eSchool++: Σχεδιασμός συστήματος διαφύλαξης Ιδιωτικότητας στην Εξ Αποστάσεως Εκπαίδευση

Ε. Κωλέτσου[1], Ε. Αγγέλη, Δ. Καλογιάννης[2]

[1] Καθηγήτρια Πληροφορικής, MSc Πληροφορική-Λογισμικό, ekoletso@cs.uoi.gr
[2] Φοιτητές Τμήματος Τεχνολογίας Πληροφορικής και Τηλεπικοινωνιών, ΑΤΕΙ Ηπείρου, {agelena88, dimkalogiannis}@gmail.com

Περίληψη

Η προστασία της ιδιωτικότητας των δημοσιευμένων δεδομένων καθορίζεται από τον τρόπο της δημόσιας παρουσίασης δεδομένων γύρω από τις δραστηριότητες ή τις πράξεις ενός συνόλου ατόμων. Το *eSchool*++ (*electronic School preservation privacy*), αποτελεί ένα σύστημα για την διαφύλαξη των προσωπικών δεδομένων, που προέρχονται από συστήματα εξ αποστάσεως εκπαίδευσης. Στόχος του *eSchool*++ είναι να επιτρέπει σε ένα σύνολο καλοπροαίρετων χρηστών να έχουν πρόσβαση σε αυτά τα δεδομένα για ερευνητικούς σκοπούς, και παράλληλα να εμποδίζει κακόβουλους εισβολείς να συσχετίσουν αυτά τα δεδομένα με δεδομένα από εξωτερικές πηγές προκειμένου να συνδυάσουν συγκεκριμένα πρόσωπα στον πραγματικό κόσμο.

Λέξεις **κλειδιά:** *διαφύλαξη ιδιωτικότητας, εξ αποστάσεως εκπαίδευση.*

1. Εισαγωγή

Τα τελευταία χρόνια παρατηρείται ότι αρκετοί οργανισμοί διαχειρίζονται ένα σημαντικό μέγεθος δεδομένων, αποθηκευμένα σε κάποιο σύστημα βάσεων δεδομένων. Τα δεδομένα αυτά πολλές φορές προέρχονται από την καταγραφή ποικίλων ανθρώπινων δραστηριοτήτων, η μελέτη των οποίων είναι πολύτιμη σε διάφορους τομείς καθώς επιτρέπει να γίνεται πιο αποδοτικός ο σχεδιασμός των υπηρεσιών που οι ίδιοι αυτοί οργανισμοί προσφέρουν. Παράλληλα, όμως, δημιουργείται ένα νέο πρόβλημα, αυτό της προστασίας της ταυτότητας των ατόμων από τα οποία προέρχονται οι καταγραφές. Επιθυμία, λοιπόν, των οργανισμών είναι, επιπλέον, η διαφύλαξη της ιδιωτικότητας σε δημοσιευμένα δεδομένα, προσφέροντας εγγυήσεις ανωνυμίας.

Αποκρύπτοντας, απλά, στοιχεία που συνδέουν άμεσα την ταυτότητα ενός προσώπου με ένα σύνολο από δεδομένα, όπως το ονοματεπώνυμο ή το ΑΦΜ, δεν εξασφαλίζεται ότι η σύνδεση αυτή δεν θα αποκαλυφθεί. Μέσω διασταύρωσης των δημοσιευμένων δεδομένων με άλλες πηγές γνώσης (εξωτερικούς καταλόγους δεδομένων), μπορεί να αναγνωριστούν με μεγάλη βεβαιότητα οι εγγραφές που αφορούν στην ταυτότητα ενός συγκεκριμένου προσώπου.

Στα συστήματα εξ αποστάσεως εκπαίδευσης, η καταγραφή πληθώρας προσωπικών δεδομένων τόσο των εκπαιδευομένων όσο και των εκπαιδευτικών είναι μία συνήθης

διαδικασία. Οι πληροφορίες που απορρέουν από την καταγραφή του ιστορικού της απόδοσης των χρηστών, των προτιμήσεων τους, των κινήσεων τους μέσα στο σύστημα και από τη δόμηση του προσωπικού τους προφίλ, είναι απαραίτητες στην εκπαιδευτική διαδικασία, ώστε να μπορέσουν να προσφέρουν μια καλύτερη μορφή εξατομικευμένης μάθησης, συστάσεων και καθοδήγησης. Όλα αυτά επιτυγχάνονται με τη χρήση ευφυών πρακτόρων, οι οποίοι είναι υπεύθυνοι για τη συλλογή και επεξεργασία όλων αυτών των δεδομένων.

