Πανεπιστήμιο Δυτικής Αττικής
Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών

ΠΜΣ Κυβερνοασφάλεια
Πτυχιακή εργασία

Υποθέσεις χρήσης τεχνικών Ψευδωνυμοποίησης στον χώρο της Δημόσιας Υγείας και της Εκπαίδευσης

Ιωάννης Μυτάκος
Επιβλέπων καθηγητής: Στέφανος Γκρίτζαλης

Η παρούσα διπλωματική εργασία παρουσιάστηκε

από τον

Ιωάννη Μυτάκο

AM: cscyb19016

Εισηγητής: Π.Γιαννακόπουλος

ΕΠΙΤΡΟΠΗ ΕΞΕΤΑΣΗΣ

Α/Α	ΟΝΟΜΑ ΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ/ΤΜΗΜΑ	ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ
1	Στέφανος Γκρίτζαλης	Καθηγητής Πανεπιστημίου Πειραιά Μέλος εξεταστικής επιτροπής	STEFANOS GKRITZALIS 25/08/2023 14:59
2	Παναγιώτης Γιαννακόπουλος	Καθηγητής Πανεπιστήμιο Δυτικής Αττικής Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών / Εισηγητής	Panagiotis Giannakopoulos Digitally signed by Panagiotis Giannakopoulos Date: 2023.08.23 15:17:40 +03'00'
3	Εμμανουήλ Μιχαηλίδης	Ακαδημαϊκός Υπότροφος Πανεπιστήμιο Δυτικής Αττικής Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών / Μέλος Εξεταστικής Επιτροπής	Emmanouil Michailidis Digitally signed by Emmanouil Michailidis Date: 2023.08.23 15:29:21 +03'00'

Δήλωση συγγραφέα μεταπτυχιακής εργασίας

Ο κάτωθι υπογράφων Ιωάννης Μυτάκος μεταπτυχιακός φοιτητής του προγράμματος σπουδών «Κυβερνοασφάλεια» του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής δηλώνω ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποίαν είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης οι πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών η λέξεων είτε ακριβώς είτε παραφρασμένες αναφέρονται στο σύνολό τους με πλήρη αναφορά στους συγγραφείς, στον εκδοτικό οίκο ή το περιοδικό συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο.

Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από εμένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου.»

Ο δηλών

IOANNIS MYTAKOS OR MITAKOS
23/08/2023 15:00

Ευχαριστούμε τον καθηγητή κ. Στέφανο Γκρίτζαλη για την υποστήριξη και την καθοδήγησή του για την ολοκλήρωση της εργασίας

Περίληψη

Η ψευδωνυμοποίηση είναι ένα μέτρο προστασίας των προσωπικών δεδομένων των χρηστών κατά την διάρκεια της επεξεργασίας τους, που προτείνεται από το Γενικό Κανονισμό Προστασίας Δεδομένων. Η πρόσφατη πανδημία δημιούργησε την ανάγκη ανάπτυξης ή χρήσης εφαρμογών που περιλαμβάνουν την συλλογή και επεξεργασία δεδομένων προσωπικού χαρακτήρα πολύ μεγάλου αριθμού προσώπων, η οποία πρέπει να γίνεται με τρόπο που να διασφαλίζει την προστασία τους. Πλέον της κανονιστικής συμμόρφωσης, η λήψη επαρκών μέτρων υπάρχει πιθανότητα να αυξήσει την εθελοντική συμμετοχή των πολιτών. Έγινε προσπάθεια να παρουσιαστούν ρεαλιστικά σενάρια χρήσης εφαρμογών στην δήλωση αποτελέσματος διαγνωστικού ελέγχου και την ινηλάτηση επαφών στον χώρο της υγείας, και την εκπαίδευση μέσω τηλεδιάσκεψης στον χώρο της εκπαίδευσης.

Το κεφάλαιο 1 περιλαμβάνει μια επισκόπηση των κανονιστικών απαιτήσεων που θέτει ο γενικός κανονισμός προστασίας δεδομένων και τις πρόσθετες απαιτήσεις που είναι δυνατόν να επιβάλλει η αντιμετώπιση μιας κατάστασης έκτακτης ανάγκης.

