PROGETTAZIONE E IMPLEMENTAZIONE DI UN'APPLICAZIONE DI MOBILE CROWDSENSING RISPETTOSA DELLA PRIVACY

Tesi di Laurea Magistrale in Ingegneria Informatica

Gaspare Mulè

Relatore: Prof. Vincenzo Agate

Correlatore: Prof. Pierluca Ferraro

UNIVERSITÀ DEGLI STUDI DI PALERMO

FACOLTÀ DI INGEGNERIA

LAUREA MAGISTRALE IN INGEGNERIA INFORMATICA

PROGETTAZIONE E IMPLEMENTAZIONE DI UN'APPLICAZIONE DI MOBILE CROWDSENSING RISPETTOSA DELLA PRIVACY

Tesi di Laurea di
Dott. Gaspare Mulè

Relatore:
Prof. Vincenzo Agate

Controrelatore:

Correlatore:
Prof. Pierluca Ferraro

Anno Accademico 2023/2024

UNIVERSITÀ DEGLI STUDI DI PALERMO
FACOLTÀ DI INGEGNERIA

LAUREA MAGISTRALE IN INGEGNERIA INFORMATICA

PROGETTAZIONE E IMPLEMENTAZIONE DI UN'APPLICAZIONE DI MOBILE CROWDSENSING RISPETTOSA DELLA PRIVACY

Tesi di Laurea di
Dott. Gaspare Mulè

Relatore:
Prof. Vincenzo Agate

Controrelatore:

Correlatore:
Prof. Pierluca Ferraro

Sommario

Il paradigma noto come Crowdsensing sta guadagnando sempre più popolarità, diventando una tecnica sempre più diffusa per la raccolta di dati provenienti da ampie indagini di diversa natura. Ad essere coinvolti nella raccolta di dati sono gli utenti, attraverso i propri dispositivi. Tuttavia, questo approccio si scontra con questioni legate alla privacy, poiché informazioni riservate riguardanti gli utenti potrebbero essere rivelate. Risulta necessario pertanto realiz- zare sistemi di crowdsensing che trattino i dati in maniera sicura e di cui gli utenti possano fidarsi. Un'ulteriore questione da tenere in considerazione però è l'affidabilità degli utenti, i quali potrebbero fornire informazioni volutamente errate per danneggiare il sistema. Un sistema di crowdsensing dovrebbe essere in grado fronteggiare il problema eliminando i dati privi di significato. Lo scopo del presente lavoro è quello di creare un'applicazione di Mobile Crowdsensing rispettosa della privacy che possa acquisire i dati tramite un sistema di questionari, preservandone la privacy e l'integrità in ogni fase, ne elabori i contenuti in maniera del tutto cifrata secondo tecniche Privacy Preserving e ne restituisca, tramite metodi di Truth Discovery, le analisi statistiche non condizionate da eventuali attività che tendono a falsificarne i risultati. Per fronteggiare la problematica si utilizzeranno dei metodi di Outlier Detection che, al variare di specifici iperparametri, riusciranno ad escludere le informazioni false, facendo in modo di tenere esclusivamente dati significativi al fine dell'analisi. Il processo di variazione di tali iperparametri sarà automatizzato da un metodo sperimentale oggetto del presente lavoro.

Indice

Introduzione

Negli ultimi anni, c'è stata una diffusione crescente di un nuovo metodo di raccolta dati basato sulla partecipazione volontaria delle persone attraverso i loro dispositivi, noto come "crowdsensing". Questo approccio ha dimostrato di essere efficace nel raccogliere una grande quantità di dati, il che è fondamentale per affrontare problemi sociali su larga scala. Tutta- via, la gestione della privacy è un aspetto cruciale di questo processo, poiché senza adeguati meccanismi e protezioni, molte persone potrebbero essere riluttanti a partecipare.

Uno strumento importante nel contesto del crowdsensing è il "Truth Discovery", che mira a stabilire una versione accurata e oggettiva della realtà, specialmente in situazioni sociali complesse. La "verità" nonchè la definizione oggettiva di un problema o di una situazione è una descrizione accurata e non soggettiva della realtà. In un problema sociale, la ground truth è spesso difficile da definire, poiché può essere influenzata da fattori soggettivi, come le opinioni, i pregiudizi e le convinzioni delle persone coinvolte. Tuttavia, è importante cercare di definire una ground truth il più precisa possibile, al fine di sviluppare soluzioni efficaci al problema. Nel presente lavoro è intesa come percezione diffusa e di maggioranza.

