer públic el conjunt de dades recopilades. Aquest procés, anomenat *anonimització* o *preservació de la privacitat*, serà l'encarregat d'assegurar que no és possible identificar un usuari dins el conjunt de dades protegides o anònimes. En la fase de publicació de les dades, el propietari de les dades recopilades publica les dades protegides per a la explotació posterior. Aquesta publicació de dades protegides es pot fer de forma pública i accessible per a qualsevol persona o entitat, o bé de forma privada a un conjunt d'empreses o centres autoritzats.

Figura 1. Escenari bàsic per a la publicació de dades



Publicació de dades mèdiques

Per exemple, un hospital recull dades dels pacients i comparteix els registres dels pacients amb un centre mèdic extern. En aquest exemple, l'hospital és el titular de les dades, els pacients són propietaris de les seves pròpies dades i el centre mèdic extern és el receptor de les dades. Les tasques de mineria de dades que el centre mèdic extern pot fer sobre les dades protegides poden ser de qualsevol tipus, des d'un simple recompte del nombre d'homes amb diabetis fins a una sofisticada anàlisi de grups de pacients segons les seves característiques fisiològiques i demogràfiques.

1.3. Preservació de la privacitat en la publicació de dades

Des del punt de vista de la privacitat o anonimització, els atributs d'un conjunt de dades es divideixen en quatre classes, segons el tipus d'informació que contenen:

- Els **identificadors** són un conjunt d'atributs que permeten identificar de forma explícita un individu. El nom, DNI o número de la seguretat social són exemples d'atributs identificadors.

- Els **quasi identificadors** són un conjunt d'atributs que potencialment podrien identificar un individu.

- Els **atributs sensibles** presenten informació específica i sensible d'un individu en concret, com ara les malalties que pateix, el seu salari o les seves preferències sexuals o religioses.

- Finalment, els **atributs no sensibles** són tots els atributs que no tenen cabuda en cap de les categories anteriors.

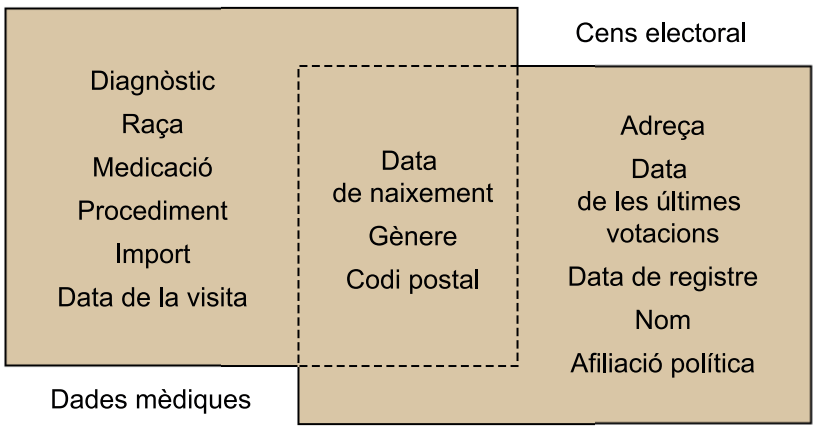
Òbviament, els atributs identificadors han de ser eliminats abans de publicar les dades. En cas contrari, la identificació dels usuaris és directa i trivial en les dades publicades.

Encara que se n'eliminaren tots els identificadors, un estudi de Sweeney l'any 2002 va aconseguir trencar la privacitat d'un governador dels Estats Units. En aquest treball, el nom i altres dades públiques del cens electoral van ser combinades amb una base de dades mèdiques utilitzant el codi postal, la data de naixement i el gènere. La figura 2 mostra la intersecció de dades entre tots dos conjunts. Cap d'aquests atributs no pot utilitzar-se per identificar una persona de manera única, però la seva combinació condueix freqüentment a identificar una única persona o un grup reduït de persones. Aquests atributs són els anomenats quasi identificadors. D'aquesta manera, els autors d'aquest treball van aconseguir identificar les dades mèdiques del governador després de combinar-los amb les dades públiques del cens electoral. Estudis posteriors van mostrar que el 87% de la població dels Estats Units podia ser identificada de manera similar utilitzant informació públicament accessible.

Lectura complementària

L. Sweeney (2002). «Achieving *k*-anonymity privacy protection using generalization and suppression». *International Journal of Uncertainty, Fuzziness, and Knowledge-based Systems* (vol. 10 (5), pàg. 571–588).

Figura 2. Combinació d'atributs quasi identificadors usat per Sweeney en 2002



En l'exemple anterior, la identitat d'un registre es veu compromesa per mitjà de la combinació de diferents quasi identificadors. Per dur a terme aquest tipus d'atacs a la privacitat dels usuaris és necessari que l'atacant tingui cert coneixement extern de l'usuari objectiu. Aquest coneixement pot ser obtingut mitjançant moltes fonts molt diferents. Per exemple, un atacant pot adonar-se que el seu cap ha estat hospitalitzat durant uns dies concrets, per tant pot saber que apareixerà en els registres que es facin públics dels pacients de l'hos-

pital. D'altra banda, no li serà molt difícil descobrir el codi postal, la data de naixement i el gènere de l'individu en qüestió, amb la qual cosa pot efectuar aquest atac per descobrir informació confidencial del seu cap.

La taula 1 presenta un exemple de conjunt de dades mèdiques. Com podem veure, els identificadors han estat eliminats del conjunt, els quasi identificadors contenen els valors originals (codi postal, gènere i edat) i l'atribut sensible es manté inalterable (malaltia), que normalment és la informació més interessant per a l'anàlisi o la mineria de dades, però també és la informació que hem d'evitar que pugui ser vinculada a un usuari concret.

Taula 1. Exemple de taula de dades mèdiques

ID	Codi postal	Gènere	Edat	Malaltia
1	08500	Home	29	Càncer
2	17600	Dona	48	Hepatitis
3	08242	Home	21	Grip
4	25128	Home	36	Càncer
5	17488	Dona	45	Diabetis
6	08401	Dona	58	Grip
7	08760	Dona	72	Hepatitis
8	43840	Dona	22	Càncer
9	43500	Home	25	Diabetis
10	25310	Dona	43	Grip

En aquestes dades d'exemple, un atacant pot identificar un usuari de forma similar a l'atac realitzat per Sweeney en l'any 2002. Utilitzant una combinació dels quasi identificadors n'obtenim registres individuals, que permeten identificar de forma única a un usuari dins el conjunt. Per exemple, si una persona coneguda per l'atacant va ser hospitalitzada durant el període previ a la publicació de les dades i l'atacant sap que és un home, que resideix a la població amb codi postal 08242 i que té vinti-un anys, llavors l'atacant pot saber, sense cap dubte, que el registre metge de l'usuari correspon al registre número 3 de les dades publicades, i que per tant aquesta persona va patir una grip.

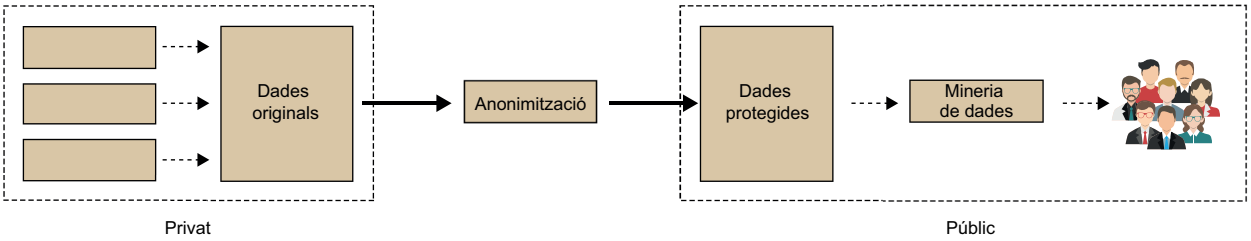
Per evitar aquest tipus d'atacs, el propietari de les dades ha d'aplicar una sèrie d'operacions sobre els quasi identificadors per impedir que se'ls pugui fer servir per identificar qualsevol usuari dins de les dades protegides.

L'anonimització o PPDP pretén ocultar la identitat i la informació sensible dels usuaris que apareixen en conjunts de dades, assumint al mateix temps que la utilitat de les dades ha de ser retinguda a les dades protegides. És a dir, l'anàlisi executada sobre les dades protegides ha de revelar informació útil i veritable, de manera similar als resultats que s'obtindrien en les mateixes anàlisis utilitzant les dades originals (no protegides).

Hi ha multitud d'estratègies d'anonimització, però en general aquest tipus de tècniques busquen les formes d'ocultar els detalls que puguin fer a un individu únic dins del conjunt de dades. L'objectiu és que un únic individu sigui indistingible respecte a un conjunt d'individus prou gran per protegir la se-

va identitat, de manera que l'atacant només pot deduir certa informació amb una certa probabilitat. La figura 3 mostra que, a diferència de l'escenari bàsic de publicació de dades vist anteriorment, en aquest cas s'afegeix la tasca d'anonimització o PPDP prèviament a la publicació de les dades.

Figura 3. Escenari bàsic considerant la preservació de la privacitat en la publicació de dades



2. Models teòrics de privacitat

En aquest apartat veurem alguns dels principals models per a la preservació de la privacitat en processos de publicació de dades. En concret, veurem tres models molt rellevants en tasques d'anonimització, com són l'aleatorització o pertorbació de les dades, el model de *k*-anonimitat i la privacitat diferencial.

2.1. Tipologia de models de protecció

Essencialment, hi ha dos enfocaments principals per limitar el risc de divulgació en processos de publicació de dades:

- **Protecció no interactiva**, mitjançant la qual es genera i s'allibera una versió protegida del conjunt de dades original recopilat dels subjectes de dades.

- **Protecció interactiva**, mitjançant la qual es fa una anàlisi de dades consultada per l'usuari en el conjunt de dades original i, a continuació, es retorna una versió protegida dels resultats a l'usuari.

Quan el tipus d'anàlisi de dades és desconegut en el moment de la protecció i publicació de dades, la protecció no interactiva és l'única solució viable. No obstant això, per a una anàlisi establerta de dades conegudes prèviament, es pot optar per la protecció interactiva, ja que permet ajustar el nivell de protecció a l'anàlisi que s'està fent, cosa que teòricament permet maximitzar la utilitat i la precisió dels resultats.

2.2. Aleatorització

El model d'*aleatorització* o *pertorbació de dades* consisteix, simplement, a introduir soroll en les dades originals, de manera que un atacant no pugui saber, del cert, si la informació que està extraient és certa o ha estat alterada durant aquest procés d'anonimització aleatòria.

D'altra banda, si s'introdueix massa soroll en les dades originals, la privacitat quedarà preservada però la utilitat de les dades pot arribar a ser nul·la. En efecte, quan el grau de pertorbació o introducció de soroll en les dades originals és molt elevat, estem generant dades aleatòries, de manera que el risc de trencar la privacitat serà nul·la, però també serà nul·la la informació que es pot extreure de les dades.

Per tant, un procés d'aleatorització ha d'introduir una quantitat de soroll que:

- sigui suficient perquè un atacant no pugui estar segur de la veracitat d'una dada en concret,

- i, al mateix temps, les dades i la informació general del conjunt de dades s'ha de preservar perquè les anàlisis realitzades sobre el conjunt de dades anònimes siguin tan properes (tant com sigui possible) al resultat obtingut utilitzant el conjunt de dades original.

Exemple de pertorbació aleatòria

Un exemple típic de pertorbació aleatòria en dades numèriques consisteix a afegir soroll als atributs numèrics. La taula 2 mostra un exemple on veiem el salari d'un conjunt de deu individus. Aquest valor, que es considera sensible, ha de ser anonimitzat per preservar la identitat dels individus, ja que si l'atacant coneix el salari d'un dels usuaris que apareix en el conjunt de dades, podria ser capaç de reidentificar l'usuari en el conjunt de dades anònim i, per tant, d'inferir altres dades associades a l'usuari (no mostrades en aquesta taula d'exemple).

La tercera columna mostra una possible configuració de valors després d'afegir soroll aleatòriament per mitjà d'una distribució normal, amb l'objectiu de pertorbar les dades originals i dificultar la possible reidentificació d'un individu.

Taula 2. Exemple de salaris i valors pertorbats mitjançant la introducció de pertorbació aleatòria

ID	Salari ($\times 10^3$)	Valor pertorbat ($\times 10^3$)
1	45	45
2	30	30
3	99	106
4	92	93
5	73	67
6	27	27
7	84	80
8	94	94
9	62	64
10	14	16

És interessant notar que, tot i que els valors han estat modificats, el rang de valors pertorbats es manté pròxim al valor original. Així, el valor mitjà dels salaris en conjunt original és de 62.0, mentre que en les dades pertorbades és de 62.2.

Una de les principals debilitats d'aquest model és que no sol preservar els valors *outliers*, és a dir, valors «extrems» o molt allunyats del rang dels altres valors. Amb la qual cosa, tot i saber que el valor ha estat pertorbat, és possible reidentificar amb una probabilitat molt alta l'usuari 10, ja que el seu valor està molt per sota dels altres.

Lectura complementària

P. Samarati (2001). *Protecting Respondents' Identities in Microdata Release* (vol. 13 (6), pàg. 1010-1027). IEEE Transactions on Knowledge and Data Engineering.