Εν συνεχεία, τα δεδομένα αυτά μπορεί να είναι στη διάθεση του φορέα υλοποίησης της εξ αποστάσεως εκπαίδευσης (σχολείου, οργανισμών ιδιωτικής εκπαίδευσης, κ.ά.) για οποιαδήποτε περεταίρω επεξεργασία, όπως για παράδειγμα στατιστικές μελέτες, κτλ. Το πρόβλημα της διαφύλαξης της ιδιωτικότητας των δεδομένων έγκειται στο γεγονός ότι τα δεδομένα αυτά δύναται να δημοσιευτούν και έτσι εκτίθενται στη διάθεση οποιουδήποτε επιθυμεί να τα επεξεργαστεί και για οποιοδήποτε σκοπό.

Η προσπάθεια στην ερευνητική περιοχή της προστασίας της ιδιωτικότητας είναι να αναπτυχθούν μέθοδοι επεξεργασίας των δεδομένων που αποτρέπουν την σύνδεση των δημοσιευμένων δεδομένων με συγκεκριμένα πρόσωπα. Παρακινούμενοι από τη σημασία της διαφύλαξης της ιδιωτικότητας των δεδομένων, παρουσιάζουμε μέσα από αυτή την εργασία το σχεδιασμό του συστήματος *eSchool++* (*electronic School preservation privacy*), ενός συστήματος επεξεργασίας προσωπικών δεδομένων που εγγυάται την ανωνυμοποίηση των δεδομένων που προέρχονται από συστήματα εξ αποστάσεως εκπαίδευσης.

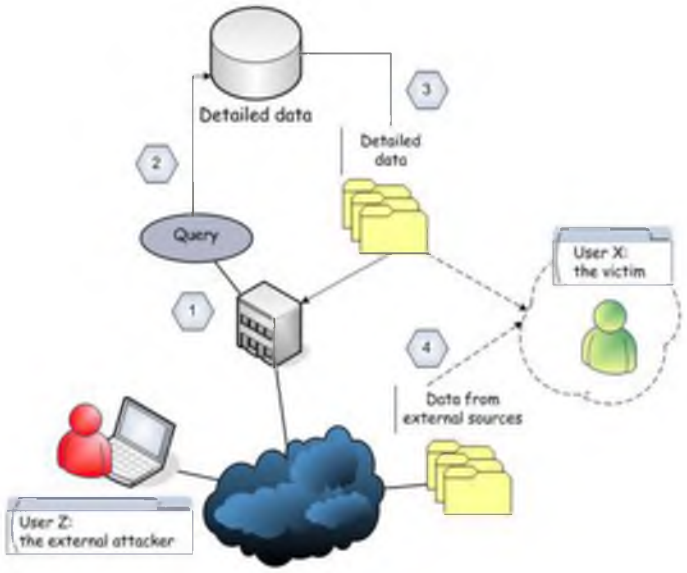
Στη συνέχεια, παρουσιάζουμε το πρόβλημα της ιδιωτικότητας σε δημοσιευμένα δεδομένα (Ενότητα 2). Επίσης, παρουσιάζουμε τεχνικές διαφύλαξης της ιδιωτικότητας (Ενότητα 3). Επιδεικνύουμε την σχεδίαση του συστήματος *eSchool++* (Ενότητα 4), καταλήγοντας σε ενδιαφέροντα συμπεράσματα (Ενότητα 5).

2. Το Πρόβλημα της Ιδιωτικότητας σε Δημοσιευμένα Δεδομένα

Η ραγδαία ανάπτυξη των τεχνολογιών της πληροφορικής και της επικοινωνίας, έχει επιφέρει σημαντικές κοινωνικοπολιτικές αλλαγές, αποτελώντας ταυτόχρονα την κύρια αιτία κινδύνων κατά των θεμελιωδών ελευθεριών και της ιδιωτικής ζωής του ατόμου. Οι κίνδυνοι αυτοί, στο πλαίσιο των συστημάτων εξ αποστάσεως εκπαίδευσης, απορρέουν από τη συλλογή και επεξεργασία προσωπικών δεδομένων, από τα αρχεία που δημιουργούν τόσο οι χρήστες μέσα στο σύστημα, όσο και από τα αρχεία που διατηρεί το ίδιο το σύστημα για τον κάθε χρήστη. Συνεπώς, δεδομένα που προέρχονται από την καταγραφή του ιστορικού των χρηστών, των προτιμήσεων και των ενδιαφερόντων τους, δύναται να είναι προσβάσιμα από όλους τους εμπλεκόμενους φορείς της εξ αποστάσεως εκπαίδευσης.