Questo studio si concentra principalmente sul contesto del crowdsensing per raccogliere dati relativi a questioni sociali su larga scala. Introduce un'applicazione web e mobile sviluppata per Android e iOS, che consente agli utenti di rispondere a questionari. Tuttavia, la privacy è una preoccupazione centrale, e per affrontarla, il lavoro si basa su un algoritmo di crittografia omomorfica a chiave asimmetrica. Questa tecnologia Privacy Preserving permette operazioni matematiche sui dati cifrati, garantendo che i dati rimangano affidabili durante l'elaborazione e impedendo al server di leggere i dati in modo diretto.

L'applicazione è suddivisa in due aree principali: una per gli utenti, in cui possono selezionare e rispondere ai questionari, e una per gli amministratori, dove possono gestire i questionari e ottenere analisi statistiche. Queste analisi statistiche sono basate su un algoritmo di clustering che è stato scelto per la sua efficienza nel suddividere i dati in cluster e per la capacità di dare in

output dei valori che indicano il valore di verità presunto di quel determinato quesito. L'output di questo algoritmo dipende da due iperparametri, uno dei quali è cruciale per il processo di eliminazione degli outlier e per migliorare la precisione della "Truth Discovery". Inizialmente, questi iperparametri vengono inseriti manualmente dall'amministratore tramite l'area riservata dell'applicativo, ma in una versione successiva dell'applicazione, vengono variati automaticamente tramite un metodo sperimentale presentato in questo studio.

Il presente lavoro presenta dei caratteri innovativi in quanto combina un approccio informatico alle tecniche di elaborazione statistica e di elaborazione della privacy già conosciute in altri ambiti. Il lavoro di sintesi tra queste tecniche è stato condotto in coerenza con gli aspetti teorici e gli ambiti di applicazione delle tecniche utilizzate di cui si propone una doverosa trattazio- ne specifica attingendo alla letteratura sinora disponibile. In particolare viene presentata una panoramica sul Crowdsensing nel primo capitolo, passando per le tecniche che permettono di raccogliere dati in modo rapido, semplice ed economico, arrivando alle varie tecniche di Privacy Preserving nel secondo capitolo, essenziali, in quanto è necessario garantire agli utenti il rispetto e la garanzia della privacy. Nel terzo capitolo invece verrà presentato un metodo sperimentale per l'ottimizzazione di un iperparametro dell'applicativo e struttura del software partendo dalla parte progettuale e arrivando alla parte realizzativa, motivando le scelte implementatitve, il flusso dei dati e la struttura dei set di dati sintetici usati per i risultati sperimentali del quarto capitolo. Verranno utilizzate varie tecnologie e strumenti informatici che, combinati con le scienze statistiche, rappresentano un mix in grado di fornire un alto valore aggiunto all'analisi dei dati e alle successive scelte che possono riguardare il mondo economico e socio/culturale ad ampio spettro. Lo studio verrà condotto tutelando la riservatezza dei dati e la protezione del loro contenuto effettuando le elaborazioni direttamente su dati cifrati al momento dell'acquisizione.

Capitolo 1

Crowdsensing

1.1 Paradigma del MCS

Il presente lavoro, il cui fine è quello di ottenere risultati affidabili e tutelati ai fini della privacy, necessita di un processo di acquisizione di dati la cui componente del costo è da tenere in particolare considerazione pertanto si è deciso di sfruttare la tecnica del Mobile Crowdsensing (MCS) per le riconosciute peculiarità in tal senso; infatti, tra gli studi sinora condotti, si evidenziano vari approcci al mobile crowdsensing come si rileva in letteratura. In questa sezione, verrà introdotto il concetto di mobile crowdsensing e saranno discussi i lavori correlati sulla QoI in altri campi diversi dal mobile crowdsensing.

Il crowdsensing mobile è un paradigma proposto per la prima volta in [3] e sfrutta gli smartphone diffusi per acquisire informazioni dettagliate e aggiornate su luoghi o eventi di interesse. Questo sistema è generalmente composto da tre componenti architettoniche principali: i partecipanti, l'applicazione di rilevamento (app) e la piattaforma di crowdsensing mobile (abbreviata in MCP).

