ESP Domande Orale

Pensa

Lui è molto tranquillo, se non sai rispondere immediatamente a una domanda ti dà un input per farti ragionare e comunque in qualche modo rispondere. Da dei songoli giudizi alle domande e il voto finale è la media del voto delle domande. Si è dimostrato molto umano nei confronti degli studenti che imbroccano le domande. Ti tiene a parlare per circa 30 minuti, ma nuovamente è molto tranquillo. Se vede che tentenni/sbagli ti aiuta come scritto sopra.

1. Privacy by default e by design
2. novità introdotte dal GDPR (pseudonon. e poi la differenza con l'anonimizzazione),
3. Ruoli definiti dal GDPR (titolare del trattamento dei dati, responsabile, data subject)
4. Responsibility and accountability
5. Come bisogna agire in caso di data breach
6. Trasferimento dati oltre UE - regole tra UE e USA
7. Rbac e Abac
8. Metodi di controllo degli accessi (Rbac e abac, con anche xacml, nel caso di due policy che vanno in conflitto, ABAC prevede dei sistemi per risolvere il conflitto?)
9. Perché un controllo degli accessi basato sugli attributi è migliore di quello basato sui ruoli
10. Esempi di attributi utili ad identificare il contesto per una richiesta di accesso
11. Come vengono realizzate le politiche ABAC
12. Auditing (obiettivo, principio del GDPR garantito dall'auditing)
13. Definizione di k-anonimity,
14. K-anonimity (Contesto e definizione, limiti k-anonimity, algoritmo a scelta)
15. Immaginiamo un dataset con informazioni mediche con il pagamento di ticket sanitari ed esenzioni ecc. Immaginiamo di voler decidere un valore di k, che ragionamento farebbe?
16. Quali sono i parametri che prenderebbe in considerazione per la scelta del k nella k-anonimity
17. problemi principali della k-anon (complessità e il fatto che non sempre sia possibile distinguere attributi QI da sensibili),
18. Linking attack perché lo abbiamo introdotto? Esempi di linking attack che abbiamo introdotto? Attraverso quale tipo di attributi avviene?
19. Definizione di quasi identifier ed esempi
20. K anonimity abbiamo visto 4 algoritmi fare un discorso sintetico di come evolvono gli algoritmi che abbiamo visto, quali aspetti vengono migliorati?
21. Come esplorano il reticolo gli algoritmi di samarati e incognito
22. Secondo lei cosa influisce di più nella complessità del calcolo degli algoritmi di k anonimizzazione?

23. Algoritmo di Mondrian e TopDown
24. Come risolve topdown i problemi di Mondrian?
25. Pur migliorando l'algoritmo di k anonimity con tutti gli strumenti che abbiamo visto, quali sono i limiti che rimangono in tutti questi tipi di approcci?
26. Limiti algoritmi su k-an (omogeneità)
27. L-diversity (attacchi e problemi su l-diversity)
28. L-diversity e suoi limiti
29. T-closness
30. Delta-presence
31. Come fare per risolvere i problemi che sorgono con la t-closeness (usando la delta-presence)
32. definizione di delta-presence e del motivo per cui è necessario andare a impostare sia delta min che delta max
33. Privacy preserving Data Mining (Principali tecniche, collegamento data obfuscation - differential privacy)
34. Algoritmi di privacy preserving data mining: alberi decisionali approcci per la privacy
35. Come funziona ID3-delta
36. Differential privacy
37. Differential privacy introduzione breve + quando è possibile usare il meccanismo esponenziale e quando invece è necessario farlo (Suppongo che quando hai query numeriche poi decidere se utilizzare Laplace o esponenziale(quindi possibile), quando hai altri tipi di query (es. Boolean o categoriche) sei obbligato ad usare il meccanismo esponenziale(quindi necessario). O magari anche quando il rumore di Laplace ti sporcherebbe così tanto la risposta da farti avere un utilità bassissima)
38. Differential privacy (obiettivo e definizione, meccanismo randomizzato e sensitività, meccanismo di Laplace e limiti, sensitività globale e problemi, sensitività locale)
39. definizione di differential privacy e l'esempio del meccanismo con la monetina
40. Meccanismo di laplace come viene applicato e su quali basi si fonda.
41. Il meccanismo di laplace non si può applicare quando la query da un risultato non numerico, le viene in mente una situazione in cui il meccanismo di laplace restituisce un risultato che non ha alcun senso nonostante il risultato sia numerico
42. meccanismo esponenziale
43. Limitazioni della differential privacy
44. Cosa non assicura la diff-privacy?
45. Differenza tra i due tipi di sensitività
46. Sensitività globale: quando è poco utile?
47. Local Sensitivity
48. Le viene in mente un caso in cui ha molto più senso una sensitività locale rispetto ad una globale? (l'esempio della mediana)