Σύμφωνα με τα παραπάνω, οι πληροφορίες που απορρέουν από τα αρχεία αυτά για τους χρήστες, δίνουν τη δυνατότητα να κατασκευαστεί μια γενική εικόνα τους. Επίσης, με κατάλληλη επεξεργασία αυτών των αρχείων και σε συνδυασμό με αρχεία που προέρχονται από εξωτερικές πηγές και είναι εύκολο να ανακτηθούν, για

παράδειγμα μέσω του διαδικτύου, θα μπορούσε να κατασκευαστεί επιπλέον μία πιο εξειδικευμένη εικόνα των χρηστών. Μέσα από αυτούς τους τρόπους, είναι δυνατόν να αποκαλυφτούν πολιτικές και θρησκευτικές πεποιθήσεις, αξίες, φιλοδοξίες, ευαίσθητα ιατρικά δεδομένα, και ακόμη, ένα μεγάλο μέρος των προτιμήσεων και γούστων των χρηστών ως καταναλωτές, γεγονός που αποκτά μεγάλη αξία για εμπορικούς σκοπούς (Εικόνα 1).



Εικόνα 1: *Το πρόβλημα της ιδιωτικότητας σε δημοσιευμένα δεδομένα*

Το πρόβλημα της προστασίας της ιδιωτικότητας των προσωπικών δεδομένων εντοπίζεται ακόμη και μέσα σε ολοκληρωμένα συστήματα εξ αποστάσεως εκπαίδευσης. Αυτό μπορεί να προκύψει βάσει καταλόγων με ποικίλη πληροφορία σχετικά με τους χρήστες, που δημοσιεύονται σε αυτά τα συστήματα και μπορεί να οδηγήσουν στην αναγνώριση ευαίσθητων δεδομένων που τους αφορούν. Μάλιστα, σοβαροί κίνδυνοι παραβίασης ενέχονται από την επεξεργασία των δεδομένων των χρηστών, πέραν του αρχικού σκοπού για τον οποίο συλλέχθηκαν. Κάτι τέτοιο θα μπορούσε, για παράδειγμα, να συμβεί στην περίπτωση που τα συστήματα εξ αποστάσεως εκπαίδευσης αποτελούν τμήματα ενός μεγαλύτερου οργανισμού ή εποπτεύεται από υψηλότερες διοικητικές δομές, και αναμένεται από αυτά να συμμετέχουν σε οποιοδήποτε είδος διασύνδεση και κοινοποίηση δεδομένων.

Πιο συγκεκριμένα, δίνουμε το ακόλουθο παράδειγμα. Έστω ένα σχολείο που έχει αποφασίσει να ενισχύσει τα μαθήματά του με εξ αποστάσεως μαθήματα, ενώ παράλληλα εποπτεύεται από υψηλότερες διοικητικές δομές (π.χ. γραφεία Α/βαθμιας ή Β/βαθμιας διεύθυνσης, περιφέρειας, υπουργείου κτλ.). Ο Πίνακας 1 περιέχει μια σειρά από δεδομένα τα οποία αφορούν τους μαθητές του σχολείου αυτού. Στον πίνακα υπάρχουν πεδία όπως το όνομα, η ημερομηνία γέννησης, το φύλο και ο

ταχυδρομικός κώδικας των μαθητών. Τα δεδομένα αυτά δημοσιεύονται από τον δάσκαλο στο χώρο του αντίστοιχου μαθήματος για ενημέρωση/καταχώρηση των μαθητών.

Πίνακας 1: *Στοιχεία μαθητών/χρηστών πλατφόρμας εξ αποστάσεως εκπαίδευσης*

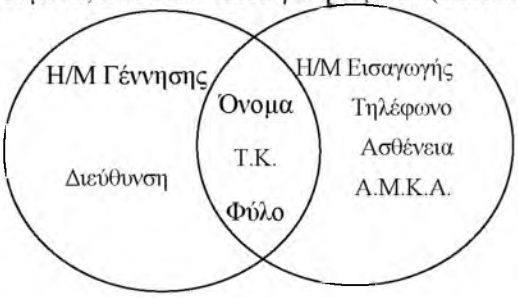
Όνομα	Η/Μ Γεν.	Φύλο	Διεύθυνση	Τ.Κ.
Ελένη	14.05.2005	Κ	Αρετής 2	34000
Μάριος	04.03.2005	Α	Ιωνίας 9	34120
Δημήτρης	05.08.2004	Α	Έβανς 7	34222
Ευαγγελία	31.01.2003	Κ	Ηπείρου 4	34345
Αλέξανδρος	25.05.2003	Α	Β.Ηπείρου 5	34090
Αμαλία	13.03.2002	Κ	Ζάρρα 6	34023
Γιάννης	05.05.2001	Α	Κιλκίς 12	34234
Γιώτα	03.03.2001	Κ	Κωστακιοί 35	34235
Μαρίνα	21.11.2000	Κ	Ανέζας 65	34234
Ελένη	15.02.2000	Κ	Ιμαρέτ 9	34212
Μάριος	16.02.2000	Α	Σόλωνος 123	34245
Θάνος	19.09.2000	Α	Δημοκρατίας 90	34267