I partecipanti sono la parte centrale dell'architettura e utilizzano i propri dispositivi intelligenti per raccogliere dati da una varietà di sensori, tra cui posizione, immagini, suoni, dati accelerometrici, dati biometrici e pressione barometrica. Possono anche fornire informazioni più complesse relative all'area di rilevamento o a fenomeni ambientali, come il traffico o le condizioni meteorologiche. I partecipanti solitamente si registrano nel sistema fornendo un nome utente e una password per identificare in modo univoco i loro contributi.

L'applicazione di rilevamento (app) è installata sugli smartphone degli utenti ed è responsa-

bile dell'interazione con gli utenti per acquisire e visualizzare dati. Gli utenti possono attivare l'acquisizione dati manualmente o su richiesta dell'applicazione, periodicamente o una tantum.

La piattaforma di crowdsensing mobile (MCP) è il cuore del sistema, gestisce l'elaborazione e la distribuzione dei dati raccolti e coordina tutte le operazioni del sistema. Solitamente, è costituita da server dedicati all'elaborazione dei dati rilevati. Inoltre, l'MCP garantisce l'archiviazione ed elaborazione efficiente dei dati, spesso utilizzando database relazionali o database appositamente progettati per dati dai sensori. L'MCP potrebbe anche includere un sistema di reputazione per valutare l'affidabilità dei dati inviati dagli utenti e un meccanismo di incentivazione per motivare la partecipazione degli utenti attraverso ricompense per i loro contributi.

1.2 Gestione delle ricompense

OMISSIS

1.3 Pericoli per i dati degli utenti

OMISSIS

1.4 Classificazione, Regressione e Clustering

OMISSIS

1.5 Tecnica del k-means

OMISSIS

1.6 Outlier Removal

OMISSIS

Capitolo 2

Privacy Preserving

1.1 Gestione della privacy

L'MCS, che vanta numerose applicazioni nei più disparati ambiti, tuttavia solleva importanti questioni di privacy, come ad esempio la divulgazione della posizione degli utenti. Per garantire la privacy, sono disponibili diversi meccanismi di protezione della privacy, tra cui l'occultamento spaziale, l'utilizzo della privacy differenziale e la crittografia (che consiste nel codificare le informazioni in modo che non possano essere lette da persone non autorizzate). Questi approcci sono stati studiati in relazione all'MCS e le loro caratteristiche sono state confrontate per valutarne l'efficacia, fermo restando che le tecniche MCS rappresentano comunque soluzioni convenienti e scalabili per la raccolta e l'elaborazione dei dati.

1.2 Crittografia

Data l'importanza della tutela della privacy nel progetto si è scelto di implementarla tramite un algoritmo di crittografia. L'obiettivo della crittografia, infatti, è garantire la riservatezza dei dati nei processi di comunicazione e archiviazione in quanto il suo uso in dispositivi vincolati ha portato a considerare funzionalità aggiuntive, come la possibilità di delegare i calcoli a computer non attendibili che, quindi, non dovrebbero entrare in possesso dei dati in chiaro. Il computer eseguirà il calcolo su questi dati crittografati, quindi senza sapere nulla sul loro valore reale e restituirà il risultato che potrà essere successivamente decifrato che, per coerenza, deve essere uguale al valore calcolato previsto se eseguito sui dati originali. Per questo motivo, lo schema

di crittografia deve presentare una struttura particolare come quella omomorfa che, sebbene inizialmente sia stata esposta a falle, ha trovato un'ottima sintesi nell'algoritmo asimmetrico di Paillier. Come noto un algoritmo di crittografia omomorfa ha la peculiarità di mantenere i dati cifrati anche durante la loro elaborazione restituendo comunque il risultato corretto mentre le differenze peculiari tra algoritmi simmetrici e asimmetrici possono essere riassunti come di seguito esposto. Un algoritmo simmetrico prevede che la cifratura e la decifratura vengono eseguite con la stessa chiave. Quindi, il mittente e il destinatario devono concordare la chiave che utilizzeranno prima di eseguire qualsiasi comunicazione sicura. Pertanto, non è possibile per due persone che non si sono mai incontrate usare tali schemi. Questo implica anche condividere una chiave diversa con tutti quelli con cui vogliamo comunicare. La scelta ricade su un algoritmo asimmetrico in quanto tali schemi sono più funzionali di quelli simmetrici poiché non è necessario che il mittente e il destinatario si accordino su nulla prima della transazione. Inoltre, spesso forniscono più funzionalità. Questi schemi, a differenza di quelli simmetrici (come AES), hanno il grande svantaggio di basarsi su calcoli matematici non banali e molto più lenti. Si stima che le differenze di velocità arrivano ad essere fino a duemila volte più veloci. Tuttavia si può pensare a delle soluzioni in cui vengono utilizzate entrambe le tipologie fruendo dei vantaggi sulla velocità offerti dalle crittografie simmetriche e, contemporaneamente, dalla maggiore sicurezza offerta da quella asimmetrica.