Lectura complementària

L. Sweeney (2002). «*k*-anonymity: a model for protecting privacy». *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* (vol. 10 (5), pàg. 557-570).

2.3. *k*-anonimitat

El model de *k*-anonimitat, introduït per Samarati i Sweeney, és un dels models de protecció no interactiva més àmpliament investigat i emprat en la publicació de dades.

La *k*-anonimitat és una propietat de les dades que garanteix que un individu no pugui ser distingit d'altres $k-1$ individus, també representats en aquestes dades.

Per aconseguir aquest objectiu, podem aplicar tècniques diferents, com ara reemplaçar valors concrets per altres valors d'una categoria més general o eliminar-ne certs valors.

Exemple de *k*-anonimitat

Inicialment, partim d'una base de dades organitzada en forma de taula, com la que es mostra a la taula 3. Aquesta taula mostra la informació referent a un grup de persones que pateixen alguna malaltia. Els quasi identificadors són el seu codi postal, l'edat i la nacionalitat, mentre que la malaltia seria l'atribut sensible que volem protegir. L'objectiu és que qualsevol combinació de quasi identificadors aparegui almenys *k* vegades.

Taula 3. Base de dades no anonimitzada

Quasi identificadors			**Atributs sensibles**
Cod. postal	Edat	Nacionalitat	Malaltia
13053	28	Russa	Arrítmia
13068	29	Espanyola	Arrítmia
13068	21	Japonesa	Infecció
13053	23	Espanyola	Infecció
14853	50	Índia	Càncer
14853	55	Russa	Arrítmia
14850	47	Espanyola	Infecció
14850	49	Espanyola	Infecció
13053	31	Espanyola	Càncer
13053	37	Índia	Càncer
13068	36	Japonesa	Càncer
13068	35	Espanyola	Càncer

La taula 4 mostra la versió *k*-anonimitzada de la taula 3. En aquest exemple, s'ha donat un valor $k = 4$, de manera que cada combinació de quasi identificadors hi apareix quatre vegades. En el cas del codi postal i de l'edat, s'ha utilitzat la tècnica de generalització. En eliminar l'última o les dues últimes xifres del codi postal, aquest segueix aportant informació sobre l'àrea geogràfica de l'individu, però ara es tracta d'una àrea més extensa. Per l'edat s'ha generalitzat indicant únicament límits (superiors o inferiors) d'edat. La tècnica de supressió ha estat aplicada a l'atribut de nacionalitat, ja que no hi ha prou combinacions per garantir *k*-anonimitat per $k = 4$.

Taula 4. Base de dades anonimitzada

Quasi identificadors			**Atributs sensibles**
Cod. postal	Edat	Nacionalitat	Malaltia
130**	<30	*	Arrítmia
130**	<30	*	Infecció
130**	<30	*	Infecció
1485*	>40	*	Càncer
1485*	>40	*	Arrítmia
1485*	>40	*	Infecció
1485*	>40	*	Infecció
130**	<40	*	Càncer
130**	<40	*	Càncer
130**	<40	*	Càncer
130**	<40	*	Càncer

La *k*-anonimitat té els seus avantatges i inconvenients. El seu principal avantatge és que un atacant no pot identificar la seva víctima amb una probabilitat superior a $\frac{1}{k}$, fet que proporciona un límit teòric associat al model de priva-

citat. D'altra banda, hem de considerar que el valor de k en aquest model és l'encarregat d'establir el nivell de privacitat, on en augmentar el seu valor (i el nivell de privacitat), es redueix la utilitat de les dades.

2.4. Privacitat diferencial

La privacitat diferencial va ser introduïda per Dwork com un model de privacitat per a la protecció interactiva en el context de les bases de dades estadístiques, és a dir, per protegir els resultats de les consultes a una base de dades. En aquest context, un mecanisme d'anonimització se situa entre l'usuari, que envia les consultes, i el controlador de base de dades, que les respon. Per preservar la privacitat dels individus, el mecanisme d'anonimització ha de garantir que la contribució de les dades d'un individu al resultat global de la consulta sigui limitada.

Lectura complementària

C. Dwork (2006). «Differential Privacy». A: *International Conference on Automata, Languages ??and Programming* (vol. 4052, pàg. 1–12). Conferència.

Formalment, un algoritme o funció A és ϵ-diferencialment privat si, i només si, per a tots els conjunts de dades D_1 i D_2 que difereixen un sol individu, es compleix:

$$\frac{Pr[A(D_1) \in S]}{Pr[A(D_2) \in S]} \leq e^{\epsilon} \quad (1)$$

on ϵ és un nombre real positiu i $S \subset rang(A)$.

D'acord amb aquesta definició, la privacitat diferencial és una condició en el mecanisme de publicació, no en el conjunt de dades en si. Intuïtivament, això significa que per a qualsevol parell de conjunts de dades similars, és a dir, que es diferencien en un sol individu, un determinat algoritme diferencialment privat es comportarà aproximadament de la mateixa manera en els dos conjunts de dades. La definició dona garanties que la presència o absència d'un individu no afectarà significativament el resultat final de l'algoritme.

La intuïció darrere de la privacitat diferencial ens diu que «la presència o absència d'un individu en un conjunt de dades no ha de modificar significativament els resultats de l'anàlisi». Aquest precepte és molt adequat per a la limitació del risc en la publicació de dades: si les dades d'un individu tenen un impacte significatiu en els resultats d'una anàlisi, probablement la privacitat d'aquest individu estigui en risc. Així, intuïtivament, la privacitat diferencial assegura que les dades estiguin adequadament protegides.

Exemple de privacitat diferencial

Un exemple senzill, però que ens permet introduir el concepte bàsic de la privacitat diferencial, és el següent: suposem que demanem a un grup de persones que respongui a la pregunta «Tens la malaltia X?»

La resposta de l'individu seguirà el procediment següent:

- Llençar una moneda.
- Si surt «cara», llavors l'individu respondrà amb honestedat a la pregunta formulada.
- Si surt «creu», després es llença la moneda de nou i es respon «sí» si surt cara, i «no» si surt creu.

La confidencialitat prové de la refutabilitat de les respostes individuals.

Quan tenim una gran quantitat de respostes, els resultats són significatius, ja que les respostes positives es donen en una quarta part per les persones que no tenen la malaltia X i tres quartes parts per les persones que realment la tenen. Així, si p és la proporció veritable de persones amb la malaltia X, llavors esperem obtenir respostes positives de:

$$\frac{1}{4}(1-p)+\frac{3}{4}p=\frac{1}{4}+\frac{p}{2}$$

Per tant, és possible estimar p sense comprometre la privacitat de cap dels usuaris que responen a la pregunta que els vam formular.

2.5. Resum

En aquest apartat hem descrit la divisió principal que hi ha entre els mètodes de protecció de dades, és a dir, mètodes interactius i no interactius. La pròpia naturalesa del problema i les dades indiquen quina és la millor alternativa en cada cas encara que, com s'ha comentat anteriorment, la protecció no interactiva és l'única solució viable quan es desitja publicar les dades per a qualsevol tipus d'anàlisi posterior, mentre que la protecció interactiva és una bona opció per a una anàlisi establerta de dades coneguda prèviament.

El model d'aleatorització permet treballar amb grans volums de dades, i fins i tot amb dades contínues (*streaming*), atès el seu baix cost de càlcul i espai. No obstant això, no existeixen garanties del nivell de protecció aconseguit per les dades, i especialment en el cas dels valors *outliers* pot ser possible la reidentificació de certs individus.

El model de la k-anonimitat, en canvi, proporciona garanties específiques sobre la probabilitat de reidentificació, que es poden ajustar mitjançant el paràmetre k. Per contra, la seva principal debilitat resideix que és necessari formular el coneixement de l'adversari sobre els quasi identificadors per poder aplicar-hi el model. Addicionalment, el cost de càlcul sol ser molt més alt que en el cas del model d'aleatorització.

Finalment, la privacitat diferencial es basa en un mètode interactiu, fet que permet respondre a certes preguntes sobre la base de dades, però no permet la publicació directa de les dades. No obstant això, malgrat la seva popularitat entre els investigadors i el pas endavant que ofereix en termes de garanties de privacitat, la privacitat diferencial només s'està desplegant en un nombre limitat d'aplicacions del món real, principalment a causa de la baixa utilitat dels resultats oferts.

Com en molts dels problemes relacionats amb les dades, com ara la pròpia mineria de dades, no existeix un algoritme òptim per a tots els problemes i tots els entorns de dades. S'haurà d'analitzar i escollir el millor (o la combinació de diversos) per a cada cas concret.

3. Anonimització de dades tabulars

Tradicionalment, les dades s'han presentat en forma de taules, on cada fila correspon a un registre o a un element (un usuari en el nostre cas) i cada columna correspon a un atribut, una propietat o una característica. Per tant, cada registre té un valor concret per a cadascun dels atributs de la taula. La taula 5 mostra el format típic d'un conjunt de dades en forma de taula. En aquest cas, podem veure un total de n files o registres i m atributs.

Taula 5. Exemple de taula amb n registres i m atributs

Atribut 1	**Atribut 2**	...	**Atribut *m***
$valor_{1-1}$	$valor_{1-2}$	...	$valor_{1-m}$
...	...	...	...
$valor_{n-1}$	$valor_{n-2}$	...	$valor_{n-m}$

En els últims anys, les tècniques de preservació de la privacitat en taules han estat àmpliament estudiades per un grup important d'investigadors de diferents institucions i universitats de tot el món.

3.1. Mètodes d'emmascarament

Els mètodes d'emmascarament permeten modificar les dades originals amb l'objectiu d'impedir o dificultar la identificació d'un usuari en les dades protegides. Aquests mètodes es poden classificar en tres categories bàsiques en funció de com es manipulen les dades originals per definir el conjunt de dades protegides.

- **Mètodes pertorbatius**. El conjunt de dades original és pertorbat d'alguna manera, i el nou conjunt de dades pot contenir informació errònia. Per exemple, es pot introduir soroll en alguns atributs, és a dir, alterar el seu valor de forma més o menys aleatòria. D'aquesta manera, algunes combinacions de valors desapareixen en el conjunt de dades protegides. Alhora, les combinacions en les dades protegides ja no corresponen a les del conjunt de dades original. Aquesta ofuscació dificulta la identificació d'usuaris en el conjunt de dades protegides per part dels atacants.

- **Mètodes no pertorbatius**. La protecció s'aconsegueix mitjançant la substitució del valor original per un altre valor que no és incorrecte però és menys específic, és a dir, més general. Per exemple, reemplacem un nombre per un interval. En general, els mètodes no pertorbatius redueixen el nivell de detall del conjunt de dades. Aquesta reducció del nivell de de-

tall provoca que diferents registres tinguin les mateixes combinacions de valors, la qual cosa dificulta la identificació d'usuaris per part d'un atacant.

- **Generadors de dades sintètiques**. En aquest cas, en lloc de distorsionar les dades originals, es creen noves dades artificials per substituir els valors originals. Formalment, els generadors de dades sintètiques construeixen un model de dades noves a partir del conjunt de dades original i, posteriorment, generen de forma aleatòria un nou (i protegit) conjunt de dades que, si bé segueix les pautes de les dades originals, no conté informació privada de cap usuari. Per exemple, podem substituir l'edat d'un conjunt d'individus per valors aleatoris generats a partir del valor mitjà i la variància observada en les dades originals.

D'altra banda, els mètodes d'emmascarament han de considerar la peculiaritats dels diferents tipus d'atributs per reduir les probabilitats d'identificació mentre es manté la utilitat de les dades protegides. Els atributs es poden classificar en dues categories bàsiques:

Lectura complementària

C. Benjamin; M. Fung; K. Wang; A. Wai-Chee Fu; S. Yu Philip (2011). *Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques*. Estats Units: CRC Press.

- Els **atributs numèrics** permeten de realitzar operacions aritmètiques entre ells com, per exemple, la sostracció o l'addició. Els ingressos i l'edat són exemples típics d'aquests atributs. Pel que fa al risc d'identificació, els valors dels atributs numèrics són propensos a ser únics en una base de dades i, per tant, poden provocar la identificació d'usuaris si no es prenen mesures oportunes per a la seva anonimització.

- Els **atributs categòrics** poden prendre valors en un conjunt finit i les operacions numèriques estàndard no tenen sentit en aquest tipus d'atributs. Podem distingir tres grups principals dins dels atributs categòrics:

– **Nominals**. El valor d'aquests atributs es representa mitjançant etiquetes que proporcionen informació. Per exemple, el color dels cabells o l'estat civil són atributs categòrics nominals.

– **Ordinals**. En aquest cas els atributs presenten un ordre o una escala rellevant entre ells. Per exemple, el nivell d'estudis (primària, secundària, batxillerat, grau, etc.) és un atribut categòric ordinal. En aquests atributs les operacions de mínim i màxim tenen sentit.

– **Estructurats**. Aquests atributs mantenen una relació de classe i subclasse entre ells. Per exemple, les professions poden seguir una jerarquia donada, on per exemple dins de la classe «metge» podem trobar moltes especialitzacions, com ara ginecòleg, pediatre, etc. En alguns casos la jerarquia pot ser explícita i en altres es pot inferir a partir dels valors donats i de les seves relacions.

A continuació, veurem cadascun dels tres tipus de mètodes de camuflament i presentarem alguns mètodes concrets i exemples que faciliten la comprensió

del seu funcionament. Així i tot, una revisió exhaustiva dels mètodes d'emmascarament queda fora de l'abast d'aquest text.