Εκτός, όμως, από τους μαθητές-χρήστες που έχουν πρόσβαση σε αυτά τα δεδομένα, έχουν εξίσου πρόσβαση και εξουσιοδοτημένοι χρήστες που ανήκουν στις υψηλότερες από το σχολείο διοικητικές δομές. Τέτοιοι χρήστες, θα μπορούσαν να ανακτήσουν αυτά τα δεδομένα για ερευνητικούς σκοπούς και στατιστικές αναλύσεις. Τι γίνεται, όμως, στην περίπτωση που θα ήθελαν αυτή την πληροφορία να την εκμεταλλευτούν συνδυαστικά με αποτέλεσμα κοινοποίησης ευαίσθητων δεδομένων; Στην περίπτωση αυτή, θα μπορούσε να γίνει ταυτοποίηση των δεδομένων με άλλους πίνακες που αναρτώνται κατά καιρούς στο διαδίκτυο και περιέχουν ευαίσθητες πληροφορίες για κάποιους από τους μαθητές. Τέτοιες πληροφορίες θα μπορούσαν να αντληθούν από νοσοκομειακούς πίνακες που είναι προσβάσιμοι σε υπαλλήλους κάποιου νοσοκομείου της περιοχής. Η ταυτοποίηση αυτή θα είναι επιτυχής, καθώς οι πίνακες περιλαμβάνουν το όνομα, τον ταχυδρομικό κώδικα, το τηλέφωνο, τον αριθμό Α.Μ.Κ.Α, την πάθηση του κάθε ασθενή, και την ημερομηνία εισαγωγής στο νοσοκομείο, του κάθε ασθενή.

Έχοντας συγκεντρώσει όλες αυτές τις πληροφορίες και λαμβάνοντας υπόψη την ημερομηνία γέννησης του μαθητή σε συνδυασμό με τα πρώτα έξι (6) ψηφία του αριθμού Α.Μ.Κ.Α, που αντιστοιχούν στην ημερομηνία γέννησης του ατόμου, καθώς και τις ημέρες απουσίας του μαθητή από το σχολείο και την ημερομηνία νοσηλείας του, μπορεί να συμπεράνει ότι, για παράδειγμα, ο μαθητής Δημήτρης που είναι αγόρι και έχει ταχυδρομικό κώδικα 49000, πάσχει από Λευχαιμία.

Με το πιο πάνω παράδειγμα, θέλουμε να δείξουμε τη δυνατότητα αναγνώρισης προσώπων κάνοντας άμεση σύνδεση πληροφορίας, που προέρχεται από

δημοσιευμένα δεδομένα, πάνω σε κοινά γνωρίσματα (Εικόνα 2).



Εικόνα 2: *Σύνδεση πληροφορίας από δημοσιευμένα δεδομένα*

Συνεπώς, με την πρόοδο της τεχνολογίας οι κίνδυνοι διευρύνονται συνεχώς και τίθεται επιτακτικά το ζήτημα της προστασίας της ιδιωτικής ζωής του ατόμου, δηλαδή της προστασίας των προσωπικών του δεδομένων (Ιγγλεζάκης, 2003). Έτσι, θα μπορούσαμε να αναφερθούμε στον κίνδυνο εκμετάλλευσης του ατόμου από τυχόν αρνητικά δεδομένα του, τα οποία μπορούν να διαιωνίζονται και μάλιστα να αξιοποιούνται ηλεκτρονικά στο διηνεκές (Σταθόπουλος, 2000). Μέσω μίας τέτοιας εκμετάλλευσης δεδομένων ελλοχεύει παράλληλα και ο κίνδυνος κατηγοριοποίησης του ατόμου από την σχηματική ταξινόμηση και κατάταξη των δεδομένων του, με συνέπεια τον ευκολότερο έλεγχό του, τις διακρίσεις σε βάρος του και τη χειραγώγησή του. Ακόμη, η εμπορευματοποίηση και η εξουδετέρωση της ατομικότητας, και της μετατροπής της ιδιωτικής ζωής σε αντικείμενο οικονομικών συναλλαγών (Ανθίμου, 2000) με υπέρμετρη πληροφοριακή εξουσία από αυτούς που ελέγχουν τις πληροφορίες, είναι πιθανοί κίνδυνοι. Τέλος, η προσβολή του δικαιώματος κάθε ατόμου στον έλεγχο και στον συμπροσδιορισμό της χρήσης των πληροφοριών που τον αφορούν, με άμεσο αντίκτυπο την αποχή του ατόμου από την άσκηση του δικαιώματος πρόσβασης στην πληροφορία, είναι θέματα που θα πρέπει να μας προβληματίζουν (Στρακαντούνα, 2004).