Alla luce delle riflessioni appena esposte e per il livello di complessità dell'applicativo sviluppato la scelta operata non fornisce particolari condizionamenti in termini di tempo di elaborazione ma, allo stesso tempo, garantisce al meglio l'aspetto fondamentale costituito dalla tutela della privacy.

La scelta dell'algoritmo ha dovuto contemperare anche degli aspetti di carattere non banale che, oltre a quanto precedentemente esposto, hanno tenuto conto anche di molti studi come quelli di Shannon, il quale ha definito il concetto di perfetta segretezza, in cui il testo cifrato non fornisce alcuna informazione sul testo in chiaro o sulla chiave. Questo è stato dimostrato nel caso del "one-time pad". Gli schemi simmetrici e asimmetrici si basano su problemi matematici difficili da risolvere, ma la stima della loro sicurezza può essere ottimistica. La valutazione della sicurezza è complessa e coinvolge contesti di attacco come attacchi su testi cifrati noti o scelti (debolezza di cui molteplici schemi soffrono). La scelta dello schema dipende dai requisiti, come velocità, memoria e sicurezza. La diversità degli schemi è cruciale per affrontare nuovi requisiti e proteggere contro attacchi che potrebbero rompere schemi simili. Schemi deterministici possono portare a svantaggi, quindi è preferibile l'uso di schemi probabilistici.

Nei cifrari simmetrici, un vettore casuale chiamato IV è introdotto nel processo di crittografia per rendere il processo probabilistico. Nei cifrari asimmetrici, la randomizzazione è complessa poiché deve essere analizzabile come gli schemi deterministici. In particolare nel presente studio si è affrontato il problema dell'elaborazione dei dati già crittografati e da non decifrare nel server di elaborazione; ciò ha comportato la scelta di schemi non deterministici che garantisce una maggiore sicurezza anche se pecca in termini di prestazioni.

Uno schema è additivo omomorfo se si considerano gli operatori di addizione, e moltiplicativamente omomorfo se si considerano gli operatori di moltiplicazione. Sono stati pubblicati molti di questi schemi omomorfi che sono stati ampiamente utilizzati in molte applicazioni. Si noti che in alcuni contesti può essere di grande interesse avere questa proprietà non solo per un operatore ma per due allo stesso tempo. Queste definizioni significano che, per una chiave fissa k, è equivalente eseguire operazioni sui testi in chiaro prima della crittografia o sui corrispondenti testi cifrati dopo la crittografia.

OMISSIS

1.3 Algoritmo omomorfico di Paillier

OMISSIS

1.4 Tecniche di blinding

OMISSIS

1.5 RDH

OMISSIS

1.6 Pericoli nell'ambito mobile

OMISSIS

1.7 Modello basato sul Cloud

OMISSIS

Capitolo 3

Sistema proposto dell'app

3.1 Descrizione del Sistema

OMISSIS

3.2 Requisiti funzionali e non funzionali

OMISSIS

3.2.1 Requisiti Funzionali

OMISSIS

3.2.2 Requisiti non funzionali

OMISSIS

3.3 Progettazione

OMISSIS

3.3.1 Scenari

OMISSIS

3.3.2 Casi d'uso

OMISSIS

3.4 Architettura sistema

OMISSIS

3.5 Schema del sistema

OMISSIS

3.6 Struttura del software

OMISSIS

3.6.1 Librerie Python per la web app

OMISSIS

3.6.2 Flusso dati

OMISSIS

3.7 Funzionamento software

OMISSIS

Capitolo 4

Risultati sperimentali

Il valore della threshold ottimale potrebbe variare in base alla distribuzione dei dati. Non avendo a disposizione un set di dati per verificare l'esattezza del processo, si è deciso di costruirlo. Sono stati effettuati dei test con delle distribuzioni di dati differenti per studiare il comportamento del software. I dati sono stati scelti secondo i valori della varianza standard sia elevata che bassa. Dopo un'attenta analisi sono stati suddivisi i set di dati in input in 3 macro categorie/distribuzioni.