3.1.1. Mètodes pertorbatius

Els mètodes pertorbatius alteren les dades i introdueixen soroll per dificultar el procés d'identificació d'un usuari per part d'un atacant. Hi ha multitud de mètodes en la literatura, encara que una revisió completa de tots ells escapa als objectius d'aquest text i només veurem aquí alguns dels mètodes més comuns i utilitzats en l'actualitat.

El primer mètode que veurem és conegut com *soroll additiu* (*additive noise*) i consisteix a afegir distorsió o soroll en les dades originals, ja sigui seguint o no la mateixa distribució de les dades originals. Un exemple simple d'aquest tipus de mètode és introduir el soroll seguint una distribució normal $N(0,p\sigma)$, on σ representa la desviació estàndard de les dades originals i p és el paràmetre que controla la quantitat de soroll introduït en les dades. Aquest mètode va ser desenvolupat originalment per tractar amb atributs numèrics, encara que posteriorment s'han afegit extensions per poder treballar amb atributs categòrics.

Distribució normal

La distribució normal, també coneguda com distribució gaussiana, és una distribució de probabilitat de variable contínua. La gràfica de la seva funció de densitat té una forma acampanada i és simètrica respecte d'un determinat punt.

La taula 6 mostra una nova versió de l'exemple que hem presentat a la taula 1. En aquest cas, hem aplicat el mètode *additive noise* a l'atribut «edat», introduint soroll en aquest atribut i disminuint, per tant, la probabilitat que un atacant pugui identificar un usuari dins de les dades protegides utilitzant informació externa.

Taula 6. Dades mèdiques després d'aplicar *additive noise* en l'atribut «edat»

ID	Codi postal	Gènere	Edat	Malaltia
1	08500	Home	32	Càncer
2	17600	Dona	46	Hepatitis
3	08242	Home	28	Grip
4	25128	Home	38	Càncer
5	17488	Dona	48	Diabetis
6	08401	Dona	61	Grip
7	08760	Dona	59	Hepatitis
8	43840	Dona	20	Càncer
9	43500	Home	24	Diabetis
10	25310	Dona	55	Grip

El segon mètode que veurem, i que és àmpliament conegut i utilitzat per empreses i administracions, és conegut com *microagregació* (*microaggregation*). Aquest mètode es basa en crear grups de dades segons la seva similitud, i posteriorment reemplaçar el valor original de cada dada pel valor mitjà de tots els valors del grup al qual pertany. Per tant, per a cada valor específic d'un o més atributs existiran sempre un conjunt de registres; mai un registre únic que permeti que un usuari sigui identificat. Aquest mètode es pot aplicar sobre un

únic atribut, i llavors es coneix com a microagregació univariant, o sobre dos o més atributs, i en aquest cas es coneix com a microagregació multivariant.

El mètode permet decidir quants registres s'ajunten en un mateix grup. És important notar que com més registres s'ajunten en cada grup, més gran és el nivell de privacitat que s'aconsegueix, però també és més gran el nivell de soroll. Igual que el mètode anterior, aquest també va ser desenvolupat per atributs numèrics, encara que posteriorment també s'han desenvolupat extensions que permeten tractar amb atributs categòrics.

La nova versió de l'exemple, després d'aplicar microagregació univariant sobre l'atribut «edat» es pot veure a la taula 7. En aquest cas hem parametritzat el mètode perquè creï grups de dos registres. Com podem veure a la taula, per a cada possible valor de l'atribut «edat», sempre hi ha dos o més registres amb el mateix valor. Per exemple, en el conjunt de dades original mostrat a la taula 1, l'edat dels usuaris de la fila 1 i 9 és de vint-i-nou i vint-i-cinc anys, respectivament. Després del procés aplicat, s'ha assignat l'edat de vint-i-set anys a tots dos usuaris. D'aquesta manera, no és possible que un atacant pugui identificar de forma única un usuari utilitzant la informació dels atributs que hem utilitzat per crear els grups, en aquest cas concret, l'edat dels individus.

Taula 7. Dades mèdiques després d'aplicar *microagregació univariant* a l'atribut «edat»

ID	Codi postal	Gènere	Edat	Malaltia
1	08500	Home	27	Càncer
2	17600	Dona	46	Hepatitis
3	08242	Home	21	Grip
4	25128	Home	39	Càncer
5	17488	Dona	46	Diabetis
6	08401	Dona	65	Grip
7	08760	Dona	65	Hepatitis
8	43840	Dona	21	Càncer
9	43500	Home	27	Diabetis
10	25310	Dona	39	Grip

Per acabar amb els mètodes pertorbatius, veurem un tercer mètode que també és molt conegut i utilitzat en entorns empresarials i governamentals. El mètode, conegut com *intercanvi de rang* (*rank swapping*, en anglès), es basa en intercanviar aleatòriament els valors d'un mateix atribut entre diferents registres. Per aconseguir que les dades no siguin excessivament pertorbades, aquest mètode ordena tots els valors de l'atribut presents a la taula, i a continuació l'intercanvi es fa entre valors que es troben dins d'un rang acotat i definit com a paràmetre del mètode. D'aquesta manera s'intenta minimitzar el soroll i mantenir la utilitat de les dades protegides. Aquest mètode pot ser aplicat a atributs numèrics i categòrics ordinals.

La taula 8 mostra els resultats d'aplicar el mètode *rank swapping* sobre l'atribut «edat». Com es pot veure, els valors de l'atribut han estat intercanviats

entre els registres amb valors pròxims. Per exemple, si ordenem els valors de l'atribut «edat» de les dades originals, mostrats en la taula 1, obtenim el vector {21,22,25,29,...}. Si apliquem aquest mètode en un rang de només dues posicions n'obtenim el vector {22,21,29,25,...},que correspon a les dades presentades a la taula 8. Els dos primers valors corresponen als registres 3 i 8 que, com es pot apreciar a la taula, han intercanviat els seus valors. De manera similar, els registres 1 i 9 també han intercanviat els seus valors durant el procés d'anonimització. I així successivament tots els registres de la taula. Com més gran és el rang que s'utilitza en l'intercanvi, més gran és el grau de privacitat i més gran és, també, la distorsió de les dades originals.

Taula 8. Dades mèdiques protegides aplicant *rank swapping* en l'atribut «edat»

ID	Codi postal	Gènere	Edat	Malaltia
1	08500	Home	25	Càncer
2	17600	Dona	45	Hepatitis
3	08242	Home	22	Grip
4	25128	Home	43	Càncer
5	17488	Dona	48	Diabetis
6	08401	Dona	72	Grip
7	08760	Dona	58	Hepatitis
8	43840	Dona	21	Càncer
9	43500	Home	29	Diabetis
10	25310	Dona	36	Grip

3.1.2. Mètodes no pertorbatius

A diferència dels mètodes vistos en el subapartat anterior, els mètodes que veurem en el proper no introdueixen soroll o distorsió en les dades originals. La informació protegida que es publica continua sent totalment veritable, encara que es generalitzen o suprimeixen algunes parts de la informació que podrien ajudar un atacant a identificar de forma única un usuari dins de les dades protegides.

Distingirem dos mètodes bàsics, d'una banda la generalització d'atributs, i per una altra la supressió d'atributs.

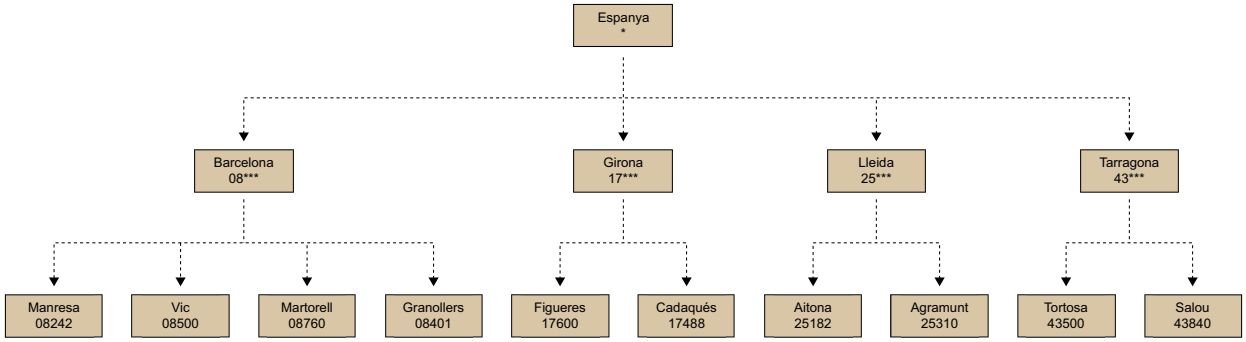
El mètode de **generalització** s'aplica normalment a atributs categòrics, encara que es pot aplicar a atributs numèrics sense cap problema. En aquest cas les dades es protegeixen reemplaçant un conjunt d'atributs per un valor més general que els inclogui tots. Per tant, aquest mètode no introdueix soroll o falseja les dades, simplement les fa més generals o menys específiques, de manera que les individualitats de cada usuari del conjunt de dades original queden difuminades entre els altres usuaris en el conjunt de dades protegit.

En el cas dels atributs numèrics, l'agregació es pot implementar mitjançant la construcció de rangs de valors. Per exemple, si tenim registres amb els valors 3,11,7,19 podem crear dos grups i associar un rang concret a cada grup. En

aquest cas, podríem crear els rangs [0,10) i [10,20), i en aquest cas el primer i el tercer registre tindrien el mateix valor [0,10) en les dades protegides.

D'altra banda, en el cas dels atributs categòrics, cal disposar de la jerarquia dels atributs, ja sigui de forma implícita o explícita. La figura 4 mostra la jerarquia de l'atribut «codi postal». Podem veure que els diferents municipis s'agrupen en províncies, i aquestes, al seu torn, s'agrupen formant l'arrel de tota l'estructura, que en aquest cas seria el nivell de país. Veiem que podríem haver inclòs un nivell extra que fes referència a les comunitats autònomes, però s'han obviat per mantenir la simplicitat de l'exemple. D'aquesta manera, si volem generalitzar la informació sense pertorbar-la, podem generalitzar el codi postal a nivell de província, de manera que les dades són absolutament correctes, però més generals que les dades originals.

Figura 4. Jerarquia de l'atribut «codi postal»



La taula 9 mostra el resultat d'aplicar la generalització sobre els atributs «codi postal» i «edat». En el primer cas, com es pot observar, s'han generalitzat tots els valors a nivell de província, de tal manera que no és possible identificar un municipi de forma única. En el segon cas, l'atribut «edat» ha estat generalitzat fent servir rangs de valors, de manera que en lloc d'indicar que l'edat de l'individu que apareix en el registre 1 és vint-i-nou anys, en les dades protegides indiquem que l'edat de l'individu està en el rang comprès entre vint i trenta anys.

Taula 9. Dades mèdiques protegides aplicant *generalització* en els atributs «codi postal» i «edat»

ID	Codi postal	Gènere	Edat	Malaltia
1	08***	Home	[20,30)	Càncer
2	17***	Dona	[30,50)	Hepatitis
3	08***	Home	[20,30)	Grip
4	25***	Home	[30,50)	Càncer
5	17***	Dona	[30,50)	Diabetis
6	08***	Dona	[50,80)	Grip
7	08***	Dona	[50,80)	Hepatitis
8	43***	Dona	[20,30)	Càncer
9	43***	Home	[20,30)	Diabetis
10	25***	Dona	[30,50)	Grip

El mètode de **supressió** d'atributs consisteix, simplement, a eliminar els atributs. Quan no és possible utilitzar la generalització o un altre mètode i l'atribut pot presentar una bretxa de privacitat, s'utilitza aquest recurs per eliminar el valor de l'atribut i indicar que el valor ha estat suprimit.

3.1.3. Generació sintètica de dades

En els últims anys, ha sorgit una nova tendència per a la preservació de la privacitat de les dades. Consisteix en la publicació de dades sintètiques en lloc de les dades originals. La idea principal és que les dades sintètiques no poden posar en perill la privacitat dels individus, ja que les dades no són «reals». Els mètodes de generació de dades sintètiques consisteixen a construir un model de les dades i a continuació generar dades artificials utilitzant el model generat.

Aquest enfocament presenta dues dificultats: d'una banda, encara que les dades són sintètiques, la identificació d'usuaris és encara possible i, per tant, el risc de divulgació ha de ser analitzat per a aquest tipus de dades. D'altra banda, com les dades sintètiques es generen a partir d'un model de dades particular, construït a partir de les dades originals, tots aquells aspectes que no estan inclosos explícitament en el model no són inclosos en les dades. A causa d'això, l'anàlisi de les dades protegides podrien portar a resultats sensiblement diferents als quals portarien les dades originals.

Hi ha multitud de mètodes per a la generació de dades sintètiques. La distorsió de dades per mitjà de la probabilitat de distribució (*data distortion by probability distribution*) va ser un dels primers procediments de protecció que poden ser classificats com a generació sintètica de dades. Aquest procediment es defineix pels tres passos següents:

Funció de densitat de probabilitat

En la teoria de la probabilitat, la funció de densitat de probabilitat d'una variable contínua descriu la probabilitat relativa segons la qual aquesta variable aleatòria prendrà un valor determinat.