Ανάγεται, λοιπόν, το συμπέρασμα ότι η ισχυρή προστασία της ιδιωτικότητας είναι βασική προϋπόθεση για την ορθή λειτουργία και εξέλιξη των σύγχρονων συστημάτων εξ αποστάσεως εκπαίδευσης, και η έλλειψή της μπορεί να έχει αντίκτυπο σε άλλα βασικά δικαιώματα των χρηστών τους.

3. Τεχνικές Διαφύλαξης της Ιδιωτικότητας

Η ανακάλυψη της ταυτότητας μιας εγγραφής (δηλαδή σε ποιο άτομο ανήκει αυτή η εγγραφή μέσα από μία σχεσιακή βάση δεδομένων) από διάφορους εξωτερικούς παράγοντες, συνεπώς, είναι εφικτή. Όπως έχουμε ήδη αναφέρει, δεν αρκεί πάντα η αφαίρεση διαφόρων αναγνωριστικών για να μας εξασφαλίσει την ανωνυμία. Η προφανής λύση στο πρόβλημα μας είναι να αφαιρέσουμε όσες ιδιότητες μπορούνε να συνδυαστούν με κάποιον εξωτερικό πίνακα. Αυτό όμως άμεσα συνεπάγεται με απώλεια πληροφορίας, γεγονός μη επιθυμητό.

Οι πιο συνήθεις εγγυήσεις για την προστασία των προσωπικών δεδομένων είναι αυτές της *k-ανωνυμίας* (*k-anonymity*) που εγγυάται ότι καμία εγγραφή δεν θα μπορεί να διακριθεί σε σχέση με άλλες *k*-1 εγγραφές (Samarati, 2001; Sweeney, 2002), και η *l-διαφορετικότητα* (*l-diversity*) που εγγυάται ότι θα υπάρχουν *l* διαφορετικές απαντήσεις σε κάθε πιθανή προσπάθεια ανάκτησης προσωπικών δεδομένων (Machanavajjhala et al., 2006). Οι περισσότερες προσεγγίσεις υποθέτουν ότι τα δεδομένα είναι οργανωμένα αυστηρά στο σχεσιακό μοντέλο, και οι περισσότερες μέθοδοι ανωνυμοποίησης δουλεύουν για τέτοια δεδομένα. Στην πράξη όμως τα δεδομένα τα οποία καταγράφουν διάφορες δραστηριότητες είναι αρκετά πιο σύνθετα. Έτσι, μία νέα τεχνική διαφύλαξης της ιδιωτικότητας είναι εκείνη της *km-ανωνυμίας* (*km-anonymity*), όπου τα δεδομένα δεν διακρίνονται σε ευαίσθητα και μη-ευαίσθητα, αλλά ότι μπορούν να λειτουργήσουν ταυτόχρονα και σαν ψευδοαναγνωστικά και σαν ευαίσθητα δεδομένα, ανάλογα με την προοπτική που αντιπάλου (Terrovitis et al., 2010).

Βασιζόμενοι στις προαναφερθείσες τεχνικές διαφύλαξης της ιδιωτικότητας των δεδομένων, προτείνουμε το σχεδιασμό του συστήματος *eSchool++*, όπου το μοντέλο ανωνυμοποίησης βασίζεται στην γενίκευση αντί στην απόκρυψη δεδομένων. Με τη γενίκευση γίνεται αντικατάσταση των τιμών των πεδίων δεδομένων με ολοένα και πιο γενικών τιμών, μέχρις ότου κάθε εγγραφή να ενταχθεί σε μια μεγάλη ομάδα εγγραφών με ίδιες τιμές. Έτσι, αντικαθιστούμε, για παράδειγμα την τιμή μιας διεύθυνσης με την πόλη ή τη χώρα στην οποία ανήκει. Ο διαχειριστής του συστήματος *eSchool++* μπορεί να επιλέξει ελεύθερα το βαθμό της γενίκευσης που θα εφαρμόσει στο δικό του σύστημα βάσεων δεδομένων.