4.1 Struttura set di dati

OMISSIS

4.2 Test

OMISSIS

Conclusioni

Sempre più frequentemente è di rilevante importanza nello studio dei comportamenti sociali l'acquisizione e la successiva elaborazione di dati ottenuti rilevando preferenze e opinioni dalla più ampia platea possibile di utenti.

Il presente lavoro è stato progettato al fine di avere un'applicazione che riuscisse ad individuare, in una sessione di crowdsensing, il valore di verità al netto di comportamenti malevoli e nella totale tutela della privacy. In particolare l'applicazione è stata pensata in ambiente mobile pertanto i software e le librerie utilizzate sono specifiche per l'ambito d'interesse.
OMISSIS

Elenco delle figure

Bibliografia

[1] A. Aladžuz, A. Delalic´ e L. Šc´eta. «Cluster Analysis in Python: An Example of Market Segmentation». In: International Conference "New Technologies, Development and Applications". Springer. 2022, pp. 1032–1041.

[2] V. Barnett, T. Lewis et al. Outliers in statistical data. Vol. 3. 1. Wiley New York, 1994.

[3] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy e M. B. Srivastava. In: «Participatory sensing» (2006).

[4] V. Agate, A. De Paola, G. Lo Re e M. Morana. «A platform for the evaluation of distributed reputation algorithms». In: 2018 IEEE/ACM 22nd International Symposium on Distributed Simulation and Real Time Applications (DS-RT). IEEE. 2018, pp. 1–8.

[5] T. Cho e S.-H. Seo. «A Strengthened Android Signature Management Method.» In: KSII Transactions on Internet & Information Systems 9.3 (2015).

[6] G. Douzas, F. Bacao e F. Last. «Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE». In: Information Sciences 465 (2018), pp. 1–20.

[7] S. Gaonkar, J. Li, R. R. Choudhury, L. Cox e A. Schmidt. «Micro-blog: sharing and querying content through mobile phones and social participation». In: Proceedings of the 6th international conference on Mobile systems, applications, and services. 2008, pp. 174–186.

[8] S. Garg e N. Baliyan. «Comparative analysis of Android and iOS from security viewpoint». In: Computer Science Review 40 (2021), p. 100372.

[9] M. Goyal e S. Kumar. «Improving the initial centroids of k-means clustering algorithm to generalize its applicability». In: Journal of The Institution of Engineers (India): Series B 95 (2014), pp. 345–350.

[10] V. Agate, A. De Paola, P. Ferraro, G. Lo Re e M. Morana. «SecureBallot: A secure open source e-Voting system». In: Journal of Network and Computer Applications 191 (2021).

[11] F. E. Grubbs. «Procedures for detecting outlying observations in samples». In: Technometrics 11.1 (1969), pp. 1–21.

[12] A. Hatamlou. «In search of optimal centroids on data clustering using a binary search algorithm». In: Pattern Recognition Letters 33.13 (2012), pp. 1756–1760.

[13] V. Hautamäki, S. Cherednichenko, I. Kärkkäinen, T. Kinnunen e P. Fränti. «Improving k-means by outlier removal». In: Image Analysis: 14th Scandinavian Conference, SCIA 2005, Joensuu, Finland, June 19-22, 2005. Proceedings 14. Springer. 2005, pp. 978–987.

[14] V. Agate, F. M. D'Anna, A. De Paola, P. Ferraro, G. Lo Re e M. Morana. «A behaviorbased intrusion detection system using ensemble learning techniques». In: ITASEC (2022).

[15] D. M. Hawkins. Identification of outliers. Vol. 11. Springer, 1980.

[16] V. Hodge e J. Austin. «A survey of outlier detection methodologies». In: Artificial intelligence review 22 (2004), pp. 85–126.

[17] C. Jost, H. Lam, A. Maximov e B. Smeets. «Encryption performance improvements of the paillier cryptosystem». In: Cryptology ePrint Archive (2015).

[18] V. Agate, S. Drago, P. Ferraro e G. Lo Re. «Anomaly Detection for Reoccurring Concept Drift in Smart Environments». In: 2022 18th International Conference on Mobility, Sensing and Networking (MSN). IEEE. 2022, pp. 113–120.