- En primer lloc, identificar una funció de densitat de probabilitat subjacent a les dades i determinar-ne els paràmetres.
- A continuació, generar sèries distorsionades mitjançant la funció de densitat de probabilitat estimada.
- Per acabar, se substitueix la sèrie original per la sèrie distorsionada en les dades protegides.

El procediment es va definir originalment per a funcions de densitat univariant (aplicat sobre una variable, en el nostre cas un atribut), encara que també pot ser aplicat a funcions de densitat de multivariants (més d'un atribut).

3.2. *k*-anonimitat en taules

El model de *k*-anonimitat no és un mètode d'emmascarament o de protecció, sinó un model o una condició que ha de ser satisfet pel conjunt de dades protegit. Tot i això, generalment aconseguim complir la *k*-anonimitat per mitjà dels mètodes de protecció o d'emmascarament que hem vist en els subapartats anteriors. És a dir, sobre el conjunt de dades originals se'ls aplica un o més dels mètodes d'emmascarament vistos anteriorment, amb la finalitat d'aconseguir un conjunt de dades protegit que compleixi les restriccions o condicions necessàries per al model de la *k*-anonimitat.

Un conjunt de dades compleix el model de la ***k*-anonimitat** si, i només si, per a qualsevol combinació d'atributs quasi identificadors existeixen *k* o més registres que comparteixen els mateixos valors.

En altres paraules, cada registre en un conjunt de dades *k*-anònim és indistingible de, com a mínim, altres $k-1$ registres pel que fa al conjunt de quasi identificadors. Per tant, la probabilitat d'identificació d'un usuari en un conjunt de dades *k*-anònim pel que fa als quasi identificadors és de, com a màxim, $\frac{1}{k}$.

La taula 10 mostra el conjunt de dades d'exemple 2-anònim, és a dir, anonimitzat per complir amb la *k*-anonimitat amb un valor de $k = 2$. Segons indica el model, per a cada possible conjunt d'atributs quasi identificadors (codi postal, gènere i edat) sempre hi ha, almenys, dos registres amb els mateixos valors. Per tant, la probabilitat d'identificació d'un usuari és de $\frac{1}{2}$ com a màxim. Com es pot apreciar a la taula, per l'emmascarament de les dades hem aplicat la generalització en el «codi postal», supressió al «gènere» i microagregació univariant en la «edat».

Taula 10. Dades mèdiques protegides que compleix amb el model de *k*-anonimitat amb valor $k = 2$

ID	Codi postal	Gènere	Edat	Malaltia
1	08***	Home	25	Càncer
2	17***	Dona	46,5	Hepatitis
3	08***	Home	25	Grip
4	25***	*	39,5	Càncer
5	17***	Dona	46,5	Diabetis
6	08***	Dona	65	Grip
7	08***	Dona	65	Hepatitis
8	43***	*	23,5	Càncer
9	43***	*	23,5	Diabetis
10	25***	*	39,5	Grip

Després del procés d'anonimització, la taula pot ser publicada amb la certesa que un atacant només podrà identificar un usuari amb una probabilitat de, com a màxim, $\frac{1}{2}$. És important subratllar que en els casos reals, on els conjunts de dades són molt grans (amb milers o milions de registres) s'ha d'escollir un

valor de k molt més gran que en l'exemple que acabem de presentar, on la probabilitat d'identificar un individu serà molt més baixa, és a dir, $\frac{1}{k} << 1$.

3.3. Privacitat diferencial en taules

La privacitat diferencial assumeix la presència d'un actor de confiança que conté el conjunt de dades, rep consultes enviades pels usuaris i retorna resultats per a aquestes consultes.

Suposem que hi ha una base de dades amb informació sobre el salari mitjà d'una població o d'una zona geogràfica específica. A més, sabem que en Pere es trasllada a aquesta zona d'aquí a poc, amb la qual cosa un atacant pot fer la consulta abans i després que sigui incorporat el registre amb la informació del Pere. La privacitat diferencial preveu l'obtenció d'informació relacionada amb el salari d'aquest individu, assegurant que ningú no pugui provar que l'individu en qüestió estigui dins o fora de la base de dades.

Considerem dues bases de dades o taules: la inicial D_1, sense la informació del Pere, que es pot veure a la taula 11; i D_2, amb la informació del Pere, que es pot veure a la taula 12.

Taula 11. Exemple de la base de dades inicial (D_1)

ID	Salari (USD)
1	50.000
2	58.000
3	72.000
4	59.000
5	68.000

Com podem veure, l'única diferència entre D_1 i D_2 és la incorporació de les noves dades, és a dir, només difereixen en un registre. En aquest cas, parlarem de *bases de dades adjacents*.

Taula 12. Exemple de la base de dades amb la informació del Pere (D_2)

ID	Salari (USD)
1	50.000
2	58.000
3	72.000
4	59.000
5	68.000
6	110.000

Perquè la base de dades sigui diferencialment privada, ens cal seleccionar una funció aleatòria, un mecanisme M, que afegeixi soroll als conjunts de dades que produiran un resultat R.

Com D_1 i D_2 són adjacents, la probabilitat que $M(D_1) = R$ és propera a la probabilitat que $M(D_2) = R$. Més formalment podem escriure:

$$\frac{P[M(D_1) = R]}{P[M(D_2) = R]} < e^{\epsilon} \tag{2}$$

Per a ϵ petit, observem que $e^{\epsilon} \approx 1 + \epsilon$ i, si les probabilitats són idèntiques, n'obtenim:

$$1 - \epsilon < \frac{P[M(D_1) = R]}{P[M(D_2) = R]} < 1 + \epsilon \tag{3}$$

La quantitat i el tipus de soroll que M afegeix està limitat per la sensibilitat global de la funció de consulta, f, que s'aplicarà a les dades. La sensibilitat global es pot escriure:

$$\Delta f = max[f(D_1) - f(D_2)] \tag{4}$$

per a tots els possibles conjunts de dades adjacents.

Si considerem una consulta de recompte, llavors $\Delta f = 1$, atès que dos conjunts de dades adjacents poden ser diferents en un registre com a màxim. Dwork va demostrar que el soroll amb una distribució laplaciana manté la privacitat diferencial si el valor del soroll laplacià és ajustat amb el paràmetre $b = \frac{\Delta f}{\epsilon}$. Per tant, una base de dades en la qual s'aplica la consulta de recompte serà diferencialment privada si utilitza un mecanisme aleatori que afegeix soroll laplacià amb $b = \frac{1}{\epsilon}$.

La pregunta següent és quin valor de ϵ hem de seleccionar. Com més gran sigui b, més soroll caldrà afegir a la resposta per aconseguir-hi la privacitat diferencial. Per tant, un ϵ més petit hi implica més soroll.

A mesura que augmentem Δf (major sensibilitat global), necessitem valors de ϵ més petits per proporcionar suficient soroll. Considerem una funció de consulta que calcula el salari mitjà. En aquest cas, la sensibilitat global és igual al salari més alt possible en els conjunts de dades (és a dir, el pitjor escenari).

No obstant això, malgrat la seva popularitat entre els investigadors i el pas endavant que ofereix en termes de garanties de privacitat, la privacitat diferencial només s'està emprant en un nombre limitat d'aplicacions del món real. La principal raó és la utilitat pobra dels resultats obtinguts mitjançant models diferencialment privats. A excepció d'una sèrie d'aplicacions de bon comportament, la privacitat diferencial té un impacte considerable en la utilitat de les dades per ser àmpliament utilitzat en l'anàlisi o la mineria de dades.

A diferència dels mètodes o algoritmes basats en aleatorització o *k*-anonimitat, que es poden aplicar sobre (gairebé) qualsevol conjunt de dades i independentment de la tasca per realitzar, en el cas de la privacitat diferencial es requereix un mètode o algoritme específic per a cada tipus de consulta que volem fer a la base de dades. Exemples de consultes poden ser: valors mitjans de certs camps (per exemple, salaris) o histogrames de valors (per exemple, edats de la població), entre molts d'altres.

3.4. Resum

En aquest apartat hem revisat els principals mètodes d'emmascarament de dades tabulars, que permeten modificar les dades originals amb l'objectiu d'impedir o dificultar la identificació d'un usuari en les dades protegides.

- Els mètodes pertorbatius es basen en la idea d'introduir soroll en les dades originals per dificultar el procés de reidentificació, essent el soroll additiu i la microagregació dos dels principals mètodes d'aquesta categoria.

- En segon lloc, hem presentat els mètodes no pertorbatius que, en comptes d'introduir soroll en les dades, introdueixen un cert grau «d'incertesa», ja sigui per mitjà del procés de generalització o bé de supressió de la informació. És a dir, el seu objectiu és presentar dades menys precises que impedeixin la reidentificació dels usuaris, sense introduir, en cap cas, dades pertorbades o falses en el conjunt de dades anònimes.

- La generació de dades sintètiques permet generar un nou conjunt de dades amb propietats similars al conjunt original. Aquest procés no és trivial, i moltes vegades requereix certa informació sobre els objectius de les tasques de mineria de dades que s'aplicaran a continuació, per tal d'ajustar les dades sintètiques per maximitzar la utilitat del conjunt de dades anònim.

A continuació, hem vist com algunes de les tècniques perturbatives i no perturbatives permeten d'aplicar el model de *k*-anonimitat en dades tabulars. El model de *k*-anonimitat és un dels models més importants, més estudiats i més utilitzats en la preservació de dades tabulars.

Finalment, hem introduït el model de la privacitat diferencial per a dades tabulars. Cal destacar la diferència d'enfocament amb els mètodes anteriors, ja que aquest model permet aconseguir un alt grau de privacitat en entorns interactius. No obstant això, com hem discutit, ha rebut certes crítiques sobre la utilitat de les dades anònimes que fa que, fins al moment, només sigui aplicable a un conjunt reduït de problemes reals.

4. Anonimització de xarxes i grafs

En els últims anys la representació de dades en format de xarxa ha experimentat un important auge en tots els nivells. Aquest format permet representar estructures i realitats més complexes que les tradicionals dades relacionals, que utilitzen el format de tuples. En un format semiestructurat cada entitat pot presentar, igual que les dades relacionals, una sèrie d'atributs en format numèric, nominal o categòric. Però a més, el format de xarxa permet representar d'una manera més rica les relacions que puguin existir entre les diferents entitats que formen el conjunt de dades. Un exemple clar d'aquesta situació el presenten les xarxes socials.

La literatura utilitza els termes *xarxa* i *graf* de manera indistinta. Generalment podem trobar referències a xarxes o grafs gairebé sense matisos, sense diferències importants en el seu significat. En aquest text s'utilitzen els termes xarxa i graf indistintament. Tot i això, alguns autors parlen d'una subtil diferència entre les dues terminologies. Per exemple, A. Barabási assenyala que la terminologia de xarxa (*network*) s'utilitza sovint per referir-se a sistemes reals, mentre que la terminologia de graf (*graph*) s'utilitza generalment per referir-nos a les representacions matemàtiques de les xarxes.

Lectura complementària

A. Barabási; M. Pósfai (2016). *Network science*. Cambridge: Cambridge University Press.

4.1. Introducció a la teoria de xarxes i grafs

En aquest subapartat introduirem la definició i notació bàsica de la teoria de grafs. Els grafs són la forma més natural de representació de les xarxes reals, i és en aquest sentit en què necessitem introduir els conceptes bàsics per poder representar les xarxes reals.

Un graf és una parella de conjunts $G = (V,E)$, on $V = \{v_1,v_2...,v_n\}$ és el conjunt de nodes o vèrtexs i $E = \{e_1,e_2,\ldots,e_m\}$ és un conjunt d'arestes que uneixen dos nodes $e_i = \{v_i,v_j\}$ de forma bidireccional, és a dir, el node v_i està connectat al node v_j i viceversa. En aquest cas, parlem de **grafs no dirigits**, **bidireccionals** o **simètrics**.

Quan les relacions no són bidireccionals, parlarem de **grafs dirigits**, **unidireccionals** o **asimètrics**. En aquest cas, es representa el graf com a parella de conjunts $G = (V,A)$, on és el conjunt de nodes o vèrtexs, igual que en el cas anterior, i $A = \{a_1,a_2,\ldots,a_m\}$ és un conjunt d'arcs que uneixen dos nodes $a_i = \{v_i,v_j\}$ de forma unidireccional, és a dir, el node v_i està connectat al node v_j.

S'anomena **ordre** de G el seu nombre de nodes, $|V|$, que per conveni és referenciat per la lletra n. Així mateix, el nombre d'arestes, $|E|$, és referenciat per la lletra m i se l'anomena **mida** del graf.

Els **nodes adjacents** o **veïns**, denotats com $\Gamma(V_i)$, es defineixen com el conjunt de nodes units a v_i per mitjà d'una aresta. En aquest cas, el **grau** d'un node es defineix com el nombre de nodes adjacents, és a dir, $|\Gamma(V_i)|$, encara que generalment el grau del vèrtex v_i es denota com $deg(v_i)$. La **seqüència de graus** (*degree sequence*) és una seqüència numèrica de n posicions en què cada posició i indica el grau del node v_i.

En un graf dirigit G es defineix els successors d'un node v_i, $\Gamma(V_i)$ com el conjunt de nodes als quals es pot arribar dibuixant un arc des de v_i. Es defineix el grau exterior d'un node com el nombre de successors $|\Gamma(V_i)|$. De manera similar, es poden definir els antecessors d'un node v_i, $\Gamma^{-1}(v_i)$ com el conjunt de nodes des dels quals és possible arribar a v_i dibuixant un arc. Es defineix el grau interior d'un node com el nombre d'antecessors, és a dir, $|\Gamma^{-1}(v_i)|$.