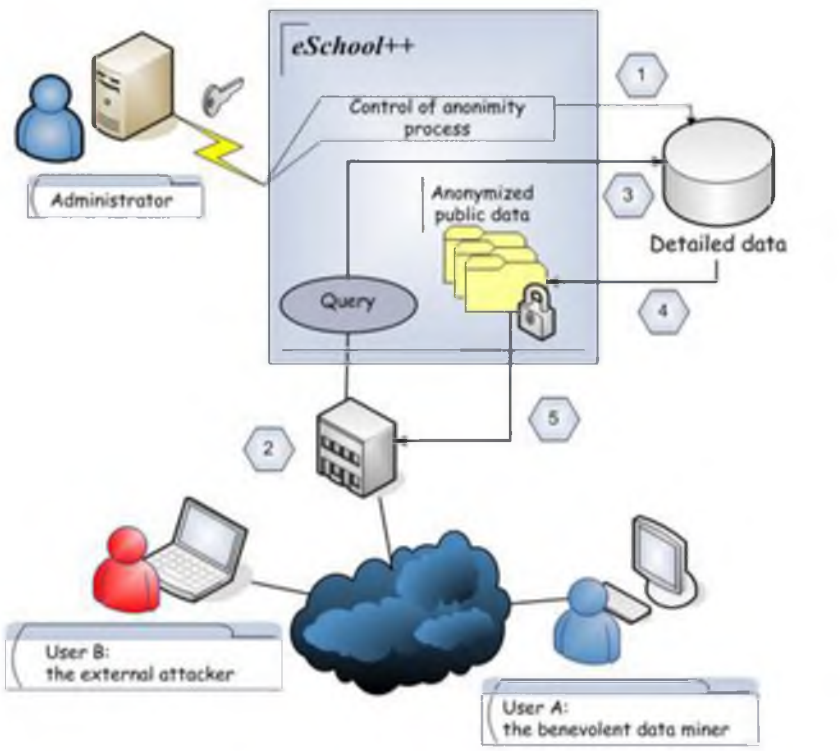
4. Το Σύστημα *eSchool++*

Η βασική ιδέα, πάνω στην οποία έγινε ο σχεδιασμός του συστήματος μας, είναι ότι το *eSchool++* μπορεί, με καθορισμένες τροποποιήσεις, να εφαρμοστεί σε οποιοδήποτε σύστημα εξ αποστάσεως εκπαίδευσης. Αυτή η δυνατότητα υπονοεί ότι το σύστημα *eSchool++* είναι ανεξάρτητο της οποιαδήποτε σχεσιακής βάσης δεδομένων, και δομείται πάνω σε αυτή (on-top) χωρίς να επηρεάζει τη δομή της. Η αρχιτεκτονική του συστήματος *eSchool++* απεικονίζεται στην Εικόνα 3.

Μέσω ενός εύχρηστου και διαδραστικού περιβάλλοντος, ο *διαχειριστής* (*administrator*) του συστήματος εξ αποστάσεως εκπαίδευσης, καλείται να ορίσει το *κριτήριο ανωνυμοποίησης* που επιθυμεί. Το *eSchool++* παρέχει τρία διαφορετικά κριτήρια ανωνυμοποίησης, που συνδέονται άμεσα με ένα σύνολο εγγραφών, διαφορετικό σε κάθε κριτήριο, και καθορίζονται δυναμικά σε σχέση με τη σχεσιακή βάση δεδομένων του συστήματος. Για να επιτευχθεί το κατάλληλο επίπεδο γενίκευσης, σύμφωνα με το κριτήριο ανωνυμοποίησης που έχει επιλεγεί από τον διαχειριστή, το αντίστοιχο σύνολο εγγραφών θα πρέπει να διαγραφεί. Όσο υψηλότερο είναι το επίπεδο γενίκευσης, τόσο περισσότερες εγγραφές διαγράφονται. Ο λόγος της απόδοσης του συστήματος, δηλαδή του καλύτερου δυνατού αποτελέσματος ανωνυμοποίησης, και του χρόνου προεπεξεργασίας των δεδομένων

είναι αντιστρόφως ανάλογος. Πιο αναλυτικά, για το χαμηλότερο επίπεδο γενίκευσης απαιτείται λιγότερος χρόνος προεπεξεργασίας των δεδομένων της σχεσιακής βάσης δεδομένων –γίνονται λιγότερες διαγραφές– ενώ για το υψηλότερο επίπεδο γενίκευσης, όπου πετυχαίνουμε πολύ καλά αποτελέσματα ανωνυμοποίησης, ο χρόνος προεπεξεργασίας των δεδομένων είναι αισθητά πολύ μεγαλύτερος. Συνεπώς, ο διαχειριστής του συστήματος μπορεί να επιλέξει τη διαγραφή περισσότερων εγγραφών για να κερδίσει περισσότερη ιδιωτικότητα, ή το αντίστροφο. Για τη μη απώλεια της αρχικής πληροφορίας που εμπεριέχεται στη σχεσιακή βάση δεδομένων, το *eSchool++* ζητάει από τον διαχειριστή του συστήματος, πριν από οποιαδήποτε τροποποίηση, να διατηρήσει αντίγραφα ασφαλείας της βάσης δεδομένων.