[19] D. Kabakchieva. «Predicting student performance by using data mining methods for classification». In: Cybernetics and information technologies 13.1 (2013), pp. 61–72.

[20] S. Kaur e U. Kaur. «A survey on various clustering techniques with K-means clustering algorithm in detail». In: Int. J. Comput. Sci. Mob. Comput 2.4 (2013), pp. 155–159.

[21] J. W. Kim, K. Edemacu e B. Jang. «Privacy-preserving mechanisms for location privacy in mobile crowdsensing: A survey». In: Journal of Network and Computer Applications 200 (2022), p. 103315.

[22] R. Kumar e R. Verma. «Classification algorithms for data mining: A survey». In: International Journal of Innovations in Engineering and Technology (IJIET) 1.2 (2012), pp. 7–14.

[23] V. Agate, A. De Paola, G. Lo Re e M. Morana. «Vulnerability Evaluation of Distributed Reputation Management Systems». In: InfQ 2016 - New Frontiers in Quantitative Methods in Informatics. ICST, Brussels, Belgium: ICST, 2016, pp. 1–8.

[24] S. Lee et al. «Assessment of malicious applications using permissions and enhanced user interfaces on Android». In: 2013 IEEE International Conference on Intelligence and Security Informatics. IEEE. 2013, pp. 270–270.

[25] V. Agate, A. De Paola, G. Lo Re e M. Morana. «DRESS: A Distributed RMS Evaluation Simulation Software». In: International Journal of Intelligent Information Technologies (IJIIT) 16.3 (2020), pp. 1–18.

[26] H. Liu, J. Li, Y. Wu e Y. Fu. «Clustering with outlier removal». In: IEEE transactions on knowledge and data engineering 33.6 (2019), pp. 2369–2379.

[27] B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer e W. F. Punch. «Predicting student performance: an application of data mining methods with an educational web-based system». In: 33rd Annual Frontiers in Education, 2003. FIE 2003. Vol. 1. IEEE. 2003, T2A–13.

[28] A. Montinaro e P. Rizzo. «Segretezza Perfetta». In: Quaderni di Matematica 2019.1 (2019), pp. 34–42.

[29] V. Agate, A. De Paola, G. Lo Re e M. Morana. «A simulation framework for evaluating distributed reputation management systems». In: Distributed Computing and Artificial Intelligence, 13th International Conference. Springer. 2016, pp. 247–254.

[30] K. Pittman. Comparison of data mining techniques used to predict student retention. Nova Southeastern University, 2008.

[31] L. Pournajaf, D. A. Garcia-Ulloa, L. Xiong e V. Sunderam. «Participant privacy in mobile crowd sensing task management: A survey of methods and challenges». In: ACM Sigmod Record 44.4 (2016), pp. 23–34.

[32] B. Rashidi e C. J. Fung. «A Survey of Android Security Threats and Defenses.» In: J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl. 6.3 (2015), pp. 3–35.

[33] F. Restuccia, S. K. Das e J. Payton. «Incentive mechanisms for participatory sensing: Survey and research challenges». In: ACM Transactions on Sensor Networks (TOSN) (2016), pp. 1–40.

[34] A. De Paola, P. Ferraro, G. Lo Re, M. Morana e M. Ortolani. «A fog-based hybrid intelligent system for energy saving in smart buildings». In: Journal of Ambient Intelligence and Humanized Computing 11.7 (2020), pp. 2793–2807.

[35] V. Agate, P. Ferraro e S. Gaglio. «A Cognitive Architecture for Ambient Intelligence Systems». In: AIC. 2018, pp. 52–58.

[36] C. Romero, S. Ventura, P. G. Espejo e C. Hervás. «Data mining algorithms to classify students». In: Educational data mining 2008.

[37] A. Rotondi, P. Pedroni e A. Pievatolo. «Analisi dei dati sperimentali». In: Probabilità, Statistica e Simulazione: Programmi applicativi scritti in R. Springer, 2021, pp. 509–562.

[38] S. Salerno, A. Sanzgiri e S. Upadhyaya. «Exploration of attacks on current generation smartphones». In: Procedia Computer Science 5 (2011), pp. 546–553.