En el cas de les xarxes socials, les dades se solen representar utilitzant els grafs, atès que permeten una representació natural de les relacions existents entre un conjunt d'usuaris (representats mitjançant nodes o vèrtexs en el context dels grafs) de la xarxa. Hi ha diversos formats de graf que permeten representar cadascuna de les xarxes existents a la realitat i que s'adapten a les particularitats de cadascuna d'elles. Per exemple, podem trobar xarxes amb relacions simètriques, com ara Facebook, on si l'usuari A és amic de l'usuari B, necessàriament B és amic d'A. D'altra banda, podem trobar exemples de xarxes asimètriques, com ara Twitter, que es representen mitjançant grafs dirigits o asimètrics, i on la relació de «seguir» d'un usuari A cap a un usuari B no ha de ser recíproca.

Representació mitjançant grafs

Suposem que volem modelar les relacions entre els usuaris de Twitter, mostrant la relació «seguir» que s'estableix entre dos usuaris de la xarxa. Per representar la informació que ens interessa podem fer servir un graf dirigit o asimètric, on creem un arc entre els nodes A i B si l'usuari A «segueix» l'usuari B. La figura 5 mostra un possible escenari on podem veure que els usuaris A, B i C «segueixen» l'usuari D. Per la seva banda, l'usuari D «segueix» els usuaris A, E i F.

A continuació veurem com modelar les relacions en una xarxa simètrica o no dirigida, com pot ser, per exemple, Facebook. En aquesta xarxa els usuaris estableixen «relacions d'amistat» bidireccionals, és a dir, si un usuari A és «amic» d'un usuari B, llavors implícitament l'usuari B també és «amic» de l'usuari A . La figura 6 mostra un possible exemple en el qual veiem les relacions d'amistat entre set usuaris. Podem veure que l'usuari A és amic de D, el qual és també amic de A i de C, E i F.

Figura 5. Grafo dirigit o asimètric

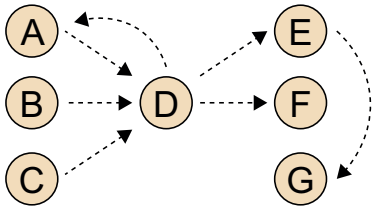
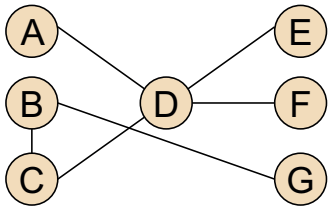


Figura 6. Graf no dirigit o simètric



En aquests casos, la pròpia estructura de la xarxa conté informació de gran utilitat per a l'anàlisi i l'estudi de les xarxes, però també pot ser utilitzat per un atacant per obtenir informació i identificar un usuari dins de les dades protegides. Els mètodes d'emmascarament específics per a xarxes socials es basen en modificar l'estructura dels grafs, ja sigui afegint o eliminant arestes o vèrtexs del graf, per introduir soroll en les dades i dificultar el procés d'identificació d'usuaris en les dades protegides.

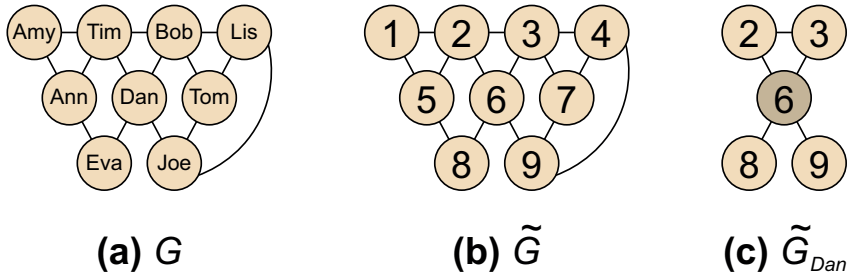
4.2. Definició del problema

Actualment, s'estan recopilant grans quantitats de dades sobre xarxes socials, que sovint contenen informació personal i privada d'usuaris i individus. Encara que es realitzen processos bàsics d'anonimització de dades, com l'eliminació de noms o altres identificadors de claus, la informació restant pot ser sensible i útil perquè un atacant torni a identificar usuaris i individus dins el conjunt de dades anònims.

Per resoldre aquest problema, s'han desenvolupat mètodes que introdueixen soroll en les dades originals per tal d'obstaculitzar els processos posteriors de reidentificació. Una estratègia natural per protegir la informació sensible és reemplaçar els atributs d'identificació amb identificadors sintètics. Ens referim a aquest procediment com *anonimització simple* (*naïve anonymization*). Aquesta pràctica mira de protegir la informació sensible trencant l'associació entre la identitat del món real i les dades sensibles.

La figura 7a mostra un exemple d'una xarxa social, on cada vèrtex representa un individu i cada aresta indica la relació d'amistat entre dos individus. La figura 7b presenta el mateix graf després d'un procés d'anonimització simple, on els identificadors de vèrtexs han estat eliminats i l'estructura es manté inalterada. En aquest escenari, un atacant pot trencar la privacitat i tornar a identificar un usuari en el graf anònim. Per exemple, si un atacant sap que l'usuari *Dan* té quatre amics i que dos d'ells són amics entre ells, llavors pot construir un subgraf a distància un del *Dan*, representat a la figura 7c. A partir d'aquest subgraf, l'atacant pot reidentificar l'usuari *Dan* en el graf anònim, atès que és l'únic node amb aquest patró de subgraf a distància un. I pot, per tant, trencar la privacitat dels usuaris.

Figura 7. Exemple d'anonimització simple, on G és el graf original, $\widetilde{G}$ és la versió anònima de la xarxa i $\widetilde{G}_{Dan}$ representa el subgraf a distància un de l'usuari *Dan*



Zhou i Pei van mostrar que, per definir el problema de la preservació de la privacitat en la publicació de dades de xarxes socials, necessitem formular les qüestions següents: en primer lloc, necessitem identificar la informació que s'ha de preservar. En segon lloc, cal modelar el coneixement que un adversari pot fer servir per atacar la privacitat de la xarxa. I, en tercer lloc, necessitem especificar l'ús de les dades publicades perquè un mètode d'anonimització pugui intentar retenir la utilitat, tant com sigui possible, mentre preserva la privacitat de la informació continguda a la xarxa.

Lectura complementària

B. Zhou; J. Pei (2008). «Preserving Privacy in Social Networks Against Neighborhood Attacks». A: *IEEE International Conference on Data Engineering* (pàg. 506-515). Conferència. Washington D. C.: IEEE Computer Society.

Respecte a la informació que s'ha de preservar en les xarxes socials, s'han identificat tres categories principals d'amenaces a la privacitat:

1) La **divulgació de la identitat** (*identity disclosure*) té lloc quan es revela la identitat d'un individu associat amb un vèrtex del graf anònim.

2) La **divulgació dels atributs** (*attribute disclosure*) no busca identificar necessàriament un vèrtex, sinó revelar atributs o dades sensibles del vèrtex. Les dades sensibles associades a cada vèrtex es veuen compromeses.

3) La **divulgació de les relacions** (*link disclosure*) té lloc quan es revela la relació sensible entre dos individus.

La divulgació d'identitat i la divulgació de les relacions s'apliquen a tot tipus de grafs. No obstant això, la divulgació dels atributs només s'aplica a les xarxes

que contenen informació associada als vèrtexs. La revelació de la identitat sovint condueix a la divulgació dels atributs, pel fet que la revelació d'identitat passa quan un individu es reidentifica dins d'un conjunt de dades, mentre que la divulgació dels atributs passa quan s'identifica informació sensible que un individu vol mantenir en privat.

Determinar el coneixement de l'adversari és el problema clau. S'ha proposat un ampli ventall de coneixements dels adversaris, en conjunció amb el seu corresponent atac i mètode de protecció. En la criptoanàlisi, els autors distingeixen entre dos tipus bàsics d'atacs, que pot ser també una classificació vàlida per als atacs a xarxes socials:

- **Atacs actius**, on l'adversari intenta comprometre la privacitat mitjançant la creació estratègica de nous comptes d'usuari i vincles amb altres usuaris abans que es publiqui la versió anònima a la xarxa, de manera que aquests nous vèrtexs i arestes estiguin presents en la versió anònima.

- **Atacs passius**, que són duts a terme per individus que miren de descobrir les identitats dels vèrtexs només després que la xarxa anònima hagi estat publicada.

4.3. Mètodes d'anonimització

Hi ha tres grans famílies de tècniques d'anonimització en grafs que es basen en la modificació de l'estructura per preservar la privacitat de les dades d'una xarxa. Aquests mètodes són:

- **Modificació d'arestes i vèrtexs**: aquestes tècniques transformen el graf mitjançant modificacions d'arestes o vèrtexs (afegint o eliminant) i després publiquen les dades pertorbades. Les dades es posen així a disposició per a qualsevol tipus d'anàlisi, sense restriccions.

- **Grafs incerts** (*uncertain graphs*): aquest enfocament està basat en l'addició o eliminació d'arestes de manera «parcial», assignant una probabilitat d'existir a cada aresta de la xarxa anònima. En lloc de crear o eliminar arestes, es considera el conjunt de totes les arestes possibles i s'assigna una probabilitat d'existir a cadascuna d'elles.

- **Mètodes de generalització** (*generalization*): aquests mètodes busquen vèrtexs similars i els agrupen en particions, de manera que els detalls sobre els individus queden ocults. El principal problema és que la utilitat del graf pot reduir-se considerablement després del procés d'anonimització, especialment quan l'objectiu és preservar les estructures locals de la xarxa.

4.3.1. Mètodes basats en la modificació d'arestes o vèrtexs

Els enfocaments de modificació d'arestes o vèrtexs anonimitzen un graf per mitjà de modificacions (és a dir, afegint o eliminant) d'arestes o vèrtexs en el graf. Aquestes modificacions poden fer-se de forma aleatòria, i en aquest cas ens referirem a ells com a *mètodes aleatoris* (*randomization*), *pertorbatius* (*random perturbation*) o *tècniques d'ofuscació aleatòria* (*obfuscation*). No obstant això, la modificació es pot fer amb la finalitat de complir amb algunes restriccions desitjades, com en el cas dels mètodes basats en la *k*-anonimitat.

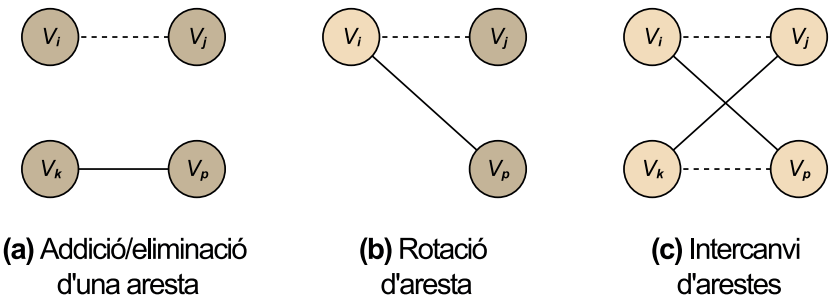
Definim tres processos bàsics de modificació d'arestes per canviar l'estructura de la xarxa, afegint o eliminant arestes. Aquests mètodes són els més bàsics, i es poden barrejar per crear combinacions complexes. Ens permeten modelar, de forma general i conceptual, la majoria dels mètodes de preservació de la privacitat. A continuació, introduirem aquests mètodes bàsics, que s'il·lustren a la figura 8. Les línies de traços representen arestes existents que s'eliminaran, mentre que les línies contínues constituiran les noves arestes. El color del node indica si canvia el seu grau (gris fosc) o no (gris clar) després del procés de modificació. Aquests processos són:

- **Addició/eliminació d'una aresta** (*edge add/del*). És la modificació d'aresta més bàsica. Simplement consisteix a eliminar una aresta existent $\{v_i,v_j\} \in E$ i afegir una altra de nova $\{v_k,v_p\} \notin E$. La figura 8a il·lustra aquest procés.

- **Rotació d'aresta** (*edge rotation*). Aquesta té lloc entre tres nodes $v_i,v_j,v_p \in V$ tal que $\{v_i,v_j\} \in E$ i $\{v_i,v_p\} \notin E$. Es defineix com la supressió de l'aresta $\{v_i,v_j\}$ i la creació de la nova aresta $\{v_i,v_p\}$, tal com s'il·lustra a la figura 8b.

- **Intercanvi d'arestes** (*edge switch*). En aquest cas, l'operació es produeix entre quatre nodes $v_i,v_j,v_k,v_p \in V$ on $\{v_i,v_j\},\{v_k,v_p\} \in E$ i $\{v_i,v_p\},\{v_k,v_j\} \notin E$. Es defineix com l'eliminació de les arestes $\{v_i,v_j\}$ i $\{v_k,v_p\}$, i la creació de les arestes $\{v_i,v_p\}$ i $\{v_k,v_j\}$, com es pot veure a la figura 8c.

Figura 8. Operacions bàsiques de modificació d'arestes



(a) Addició/eliminació d'una aresta

(b) Rotació d'aresta

(c) Intercanvi d'arestes

Per a les tres operacions presentades, el nombre de nodes i arestes es manté inalterable, però la distribució dels graus canvia en el cas de la rotació d'arestes. Clarament, l'addició/eliminació d'arestes és el concepte més general i totes les altres pertorbacions poden ser modelades com un cas particular de la mateixa.