Σε δεύτερη φάση, διάφοροι εξουσιοδοτημένοι χρήστες μέσω της κατάλληλης διεπαφής μπορούν να συνδεθούν στην πλατφόρμα της εξ αποστάσεως εκπαίδευσης, και να ζητήσουν πληροφορίες που είναι καταγεγραμμένες στη σχεσιακή βάση δεδομένων του συστήματος. Οι χρήστες μπορεί να είναι είτε καλοπροαίρετοι (*User A*) και να επιθυμούν πρόσβαση σε αυτά τα δεδομένα για ερευνητικούς σκοπούς, είτε κακόβουλοι εισβολείς που επιθυμούν να ανακτήσουν πληροφορία και να τη συσχετίσουν με δεδομένα από εξωτερικές πηγές προκειμένου να συνδυάσουν συγκεκριμένα πρόσωπα στον πραγματικό κόσμο (*User B*). Ανεξάρτητα, λοιπόν, με τις προθέσεις των χρηστών, τα ερωτήματα για ανάκτηση πληροφορίας μέσω του *eSchool++* καταγράφονται και με δυναμική δημιουργία κατάλληλων όψεων προωθούνται στο σύστημα διαχείρισης της βάσης δεδομένων. Το σύνολο των αποτελεσμάτων που επιστρέφεται τελικά στους χρήστες δεν είναι τα λεπτομερή δεδομένα που είναι πραγματικά καταγεγραμμένα στο σύστημα, αλλά ένα σύνολο από πιο γενικευμένα αποτελέσματα (Εικόνα 3).



Εικόνα 3: *Αρχιτεκτονική του συστήματος eSchool++*

Για να γίνει πιο κατανοητό, δίνουμε τη γενικευμένη μορφή του παραδείγματος που αναλύσαμε προηγουμένως στην Ενότητα 2. Έστω, λοιπόν, ένας στατιστικός αναλυτής, είτε με καλές προθέσεις είτε κακόβουλος, ο οποίος έχει δικαίωμα πρόσβασης στα δεδομένα της πλατφόρμας εξ αποστάσεως εκπαίδευσης. Έχοντας συνδυάσει δεδομένα από εξωτερικές πηγές, για παράδειγμα στατιστικά νοσοκομείου, μπορεί να βγάλει συμπεράσματα και να αποκαλύψει προσωπικές ευαίσθητες πληροφορίες για κάποιον μαθητή. Με τη μέθοδο γενίκευσης που χρησιμοποιούμε στο σύστημα *eSchool++*, προσπαθούμε να αποτρέψουμε την εξαγωγή οποιασδήποτε πληροφορίας, μέσω της ταυτοποίησης προσώπων, όταν αυτή συνδυάζεται με δεδομένα από άλλες εξωτερικές πηγές. Μέσω της ομαδοποίησης δεδομένων που αντιπροσωπεύουν τους μαθητές, δίνουμε στους χρήστες *User A* και *User B*, αποτελέσματα τέτοια που μπορούν να χρησιμοποιηθούν για ερευνητικούς σκοπούς, στατιστικές αναλύσεις κτλ., προστατεύοντας παράλληλα την ιδιωτικότητα των προσωπικών δεδομένων των μαθητών. Παρακάτω παρουσιάζεται ομαδοποιημένος ο αρχικός πίνακας με τα δεδομένα των μαθητών μέσα στο σύστημα εξ αποστάσεως εκπαίδευσης που χρησιμοποιείται από το σχολείο, όπως θα παρουσιαζόταν σε κάποιον εξωτερικό/εξουσιοδοτημένο χρήστη της πλατφόρμας, καλόβουλο ή κακόβουλο, προφυλάσσοντας έτσι όλα τα προσωπικά δεδομένα των μαθητών (Πίνακας 2).