[39] A. De Paola, P. Ferraro, S. Gaglio, G. Lo Re e S. K. Das. «An adaptive bayesian system for context-aware data fusion in smart environments». In: IEEE Transactions on Mobile Computing 16.6 (2016), pp. 1502–1515.

[40] Y. Shoukry, K. Gatsis, A. Alanwar, G. J. Pappas, S. A. Seshia, M. Srivastava e P. Tabuada. «Privacy-aware quadratic optimization using partially homomorphic encryption». In: 2016 IEEE 55th Conference on Decision and Control (CDC). IEEE. 2016, pp. 5053– 5058.

[41] B. Shrestha, D. Ma, Y. Zhu, H. Li e N. Saxena. «Tap-wave-rub: Lightweight human interaction approach to curb emerging smartphone malware». In: IEEE Transactions on Information Forensics and Security 10.11 (2015), pp. 2270–2283.

[42] V. Agate, F. Concone e P. Ferraro. «A Resilient Smart Architecture for Road Surface Condition Monitoring». In: The Proceedings of the International Conference on Smart City Applications. Springer. 2021, pp. 199–209.

[43] I. Souiden, M. N. Omri e Z. Brahmi. «A survey of outlier detection in high dimensional data streams». In: Computer Science Review 44 (2022), p. 100463.

[44] V. Agate, P. Ferraro, G. Lo Re e S. K. Das. «BLIND: A privacy preserving truth discovery system for mobile crowdsensing». In: Journal of Network and Computer Applications (2023), p. 103811.

[45] P. Strecht, L. Cruz, C. Soares, J. Mendes-Moreira e R. Abreu. «A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance.» In: International educational data mining society (2015).

[46] P. Strecht, J. Mendes-Moreira e C. Soares. «Merging Decision Trees: a case study in predicting student performance». In: Advanced Data Mining and Applications: 10th International Conference, ADMA 2014, Guilin, China, December 19-21, 2014. Proceedings Springer. 2014, pp. 535–548.

[47] F. Stulp e O. Sigaud. «Many regression algorithms, one unified model: A review». In: Neural Networks 69 (2015), pp. 60–79.

[48] V. Agate, A. De Paola, G. Lo Re e A. Virga. «Reliable Reputation-Based Event Detection in V2V Networks». In: International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability. Springer. 2023, pp. 267–281.

[49] C. Szongott, B. Henne e M. Smith. «Evaluating the threat of epidemic mobile malware». In: 2012 IEEE 8th international conference on wireless and mobile computing, networking and communications (WiMob). IEEE. 2012, pp. 443–450.

[50] P. Teufl, T. Zefferer e C. Stromberger. «Mobile device encryption systems». In: Security and Privacy Protection in Information Processing Systems: 28th IFIP TC 11 International Conference, SEC 2013, Auckland, New Zealand, July 8-10, 2013. Proceedings 28. Springer. 2013, pp. 203–216.

[51] Z. Wang, X. Pang, J. Hu, W. Liu, Q. Wang, Y. Li e H. Chen. «When mobile crowdsensing meets privacy». In: IEEE Communications Magazine 57.9 (2019), pp. 72–78.

[52] H.-T. Wu, Y.-m. Cheung e J. Huang. «Reversible data hiding in Paillier cryptosystem». In: Journal of Visual Communication and Image Representation 40 (2016), pp. 765–771.

[53] V. Agate, F. Concone, A. De Paola, P. Ferraro, G. Lo Re e M. Morana. «Bayesian Modeling for Differential Cryptanalysis of Block Ciphers: A DES Instance». In: IEEE Access 11 (2023),

pp. 4809–4820.

[54] A. Zafra e S. Ventura. «Predicting Student Grades in Learning Management Systems with Multiple Instance Genetic Programming.» In: International working group on educational data mining (2009).

[55] J. Zdziarski. «Identifying back doors, attack points, and surveillance mechanisms in iOS devices». In: Digital Investigation 11.1 (2014), pp. 3–19.

[56] Y. Zhang, N. Meratnia e P. Havinga. «A taxonomy framework for unsupervised outlier detection techniques for multi-type data sets». In: Computer 49.3 (2007), pp. 355–363.

[57] J. Zimmermann, K. H. Brodersen, J.-P. Pellet, E. August e J. M. Buhmann. «Predicting Graduate-level Performance from Undergraduate Achievements.» In: EDM. Citeseer. 2011, pp. 357–358.