Mètodes aleatoris

Aquests mètodes es basen en la introducció de soroll aleatori en les dades originals i han estat àmpliament investigats per les dades estructurades o relacionals.

Els enfocaments basats en la introducció de soroll aleatori, en general, protegeixen contra la reidentificació d'una manera probabilística. Específicament, els mètodes basats en l'addició/eliminació o rotació d'arestes preserven la identitat dels usuaris (*identity disclosure*), quan se suposa que el coneixement de l'adversari es basa en la informació del grau o dels vèrtexs veïns de l'objectiu, i també contra la divulgació de les relacions (*link disclosure*). Els mètodes basats en l'intercanvi d'arestes, però, no protegeixen contra la revelació de la identitat quan un adversari té coneixement sobre el grau dels vèrtexs, atès que la seqüència de graus del graf anònim segueix sent la mateixa.

Possiblement, les dues estratègies més bàsiques de pertorbació aleatòria en grafs són:

- Addició/eliminació aleatòria (*rand add/del*). Aquesta tècnica aplica, reiteradament, l'operació d'addició/eliminació d'arestes a l'atzar, considerant tot el conjunt d'arestes, sense restriccions. Aquesta estratègia conserva el nombre d'arestes en el graf anònim.

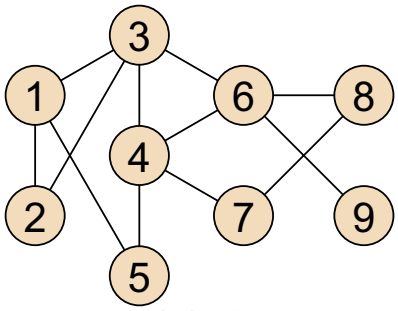
- Intercanvi aleatori (*rand switch*). Aquesta estratègia canvia aleatòriament i reiteradament parells d'arestes existents seguint la descripció de l'intercanvi d'arestes. En aquest cas, es preserva el nombre d'arestes del graf, però també el grau de tots els vèrtexs.

Exemple d'un procés d'anonimització aleatori

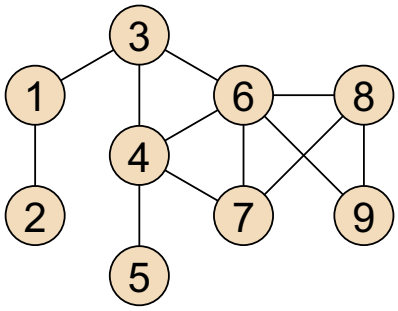
La figura 9 mostra un exemple d'anonimització aleatòria, on la xarxa original es presenta a la figura 9a. A continuació, la figura 9b mostra una versió pertorbada de la mateixa xarxa utilitzant el mètode d'addició/eliminació aleatòria (*rand add/del*). Durant el procés d'anonimització, s'han eliminat dues arestes ($\{1,5\}$ i $\{2,3\}$) i s'han creat dues noves arestes ($\{6,7\}$ I $\{8,9\}$). Una versió alternativa, en aquest cas emprant el mètode d'intercanvi aleatori (*rand switch*), es presenta a la figura 9c, on les arestes $\{1,2\}$ I $\{4,5\}$ han estat intercanviades, donant com a resultat les arestes $\{1,4\}$ i $\{2,5\}$.

Tots dos mètodes preserven el nombre de vèrtexs i arestes. A més, el mètode d'intercanvi aleatori també preserva la seqüència de graus, és a dir, $d(G) = d(\widetilde{G}_{sw}) = \{3,2,4,4,2,4,2,2,1\}$, mentre que el mètode d'addició/eliminació aleatori no té aquesta propietat, és a dir $d(\widetilde{G}_{ra}) = \{2,1,3,4,1,5,3,3,2\}$.

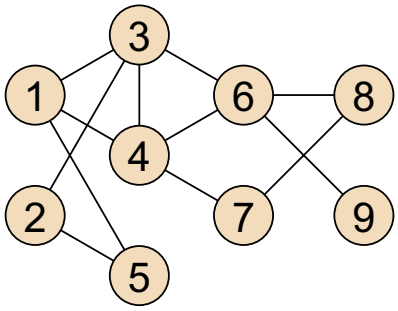
Figura 9. Exemple de pertorbació aleatòria, on G és el graf original, $\widetilde{G}_{ra}$ i $\widetilde{G}_{sw}$ són les versions anonimitzades emprant el mètode d'addició/eliminació aleatòria (*rand add/del*) i l'intercanvi aleatori (*rand switch*), respectivament



(a) G



(b) $\widetilde{G}_{ad}$



(c) $\widetilde{G}_{sw}$

Hay i altres proposen un mètode anomenat *random perturbation* per anonimitzar els grafs simètrics fent ús de l'estratègia d'addició/eliminació d'arestes, eliminant aleatòriament p arestes i després afegint aleatòriament p arestes falses no existents en el graf. Els principals avantatges d'aquest mètode són la seva simplicitat i la seva baixa complexitat. Per contra, els usuaris amb un nombre de veïns molt superior a la mitjana (*hubs*) no estan ben protegits i poden ser reidentificats dins del graf anònim.

Lectura complementària

M. Hay; G. Miklau; D. Jensen; P. Weis; S. Srivastava (2007). *Anonymizing Social Networks*. Amherst: University of Massachusetts.

Altres treballs posteriors intenten aplicar els mètodes anteriorment comentats, però en lloc d'aplicar-los de forma aleatòria entre totes les arestes del graf original, intenten preservar les arestes que contribueixen a preservar la utilitat del graf anònim. En aquest sentit, els autors Ying i Wu van desenvolupar dos mètodes dissenyats específicament per preservar les característiques espectrals del graf original, anomenats *spctr add/del* i *spctr switch*. En la mateixa línia, Casas-Roma proposa una estratègia que té com a objectiu preservar les arestes més importants de la xarxa, mirant de maximitzar la utilitat de les dades alhora que se n'assoleix el nivell de privacitat desitjat.

Lectura complementària

X. Ying; X. Wu (2008). *Randomizing Social Networks: a Spectrum Preserving Approach*. A: *SIAM Conference on Data Mining* (pàg. 739-750). Conferència. Atlanta: SIAM.

Lectura complementària

J. Casas-Roma (2014). *Privacy-Preserving on Graphs Using Randomization and Edge-Relevance*. A: V. Torra (ed.). *International Conference on Modeling Decisions for Artificial Intelligence* (pàg. 204-216). Conferència. Tokio: Springer International Publishing Switzerland.

k-anonimitat i derivats

Una altra estratègia àmpliament adoptada dels enfocaments de modificació d'arestes i vèrtexs té com a objectiu complir amb determinades restriccions de privacitat. És a dir, les modificacions no es fan de forma aleatòria, sinó que, al contrari, es busca realitzar les mínimes modificacions que permetin complir amb les restriccions de privacitat desitjades. Probablement, el model de *k*-anonimitat és el més conegut i utilitzat en aquest grup, encara que recentment s'han desenvolupat altres models i extensions.

El model de *k*-anonimitat, com s'ha descrit en el subapartat 2.3., indica que un atacant no podrà distingir un usuari entre un grup de k usuaris i, en conseqüència, no pot reidentificar a cap individu amb una probabilitat més gran que $\frac{1}{k}$.

Es poden emprar diferents conceptes com quasi identificadors per aplicar el model de *k*-anonimitat en xarxes o grafs. Una opció àmpliament utilitzada és fer servir el grau dels vèrtexs com un quasi identificador, entenent que és relativament fàcil saber el nombre de «amics» o «seguidors» que un usuari en concret té en una xarxa social. En conseqüència, suposem que l'atacant sap el grau d'alguns vèrtexs objectiu. Si l'atacant identifica un vèrtex únic amb el mateix grau en el gràfic anònim, llavors ha reidentificat aquest vèrtex. És a dir, $deg(v_i) \neq deg_(v_j) \forall j \neq i$.

Lectura complementària

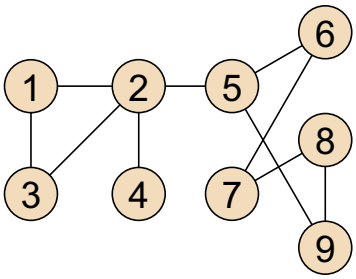
K. Liu; E. Terzi (2008). *Towards identity anonymization on graphs*. A: *ACM SIGMOD International Conference on Management of Data* (pàg. 93-106). Conferència. Nova York: ACM Press.

Aquest model es diu *k*-anonimitat basada en el grau (*k-degree anonymity*) i va ser introduïda per Liu i Terzi. Per tant, aquests mètodes es basen en la modificació de l'estructura del grau (mitjançant modificacions d'arestes o vèrtexs) per assegurar que tots els vèrtexs satisfan el model de *k*-anonimitat quan es

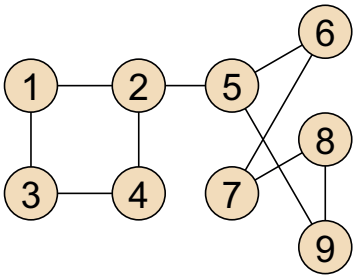
considera el grau com un quasi identificador. En altres paraules, l'objectiu principal és que tots els vèrtexs tinguin, almenys, $k-1$ altres vèrtexs compartint el mateix grau.

Una xarxa $G = (V,E)$ és k-anònima en el grau si, i només si, tots els seus vèrtexs compleixen la propietat de la k-anonimitat basada en el grau.

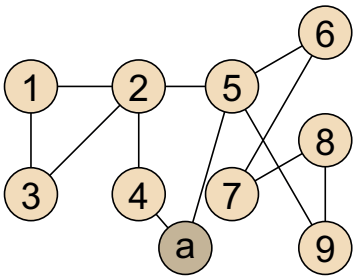
Figura 10. Exemple de k-anonimitat, on G representa el graf original, $\widetilde{G}_{em}$ i $\widetilde{G}_{va}$ són versions 2-anònimes basades en el grau a partir de modificacions en les arestes i els vèrtexs, respectivament



(a) G



(b) $\widetilde{G}_{em}$



(c) $\widetilde{G}_{va}$

Exemple de *k*-anonimitat basada en el grau

Un exemple de *k*-anonimització basada en el grau s'il·lustra a la figura 10. La xarxa original G, representada a la figura 10a, és a $k = 1$ anònima respecte al grau, atès que la seva seqüència de graus és $d(G) = \{2,4,2,1,3,2,2,2,2\}$.

Un exemple d'una xarxa anònima amb $k = 2$ es presenta a la figura 10b. En aquest cas, s'ha emprat la modificació d'arestes per complir amb el model de *k*-anonimitat basada en el grau. Per tant, el nombre de vèrtexs és el mateix, és a dir $\widetilde{n} = n$, i la pertorbació s'aconsegueix afegint i eliminant arestes. La seva seqüència de graus és $d(\widetilde{G}_{em}) = \{2,3,2,2,3,2,2,2,2\}$. Per tant, és una seqüència $k = 2$ anònima a causa del fet que cada valor de grau apareix, almenys, dues vegades en la seqüència de graus.

Alternativament, podem veure una altra xarxa $k = 2$ anònima basada en el grau a la figura 10c, però en aquest cas s'ha emprat l'addició de vèrtexs. Com es mostra, l'estructura original roman igual, però s'ha afegit un nou vèrtex (gris fosc) i s'han creat dues arestes $\{a,4\}$ I $\{a,5\}$ per complir amb la propietat de 2-anonimitat basada en el grau. La seva seqüència de graus és $d(\widetilde{G}_{va}) = \{2,4,2,2,4,2,2,2,2,2\}$. Mitjançant aquest model, el nombre de vèrtexs s'incrementa en un ($\widetilde{n} = n + 1$) i el nombre d'arestes en dos ($\widetilde{m} = m + 2$).

Els mateixos autors que van plantejar el concepte de *k*-anonimitat basada en el grau, Liu i Terzi, van proposar un mètode basat en la programació lineal sencera i en l'intercanvi d'arestes per construir un nou graf *k*-anònim basat en el grau, mantenint un nombre de vèrtexs i mirant de minimitzar el nombre d'arestes modificades durant el procés, és a dir $V = \widetilde{V}$ i $E \cap \widetilde{E} \approx E$. El treball de Liu i Terzi va inspirar molts altres autors que van millorar aquest treball seminal, tant en termes de velocitat com d'escalabilitat (permetent abordar conjunts de dades més grans), utilitzant diferents tipus d'heurístiques per a la selecció de les arestes que afegir o eliminar en el procés de construcció del graf anònim.

Lectura complementària

S. Chester; B. M. Kapron; G. Ramesh; G. Srivastava; A. Thomo; S. Venkatesh (2013). «Why Waldo befriended the dummy? *k*-Anonymization of social networks with pseudo-nodes». *Social Network Analysis and Mining* (vol. 3 (3), pàg. 381-399).

Els autors Chester i altres van plantejar el mateix problema, però permetent modificacions al conjunt de vèrtexs, en lloc de només al conjunt d'arestes, i això ofereix algunes diferències pel que fa a la utilitat del graf anònim. En aquest treball van crear noves arestes entre vèrtexs falsos i reals o únicament entre vèrtexs falsos. No obstant això, els resultats van mostrar que la pèrdua d'informació augmenta quan s'utilitza simultàniament l'addició de nous vèrtexs falsos i la creació d'arestes.