Orale del 11 dicembre 2020

Studente 1 - voto 28

1. K-anonimity, cos'è, da cosa ci protegge, definizione formale
2. Secondo lei qual è una probabilità accettabile di successo di un linking attack (confidenza minore di 1/k). "no, apetti intendevo, nel senso, qual è secondo lei un K accettabile, con che procedimento sceglierebbe un k accettabile?".
3. Mi parli dell'algoritmo Incognito
4. Differential Privacy - Quali sono le motivazioni dietro all'introduzione del meccanismo esponeziale.
5. Cosa farebbe lei nel progettare un algoritmo che rispetti il meccanismo esponenziale? (lode?)
6. quando siam obbligati a usare un meccanismo alternativo a Laplace? (lo studente non è stato chiaro su questo punto)
7. Metodi di controllo degli accessi, differenze tra RBAC e ABAC, e in quale caso utilizzerebbe i due sistemi?

Studente 2 - voto 27

1. Secondo lei, qual è la soglia k accettabile per la K-anonimity? il suo collega è stato un po' impreciso, come sceglierebeb k? Che ragionamento farebbe? (scelta basata sulla variabilità del dominio, numero di attributi dei record, attributi QI simili, dimensione del dataset)
2. Il suo collega ha parlato di incognito. Ora, c'è una differenza sostaziale tra Samarati ed Incognito, qual'è? (fa fatica a rispondere, Pensa lo aiuta con le seguenti domande)
   - che cos'è un QI?
   - risposta: a samarati bisogna passargli i QI dall'esterno (o comunque costruire/passargli il reticolo). Ingcognito fa tutto da solo data una tabella.
3. Differential Privacy: definizione (intuitiva e formale). Affinché il meccanismo (qualsiasi) rispetti la DP, il valore di epsilon quale deve essere? Nel senso, che valore deve assumere epsilon per garantire la DF? (valori molto piccoli) (lo studente ha sbagliato la domanda, ma pensa l'ha corretto facendolo ragionare, in modo molto tranquillo)
4. Meccanismo di Laplace (lo studente ha completamente imbroccato la domanda dando uan definizione diversa ed essendo convinto dell'accuratezza della sua risposta, pensa l'ha corretto)
5. Nel meccanismo di Laplace oltre alla sensitività c'è un altro parametro, mi dice qualcosa su di lui? (epsilon)

Studente 3 - voto 30

1. K-anonimity, come si superano i limiti dell'homogeny attack e della conoscenza di background? (disorso su l-diversity)

2. Dati i criteri elencati dal cooolega precedente, come li andrebbe ad usare nell'applicare la k-anonimity? Che ragionamento/analisi applicherebbe? (**analisi del rischio!!** Non l'abbiamo trattato a lezione, ma ti fa ragionare per arrivarci!)
3. DP - la DP gode alcune proprietà gode di alcune proprietà utili per progettare un sistema differenzialmente privato, quali sono? (teorema di combinazione delle tecniche di DP)
4. C'è un caso di meccanismo che abbiamo dimostrato la cui dimostrazione è molto semplice, mi sa dire qual'è (meccanismo della monetina)
5. Parlando di DP - il numero di monetine da lanciare a cosa lo associa? C'è un meccanismo più accurato? Le monetine che cosa sono nell'ambito della DP? (simplesso di probabilità)
6. Auditing nel DB, a cosa servono e quale aspetto del GDPR vanno a coprire?

Studente 4 - voto 29

1. Delta-presence
2. La delta presence ha un affinità con la DF, in termini di presenza-assenza individui nel datset. Mi dica in cosa sono dverse queste due fedinizioni di privacy, sia dal punto di vista applicativo che dal punt odi vista teorico. (DP applicata alla query, Delta-presence a una tabella)
3. DP - Global vs local sensitivity
4. Esempio della mediana, le viene in mente qualche esempio pratico in cui ha più senso utilizzare la sensitività locale rispetto alla globale.
5. Metodi per il calcolo della sensitività locale, SOLO NOMI. (PTR, PBLS, Smooth Sensitivity)
6. GDPR: mi parli della privacy by design. Mi dice anche i 7 principi cardine su cui si basa la privacy by design.

Studente 5 - voto 30 e lode

1. Privacy Preserving Data Mining, cos'è e quali approcci abbiamo visto (definizione del problema, 3 approcci, metodo di bootstrap)
2. Mi parla di ID3-delta?
3. Pseudonimizzazione nel GDPR
4. Secondo lei come si potrebbe migliorare il GDPR? (domanda molto larga, senza una risposta precisa).