ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ

Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

ΠΡΟΣΤΑΣΙΑ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΑΠΟ ΕΠΙΤΙΘΕΜΕΝΟΥΣ ΜΕ ΣΥΝΑΘΡΟΙΣΤΙΚΗ ΓΝΩΣΗ

ΠΛΗΡΟΦΟΡΙΕΣ: Όλγα Γκουντούνα, 210 772 1442, olga@dblab.ece.ntua.gr

Μανώλης Τερροβίτης, 210 6990522, mter@imis.athena-innovation.gr

ΠΕΡΙΛΗΨΗ: Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη και υλοποίηση αλγορίθμων ανωνυμοποίησης με σκοπό την προστασία από επιθέσεις που εκμεταλλεύονται γνώση συναθροιστικών συναρτήσεων πάνω στα δεδομένα.

ΑΤΟΜΑ: 1

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: C++

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Πολλοί οργανισμοί συλλέγουν δεδομένα από χρήστες τα οποία θα μπορούσαν να δημοσιευθούν ή να διανεμηθούν για ερευνητικούς σκοπούς. Η διανομή και χρήση αυτών των δεδομένων εγκυμονεί σοβαρούς κινδύνους για την παραβίαση τις ιδιωτικότητας. Ακόμα και με την απαλοιφή στοιχείων όπως το ονοματεπώνυμο, ΑΦΜ, κτλ., τα οποία προσδιορίζουν μοναδικά ένα άτομο, ο συνδυασμός άλλων στοιχείων, όπως ηλικία, φύλο και ταχυδρομικός κώδικας θα μπορούσαν να λειτουργήσουν ως *ψευδο-αναγνωριστικά* και να οδηγήσουν στην ταυτοποίηση του ατόμου και στην ταυτόχρονη αποκάλυψη ευαίσθητων προσωπικών στοιχείων του.

Για την προστασία της ιδιωτικότητας των ατόμων έχουν προταθεί διάφορες εγγυήσεις στη βιβλιογραφία [1, 2, 3]. Η ικανοποίηση των εγγυήσεων διασφαλίζεται με την εφαρμογή αλγορίθμων ανωνυμοποίησης στα προς δημοσίευση δεδομένα.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση όπου ένας κακόβουλος τρίτος χρησιμοποιεί κάποια συνολική γνώση που έχει για ένα πρόσωπο (π.χ. το συνολικό του εισόδημα) για να το αναγνωρίσει σε ένα σύνολο εγγραφών που περιέχουν αναλυτική πληροφορία (π.χ. εισοδήματα ανά δραστηριότητα). Το ερώτημα είναι αν μια τέτοια βάση δεδομένων θα μπορούσε να ανωνυμοποιηθεί κατάλληλα ώστε να μην ελλοχεύει κίνδυνος παραβίασης ιδιωτικότητας από κάποιο κακόβουλο επιτιθέμενο, ο οποίος μπορεί να εκμεταλλευτεί τη γνώση του συνολικού εισοδήματος.

Στην εργασία αυτή θα ερευνηθεί το μοντέλο επίθεσης με διάφορες συναθροιστικές συναρτήσεις πάνω στα δεδομένα και θα υλοποιηθούν αλγόριθμοι ανωνυμοποίησης που ελαχιστοποιούν την απώλεια πληροφορίας σε κάθε περίπτωση.

ΣΧΕΤΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ:

[1] L. Sweeney, *"k-Anonymity: A Model for Protecting Privacy."* International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, pp. 557-570, 2002.

[2] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam,*"l-Diversity: Privacy Beyond k-Anonymity."* in Proc. of ICDE, 2006.

[3] M. Terrovitis, N. Mamoulis, and P. Kalnis. *"Privacy-preserving anonymization of set-valued data."* in Proc. of VLDB, 2008.

[4] K. LeFevre, D. DeWitt, and R. Ramakrishnan. *"Incognito: Efficient full-domain k-anonymity."* in Proc. of SIGMOD, 2005.

[5] K. LeFevre, D. DeWitt, and R. Ramakrishnan. *"Mondrian multidimensional k-anonymity."* in Proc. of ICDE, 2006.