El model de la *k*-anonimitat basada en el grau ha estat àmpliament estudiat, però també s'han considerat altres models derivats que incrementen el nivell de privacitat, normalment incrementant també el nivell de complexitat de càlcul del graf anònim, i en conseqüència dificulten l'aplicació d'aquests mètodes en grafs de grans dimensions (centenars de milers o milions de vèrtexs i arestes).

En lloc de fer servir el grau de vèrtex com quasi identificador, altres treballs consideren el subgraf de veïnatge a distància un dels vèrtexs objectius com quasi identificador. Per a un vèrtex $v_i \in V$, v_i és *k*-anònim a G si hi ha almenys altres $k-1$ vèrtexs $v_1, \ldots, V_{k-1} \in V$ tal que $\Gamma(V_i), \Gamma(V_1), \ldots, \Gamma(V_{k-1})$ que són isomorfs. Llavors, G es considera *k*-anònim basat en el veïnatge a distància un si tots els vèrtexs de G són *k*-anònims considerant els veïns a distància un del vèrtex objectiu.

Altres autors han modelat un coneixement de l'adversari més complex i, en conseqüència, han creat models de *k*-anonimitat basats en quasi identificadors que impliquen un coneixement més gran per part de l'adversari. Per exemple, Hay i altres van proposar un mètode anomenat *k*-anonimitat basat en un conjunt de candidats. En aquest mètode, un vèrtex v_i és *k*-anònim respecte a una pregunta Q si n'hi ha, almenys, altres $k-1$ vèrtexs en el graf amb la mateixa resposta a la pregunta Q. Formalment, $|cand_Q(v_i)| \geq k$ on $cand_Q(v_i) = \{v_j \in V : Q(v_i) = Q(v_j)\}$. Un graf és *k*-anònim respecte a una pregunta Q, si tots els seus vèrtexs són *k*-anònims que fa a la pregunta Q. Aquest model és més general que els vistos anteriorment, i permet modelar el coneixement de l'adversari amb una consulta Q, que pot referir-se a qualsevol propietat estructural de la xarxa o graf.

Lectura complementària

M. Hay; G. Miklau; D. Jensen; D. Towsley; P. Weis (2008). «Resisting structural re-identification in anonymized social networks». *Proceedings of the VLDB Endowment* (vol. 1 (1), pàg. 102-114).

4.3.2. Grafs incerts

En lloc d'anonimitzar els grafs afegint o eliminant arestes i vèrtexs per satisfer determinats paràmetres de privacitat, mètodes recents han aprofitat la semàntica dels grafs incerts (*uncertain graphs*) per aconseguir la protecció de la privacitat desitjada. Considerem el graf $G = (V,E)$ i denotem V_2 com el conjunt de tots els parells de vèrtexs $\binom{n}{2}$ desordenats de V, és a dir $V_2 = \{(V_i,v_j) : 1 \leq i < j \leq n\}$. Un graf incert és un parell $\widetilde{G} = (V,p)$, on $p : V_2 \rightarrow [0,1]$ és una funció que assigna les probabilitats existents a tots les arestes possibles. Aquestes tècniques anonimitzen un gràfic determinístic, convertint-lo en una forma incerta on totes les possibles arestes existeixen amb una certa probabilitat en el rang [0,1].

Exemple d'anonimització mitjançant un graf incert

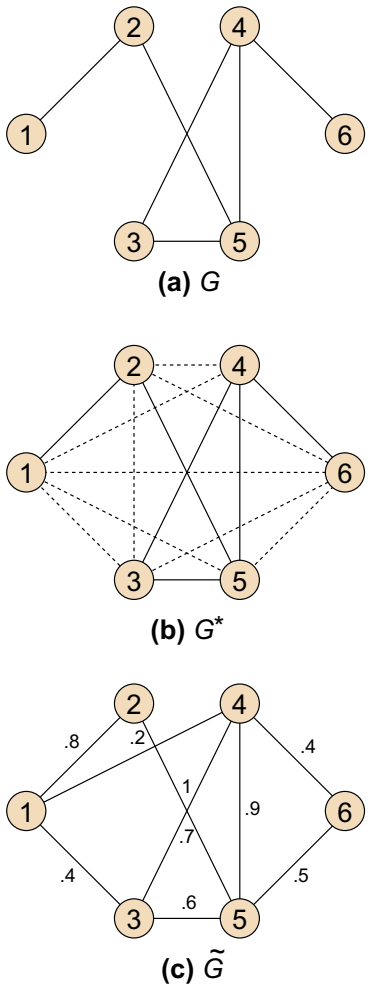
La figura 11 mostra el procés d'anonimització sota el model de graf incert. El graf original G està representat a la figura 11a, i la versió incerta del mateix graf es mostra a la figura 11b. Com es pot veure, hi ha totes les arestes possibles, és a dir, $\binom{6}{2}$, i cadascuna està assignada a una probabilitat igual a un (línies negres) o zero (línies de punts). Així, G^* és la representació de G sota el model de graf incert, però no és pertorbat ni anonimitzat. La versió anònima es presenta a la figura 11c, on la probabilitat de cada aresta es pertorba i s'estableix en el rang [0,1]. Les arestes amb probabilitat igual a zero no es representen en $\widetilde{G}$ per preservar una visualització clara del graf incert pertorbat.

El primer enfocament va ser proposat per Boldi i altres i es basa a injectar incertesa en les xarxes originals i publicar els grafs incerts resultants. Des d'una perspectiva probabilística, afegir una aresta no existent $\{v_i,v_j\}$ correspon a canviar la seva probabilitat $p(\{v_i,v_j\})$ de zero a un, mentre que eliminar una aresta correspon a canviar la seva probabilitat d'un a zero.

Lectura complementària

P. Boldi; F. Bonchi; A. Gionis; T. Tassa (2012). «Injecting Uncertainty in Graphs for Identity Obfuscation». *Proceedings of the VLDB Endowment* (vol. 5 (11), pàg. 1376-1387).

Figura 11. Exemple d'anonimització d'un graf incert, on G representa la xarxa original. El graf incert és G^*, on les arestes existents tenen associada una probabilitat igual a un (línies negres) i les inexistents una probabilitat igual a zero (línies de punts). $\widetilde{G}$ representa un possible graf incert després del procés d'anonimització



En aquest mètode, en lloc de considerar només les probabilitats binàries de les arestes, permeten que les probabilitats prenguin qualsevol valor en el rang [0,1]. Per tant, cada aresta està associada a una probabilitat específica en el gràfic incert. Per mantenir la utilitat del graf anònim, el mètode injecta incertesa (o soroll) només en un petit subconjunt de parells de vèrtexs E_c, i suposa que no hi ha altres parells de vèrtexs, és a dir $p(v_i,v_j) = 0 \forall (V_i,v_j) \notin E_c$.

Un graf incert és (k,ϵ)-ofuscat respecte a la propietat si l'entropia de la distribució $Y_{P(v)}$ sobre almenys $(1-\epsilon)n$ vèrtexs de $\widetilde{G}$ és més gran o igual que $log_2(k)$, és a dir, $H(Y_{P(v)}) \geq log_2(k)$.

És important remarcar que les estadístiques i les mètriques han de ser definides (o redefinides) per a ser aplicades en aquest tipus de grafs, ja que gairebé totes van ser dissenyades per treballar amb grafs binaris i no poden aplicar-se directament en aquest tipus de grafs incerts. En aquesta direcció, el càlcul de les estadístiques basades en el grau, com ara el nombre d'arestes, el grau mit-

jà, el grau màxim i la variància de grau van ser proposades en el treball dels mateixos autors.

4.3.3. Mètodes de generalització

Els mètodes de generalització, també coneguts com *enfocaments basats en clústers* (*generalization* o *clustering-based approaches*), es basen, essencialment, a agrupar vèrtexs i arestes en particions anomenades *supervèrtexs* i *superaristes*. Els detalls sobre els individus poden ocultar-se adequadament, però el graf queda reduït considerablement després del procés d'anonimització, la qual cosa pot no ser desitjable per analitzar les estructures locals. El graf generalitzat, que conté les estructures d'enllaç entre les particions, a més de la descripció agregada de cada partició, encara es pot utilitzar per estudiar les macropropietats de la xarxa original. Encara que contingui les propietats del graf original, no té la mateixa granularitat i escala. Els mètodes de generalització ofereixen un bon nivell d'anonimat, però al preu de disminuir considerablement la utilitat del graf anònim.

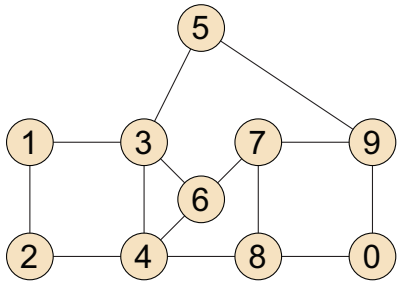
Els processos de generalització redueixen la mida del graf, tant en el nombre de vèrtexs com en el nombre de les arestes. En essència aquests mètodes produeixen un resum de la xarxa original, que també pot ser útil per reduir el temps de càlcul en els processos de mineria de grafs posteriors. No obstant això, tots els mètodes desenvolupats fins ara necessiten disposar del graf complet per poder analitzar-lo i decidir les particions de vèrtexs que realitzarà en el procés de generalització. En conseqüència, no són capaços de tractar amb grans volums de dades o amb dades en temps real (*streaming data*).

Com tots els mètodes abans esmentats, els enfocaments de generalització també protegeixen contra la revelació d'identitat. D'altra banda, és interessant subratllar que els enfocaments de generalització també preserven contra la divulgació d'atribut i de relació, ja que dos vèrtexs de qualsevol grup són indistingibles en funció de les seves relacions o els seus atributs.

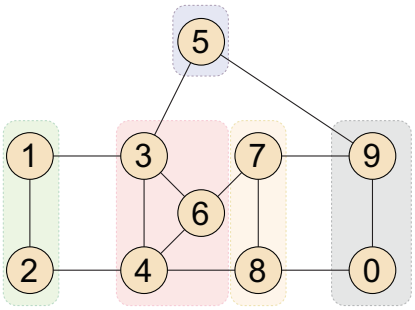
Exemple de generalització

La figura 12 presenta un exemple de procés d'anonimització basat en el mètode de generalització, on G és la xarxa original. En primer lloc, aquests mètodes calculen una partició de tot el conjunt de vèrtexs, com ara la partició que es mostra a la figura 12b, on tots els vèrtexs han estat assignats a un dels cinc clústers. Aquest és el procés més important, ja que agrupa vèrtexs amb característiques similars, condueix el procés de generalització a millors resultats en termes d'utilitat de dades i de pèrdua d'informació. En segon lloc, una vegada que es creen les particions, aquests mètodes agrupen tots els vèrtexs de la mateixa partició en un supervèrtex i creen superaristes entre ells. Una versió generalitzada de G es pot veure a la figura 12c. Com es mostra, cada supervèrtex conté informació sobre el nombre de vèrtexs i arestes internes entre ells. Generalment, cada superarista s'etiqueta d'acord amb el nombre total d'arestes entre tots els vèrtexs de cada supervèrtex.

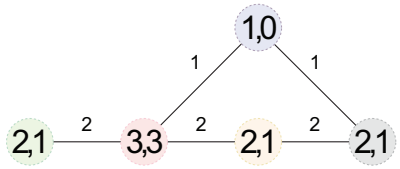
Figura 12. Exemple de generalització, on G representa la xarxa original. Es presenta una partició de mostra del conjunt de vèrtexs i s'utilitza per crear un gràfic generalitzat $\widetilde{G}$



(a) G



(b) Partició del conjunt de vèrtexs



(c) $\widetilde{G}$

Els autors Hay i altres van aplicar els mètodes de generalització utilitzant la mida d'una partició per assegurar l'anonimat dels vèrtexs. El seu mètode obté un supergraf *k*-anònim agrupant nodes en supervèrtexs i arestes en superaristes. Cada supervèrtex conté un conjunt més gran o igual a *k* vèrtexs i cada superarista representa totes les arestes entre vèrtexs en dues supervèrtexs. Només es publica la densitat de les arestes per a cada partició, de manera que serà difícil distingir entre individus dins d'una partició. Els autors van avaluar l'efectivitat de les consultes estructurals en xarxes reals de diversos dominis i grafs aleatoris. Els seus resultats van mostrar que les xarxes són diverses en la seva resistència als atacs: les xarxes socials i de comunicació tendeixen a ser més resistents del que alguns models aleatoris podrien suggerir.

Lectura complementària

M. Hay; G. Miklau; D. Jensen; D. Towsley; P. Weis (2008). «Resisting structural re-identification in anonymized social networks». *Proceedings of the VLDB Endowment* (vol. 1 (1), pàg. 102-114).

Un dels mètodes més coneguts de generalització va ser introduït per Campan i Truta i és aplicable a xarxes no dirigides amb vèrtexs etiquetats i arestes no etiquetades. Els atributs dels vèrtexs poden ser de tres tipus principals: identificadors, quasi identificadors i atributs sensibles. Apliquem el model de *k*-anonimitat als quasi identificadors per aconseguir vèrtexs indistingibles en els seus atributs o relacions entre ells. Els autors van desenvolupar un nou mètode, anomenat SaNGreeA, dissenyat per anonimitzar la informació estructural. Agrupa vèrtexs en diversos grups i, a continuació, s'assigna una etiqueta per a cada partició amb informació resumida (com el nombre de nodes en la partició).