Πίνακας 2: *Στοιχεία μαθητών/χρηστών πλατφόρμας εξ αποστάσεως εκπαίδευσης, όπως παρουσιάζονται σε εξωτερικούς/εξουσιοδοτημένους χρήστες της πλατφόρμας*

Όνομα	Ηλικία	Φύλο	Περιοχή	Τ.Κ.
*	< 10	Α	Αττική	34***
*		Α		34***
*		Α		34***
*		Κ		34***
*		Κ		34***
*		Κ		34***
*	1*	Α		342**
*		Α		342**
*		Α		342**
*		Κ		342**
*		Κ		342**
*		Κ		342**

Από το σχήμα και τον πίνακα με τα τροποποιημένα δεδομένα, γίνεται σαφής ο τρόπος με τον οποίο προστατεύουμε τις ευαίσθητες πληροφορίες που αφορούν μοναδικά κάθε άτομο. Όπως αναφέρθηκε προηγουμένως, είναι σίγουρο πως με τη γενίκευση των δεδομένων, καμία πληροφορία δεν πρόκειται να διαρρεύσει και η χρήση της από οποιονδήποτε χρήστη είναι ασφαλής.

5. Συμπεράσματα

Μέσα από αυτή την εργασία έγινε μία παρουσίαση του σχεδιασμού του συστήματος *eSchool++*, ενός συστήματος που παρέχει μεθόδους διαφύλαξης της ιδιωτικότητας σε δημοσιευμένα δεδομένα, τα οποία προέρχονται από συστήματα εξ αποστάσεως εκπαίδευσης. Μέσω του *eSchool++* παρέχεται στο διαχειριστή του συστήματος εξ αποστάσεως εκπαίδευσης η δυνατότητα επιλογής του κριτηρίου ανωνυμοποίησης των δεδομένων, επιτρέποντάς του να αποφασίσει είτε τη διαγραφή περισσότερων εγγραφών για να κερδίσει περισσότερη ιδιωτικότητα, είτε το αντίστροφο. Σε κάθε περίπτωση, το *eSchool++* διατηρεί αντίγραφα ασφαλείας της αρχικής δομής και πληροφορίας της σχεσιακής βάσης δεδομένων. Έτσι, τόσο καλόβουλοι χρήστες όσο και κακόβουλοι εισβολείς, που έχουν πρόσβαση στο σύστημα, μπορούν να κάνουν πλέον ανάκτηση γενικευμένων δεδομένων, τα οποία συνδυάζοντας τα με δεδομένα εξωτερικών πηγών δεν δύναται να οδηγήσουν στην αναγνώριση και ταυτοποίηση συγκεκριμένων ατόμων στον πραγματικό κόσμο.

Θεωρώντας πως το *eSchool++*, θα επιτελέσει το έργο του στο μέγιστο βαθμό, θα υπάρχει μία συνεχής προσπάθεια για την κάλυψη των νέων αναγκών, που πιθανώς να προκύψουν, κατά τη φάση υλοποίησης, αλλά και συντήρησής του, με ενδεχόμενη επέκταση των δυνατοτήτων του.

Βιβλιογραφία

Ανθίμου, Κ. (2000). *Προστασία προσωπικών δεδομένων και ΜΜΕ*, Κριτική Επιθεώρηση (ΚριτΕ) 1 (σελ. 271-318).

Ιγγλεζάκης, Ι. (2003). *Ευαίσθητα προσωπικά δεδομένα*, Εκδόσεις Σάκκουλα (σελ. 10). Αθήνα.

Machanavajjhala, A., Gehrke, J., Kifer, D. (2006). *l-diversity: Privacy beyond k-anonymity*. In ICDE.

Samarati, P. (2001). *Protecting respondents' identities in microdata release*. IEEE Trans. Knowl. Data Eng. (TKDE), 13(6):1010–1027.

Σταθόπουλος, Μ., (2000). *Η χρήση προσωπικών δεδομένων και η διαπάλη μεταξύ των κατόχων τους και ελευθεριών των υποκειμένων τους*, Νομικό Βήμα (NoB) , 48(1) (σελ. 1-19).

Στρακαντούνα, Β. (2004). *Επεξεργασία προσωπικών δεδομένων και προστασία της ιδιωτικότητας στο σύγχρονο περιβάλλον των βιβλιοθηκών και υπηρεσιών πληροφόρησης*. Ιόνιο Πανεπιστήμιο (σελ. 39-41). Ανακτήθηκε 10/11/2010, από τη διεύθυνση http://dlib.ionio.gr/mtheses/strakantouna_privacy.pdf

Sweeney, L. (2002*). k-Anonymity: A Model for Protecting Privacy*. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10(5): 557-570.

Terrovitis, M., Mamoulis, N., Kalnis, P. (2010). *Local and Global Recoding Methods for Anonymizing Set-valued Data*. The VLDB Journal.