Lectura complementària

A. Campan; T. M. Truta (2009). «Data and Structural *k*-Anonymity in Social Networks». A: F. Bonchi; E. Ferrari; W. Jiang; B. Malin (eds.). *Privacy, Security, and Trust in KDD* (pàg. 33-54). Berlín: Springer-Verlag.

4.4. Resum

En aquest apartat hem revisat la problemàtica de la publicació de dades en format semiestructurat, o format de xarxa o graf. En aquest tipus de dades, més enllà dels atributs que pot contenir cada individu (representat per un node o vèrtex) i que poden ser anonimitzades mitjançant les tècniques vistes en l'apartat 3, un atacant pot emprar informació estructural per reidentificar els usuaris dins de la xarxa anònima. És a dir, el nombre de relacions i el subgraf a distància $d \geq 1$ poden aportar informació molt rellevant que l'atacant pot emprar per a la reidentificació dels usuaris objectiu.

En els últims anys, aquest ha estat un camp molt actiu en la investigació sobre privacitat i utilitat de les dades. S'han desenvolupat gran quantitat de models basats en la modificació d'arestes i vèrtexs, especialment models basats en la *k*-anonimitat, on s'ha buscat reduir la pèrdua d'informació i mantenir un nivell de privacitat adequat. Tot i això, alguns d'aquests models presenten un cost computacional molt elevat i no són aplicables a xarxes de centenars de milers o milions de nodes i arestes.

Els grafs incerts han estat introduïts recentment en el context de la privacitat. Tot i que el model presenta interessants propietats, també presenten la problemàtica que la majoria dels models de mineria de dades (concretament mineria de grafs, *graph mining*) no permeten treballar amb aquest tipus de grafs.

Els models de generalització, d'altra banda, presenten interessants propietats de privacitat, en unir tots els individus similars en un únic supernode amb la informació agregada, però no permeten anàlisis relacionades amb les estructures locals de les xarxes, ja que aquestes es perden en el procés d'agrupament.

5. Conclusions

Dades estructurades

Les dades estructurades són aquelles que segueixen un patró igual per a tots els elements i que a més és conegut *a priori*. Per exemple, les dades d'un full de càlcul presenten els mateixos atributs per a cada fila.

Tradicionalment les dades s'han emmagatzemat en format de taules. Durant molt de temps aquest ha estat el format «estàndard» d'emmagatzematge i transferència de la informació. A causa d'això, la preservació de la privacitat s'ha centrat durant molts anys en les dades en format de taules, també anomenades *dades estructurades*, com ara les dades contingudes en bases de dades relacionals o en fulls de càlcul.

Tots els mètodes vistos fins ara han estat desenvolupats per treballar amb dades estructurades i semiestructurades (de tipus xarxa o graf), encara que molts d'ells han estat estesos per poder treballar amb altres tipus de dades.

En el context social actual, es generen dades en multitud de formats diferents, que van molt més enllà de les tradicionals taules. D'una banda, és innegable que les dades que generen els mateixos ciutadans en les diferents xarxes socials són una font d'informació molt important per a l'administració i institucions de tota índole. D'aquestes xarxes podem obtenir dades sobre les opinions o preferències dels ciutadans, les seves localitzacions en diferents moments del dia i un llarg etcètera que creix dia a dia amb l'aparició de noves xarxes socials. D'altra banda, és habitual que tot tipus d'administracions públiques de la ciutat, especialment en el context de les ciutats intel·ligents (*smart cities*), generin i publiquin informació en diferents formats: des de documents censals dels comerços de la ciutat fins a tot tipus de serveis que s'ofereixen als ciutadans. Per tant, en aquest context és necessari, però no suficient, tractar la privacitat de les dades en format estructurat i semiestructurat (de tipus xarxa o graf). La diversitat de formats de les dades anirà en augment en els propers anys, i s'haurà de tractar de forma específica cadascun dels nous formats, considerant les seves característiques i especificacions.

A continuació enumerarem, molt breument, altres tipus de dades que per les seves característiques o auge recent han rebut molta atenció des del punt de vista de la privacitat.

5.1. Localització i temps

En alguns casos les dades porten associada informació referent a la localització i el moment exacte en què s'han produït. En aquests casos, generalment, la informació relativa a la localització i el moment en què es produeix és molt important i pot resultar de gran interès per a l'estudi del comportament humà. Això no obstant, aquestes dades també poden comportar una bretxa en la

privacitat d'un usuari i poden ser utilitzades per un atacant per aconseguir la seva identificació. Per exemple, suposem que es publiquen dades relacionades amb la mobilitat i la localització dels ciutadans. La xarxa social Foursquare* és un exemple de servei de localització i mobilitat, on els usuaris marquen (*check-in*) llocs específics on es troben i els comparteixen amb la resta d'usuaris. En el conjunt de dades protegides, els diferents registres d'un mateix usuari indiquen el conjunt de llocs on un usuari ha estat en un determinat interval de temps. Si un usuari és identificat dins d'un conjunt protegit, l'atacant pot obtenir informació dels llocs en què aquest ha estat i el moment exacte en què ha estat en cada lloc.

*http://foursquare.com

Foursquare

Foursquare és una xarxa social creada l'any 2009 i basada en la localització dels usuaris.

Per preservar la privacitat en aquest context es pot entendre la localització com un atribut numèric (per exemple, per mitjà de la coordenades GPS del punt) i aplicar sobre aquesta dada els mètodes d'emmascarament per dades numèriques, o també entendre la localització com un punt dins d'un municipi, una comarca i una província i, per tant, convertir aquestes dades en atributs categòrics i aplicar el concepte de generalització per emmascarar la localització real de l'usuari. Una aproximació similar pot ser emprada per a tractar amb les dades referents al temps.

GPS

El sistema de posicionament global o GPS (*Global Positioning System*, en anglès) és un sistema que permet determinar la posició d'un objecte facilitant la seva latitud i longitud.

5.2. Registres de cerca i accés

Cada vegada que un usuari realitza una recerca mitjançant un motor de cerca d'internet, aquest emmagatzema informació sobre l'usuari, el moment i el dispositiu des d'on es produeix la recerca i els termes introduïts. Aquesta informació és molt útil per a l'estudi del comportament humà i fins i tot per a la predicció d'epidèmies, com el projecte *Google Flu Trends**.

*https://www.google.org / Flutrends / about /

Google Flu Trends

Mitjançant l'agregació de registres de recerca en el motor de Google, el projecte intenta fer prediccions precises sobre l'activitat de la grip durant els anys 2008-2014.

L'exemple d'AOL de què hem parlat anteriorment va demostrar que una insuficient anonimització de les dades pot conduir un atacant a identificar usuaris dins de les dades protegides mitjançant les seves recerques. La particularitat d'aquest tipus de dades és que hem de tractar amb la semàntica de les consultes. Per exemple, si es vol poder aplicar el concepte de generalització, cal «entendre» el significat d'una consulta per poder generalitzar i així evitar les particularitats que poden conduir a la identificació d'un usuari. Però la semàntica de les consultes pot resultar molt complexa i cal disposar d'ontologies o altres eines semàntiques per poder establir equivalències i jerarquies entre les diferents consultes.

5.3. Documents

La publicació de documents també porta implícits alguns problemes de privacitat. A diferència dels casos fins ara comentats, quan es publiquen documents estem treballant amb dades no estructurades, a diferència de les taules o registres, que són dades estructurades. Certament, els documents no presenten

patrons d'atributs coneguts *a priori*; al contrari, els documents no presenten cap tipus d'estructura i per a la seva anàlisi cal analitzar tot el text considerant la gramàtica i la semàntica per poder extreure coneixement del text. En aquest sentit, és habitual crear índexs que permetin associar documents o parts dels mateixos a determinades paraules o conjunts de paraules.

La preservació de la privacitat en el cas dels documents ha de ser estudiada i analitzada des d'aquest doble punt de vista: d'una banda el propi document i, de l'altra, els índexs creats a partir del conjunt total de documents que poden guiar l'atacant a la identificació de dades personals privades dins dels mateixos.

Aquests són només alguns exemples de les problemàtiques que s'hauran de resoldre en els propers anys relacionades amb la privacitat i l'anonimització de les dades. Les dades són una font innegable d'informació, que pot ser utilitzada per les administracions públiques per millorar el rendiment de les ciutats i la vida dels seus ciutadans, a més de per part de qualsevol tipus d'empresa, que pot optimitzar el seu procés de màrqueting i venda. Però en cap cas no s'ha de permetre que aquestes «millores» siguin al preu de sacrificar la privacitat dels usuaris, que es poden veure discriminats davant de certes situacions pels resultats d'anàlisi i els processos de mineria de dades.

Resum

En aquest mòdul didàctic hem presentat la problemàtica bàsica que implica la publicació de dades en relació amb la preservació de la privacitat. Hem vist els models bàsics d'anonimització, és a dir, aleatorització, *k*-anonimitat i privacitat diferencial, i també els mètodes bàsics d'emmascarament.

Hem introduït el model de *k*-anonimitat, que és probablement el model més utilitzat en l'actualitat, tant en el sector industrial com en el sector acadèmic, per anonimitzar dades prèviament a la seva publicació. També hem presentat el model de privacitat diferencial, que en els últims anys està atraient una gran part de recerca en el camp de l'anonimització i preservació de la privacitat en processos de publicació de dades.

Durant el contingut d'aquest mòdul didàctic hem aprofundit en la problemàtica i les solucions de les dades estructurades en format de taula i les dades semiestructurades en format de xarxa o graf. Però com ja hem comentat, el ventall de tipologies de dades i casuístiques de problemes no para d'augmentar, de la mateixa manera que augmenten les dades accessibles per a qualsevol entitat o institució pública o privada. En els propers anys, possiblement, assistirem a un important auge en la investigació de mètodes de preservació de la privacitat, que permetin emprar la gran quantitat de dades que es generen actualment en un entorn que respecti la privacitat dels usuaris implicats en les anàlisis.

Glossari

Atributs identificadors. Conjunt d'atributs que permeten identificar de forma explícita un individu, com per exemple mitjançant el nom, el DNI o el número de la seguretat social.

Atributs quasi identificadors. Conjunt d'atributs que potencialment podrien identificar un individu.

Atributs sensibles. Conjunt d'atributs que presenten informació específica i sensible d'un individu en concret, i com a tals han de poder ser associats a un únic individu.

Dades estructurades. Aquelles que segueixen un patró igual per a tots els elements i que, a més, és conegut *a priori*. Per exemple, les dades d'un full de càlcul presenten els mateixos atributs per a cada fila.

Dades semiestructurades. Forma de dades que no conté una estructura fixa predefinida *a priori*, però conté etiquetes o altres marcadors per separar els elements semàntics i fer complir jerarquies de registres i camps de les dades. Per exemple, els documents JSON o HTML.

Dades no estructurades. Aquelles que no segueixen cap tipus de patró conegut *a priori*. Per exemple, dos documents de text o imatges.

Graf. Parella de conjunts $G = (V,E)$, on $V = \{v_1,v_2...,v_n\}$ és el conjunt de nodes o vèrtexs i $E = \{e_1,e_2,\ldots,e_m\}$ és un conjunt d'arestes o arcs que uneixen dos nodes $e_i = \{v_i,v_j\}$.

GPS. El sistema de posicionament global o GPS (*Global Positioning System*, en anglès) és un sistema que permet determinar la posició d'un objecte facilitant la seva latitud i longitud.

Mètode perturbatiu. Mètode d'emmascarament que altera les dades introduint cert soroll o distorsió per dificultar el procés d'identificació d'un usuari per part d'un atacant.

Mètode no perturbatiu. Mètode d'emmascarament basat en la generalització o l'esborrat d'algunes parts de la informació que podrien conduir un atacant a identificar de forma única un usuari dins de les dades protegides.

PPDP. Procés de preservació de la privacitat en la publicació de dades (*privacy-preserving data publishing*), que estudia com publicar dades, de manera que un cop publicades mantinguin la seva utilitat a la vegada que preservin la privacitat dels usuaris que hi apareixen.

Bibliografia

Benjamin, C.; Fung, M.; Wang, K.; Wai-Chee Fu, A.; Yu, P. S. (2011). *Introduction to Privacy-Preserving Data Publishing: Concepts and Techniques*. Estats Units d'Amèrica: CRC Press.

Casas-Roma, J.; Herrera-Joancomartí J.; Torra, V. (2016). «A survey of graph-modication techniques for privacy-preserving on networks». A: *Artificial Intelligence Review* (vol. 47 (3), pàg. 341-366). DOI: 10.1007/s10462-016-9484-8.

Charu, C.; Yu, P. S. (2008). *Privacy-Preserving Data Mining: Models and Algorithms*. Nova York: Springer.

Navarro-Arribas, G.; Torra, V. (2015). *Advanced research in data privacy*. Nova York: Springer.

Torra, V. (2010). *Privacy in Data Mining. Data Mining and Knowledge Discovery Handbook* (pàgs. 687-716). Nova York: Springer.

Torra, V.; Navarro-Arribas, G. (2014). «Data privacy». A: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (vol. 4 (4), pàg. 269-280).