Στο κεφάλαιο 2 εξετάζονται οι υπάρχουσες τεχνικές ψευδωνυμοποίησης προσπαθώντας να κωδικοποιήσουμε τα κριτήρια επιλογής των τεχνικών μέτρων, των πολιτικών καθώς και των οργανωτικών μέτρων εφαρμογής της ψευδωνυμοποίησης, σε σχέση με τις απαιτήσεις.

Στο κεφάλαιο 3 διατυπώνονται σενάρια χρήσης ψευδωνυμοποίησης σε εφαρμογές υγείας. Εξετάζεται η δήλωση αποτελέσματος διαγνωστικού ελέγχου και ακολούθως οι εφαρμογές Αυτόματης Ιχνηλάτησης Επαφών, όπου παρουσιάζονται τρία αντιπροσωπευτικά ανοιχτά πρωτόκολλα, το PEPP-PT, το DP-3T και το Pronto-C2.

Στο κεφάλαιο 4 διατυπώνονται σενάρια χρήσης ψευδωνυμοποίησης σε εφαρμογές εκπαίδευσης μέσω τηλεδιάσκεψής, όπως η ψευδωνυμοποίηση των αναγνωριστικών χρήστη και των IP διευθύνσεων.

Λέξεις - κλειδιά

Pseudonymization, general data protection regulation, automatic contact tracing, digital contact tracing, distance education, remote teaching

Μεθοδολογία

Για την παρουσίαση ρεαλιστικών σεναρίων χρήσης δεν χρησιμοποιήθηκαν υπηρεσιακές, εσωτερικές ή διαβαθμισμένες πληροφορίες, αλλά μόνο πληροφορίες με ευρεία δημοσιοποίηση που δημοσιεύθηκαν από επίσημες πηγές στο διαδίκτυο.

1 Εισαγωγή

1.1 Ψευδωνυμοποίηση και ΓΚΠΔ

Η έναρξη ισχύος του Γενικού Κανονισμού Προστασίας Δεδομένων (General Data Protection Regulation - GDPR) άλλαξε τον τρόπο επεξεργασίας των δεδομένων προσωπικού χαρακτήρα των Ευρωπαίων πολιτών. Λόγω του παγκόσμιου χαρακτήρα του Διαδικτύου αλλά και άλλων δικτύων που επιτρέπουν την διενέργεια συναλλαγών ανεξαρτήτως της απόστασης των συναλλασσόμενων, ο GDPR δημιούργησε υποχρεώσεις σε οργανισμούς και φορείς σε όλο τον κόσμο, εφόσον αυτοί επεξεργάζονται δεδομένα προσωπικού χαρακτήρα Ευρωπαίων πολιτών.

Η ψευδωνυμοποίηση είναι η διαδικασία που μπορεί να περιλαμβάνει ποικίλες τεχνικές με σκοπό την απόκρυψη της ταυτότητας των υποκειμένων των δεδομένων με την αντικατάσταση των στοιχείων ταυτοποίησης από ψευδώνυμα στοιχεία και ορίζεται στην περίπτωση 5 του Άρθρου 4 του GDPR ως εξής:

«η επεξεργασία δεδομένων προσωπικού χαρακτήρα κατά τρόπο ώστε τα δεδομένα να μην μπορούν πλέον να αποδοθούν σε συγκεκριμένο υποκείμενο των δεδομένων χωρίς τη χρήση συμπληρωματικών πληροφοριών, εφόσον οι εν λόγω συμπληρωματικές πληροφορίες διατηρούνται χωριστά και υπόκεινται σε τεχνικά και οργανωτικά μέτρα προκειμένου να διασφαλιστεί ότι δεν μπορούν να αποδοθούν σε ταυτοποιημένο ή ταυτοποιήσιμο φυσικό πρόσωπο.»

Πρακτικά, τα δεδομένα που προσδιορίζουν την ταυτότητα του υποκειμένου αντικαθίστανται από «ψευδώνυμα» δεδομένα, τα οποία δεν είναι δυνατόν να αντιστοιχιστούν στα αρχικά δεδομένα ταυτοποίησης χωρίς επιπρόσθετες πληροφορίες.

Σύμφωνα με τον παραπάνω ορισμό, μπορούμε να παρατηρήσουμε ότι προβλέπεται η διατήρηση των συμπληρωματικών πληροφοριών, συνεπώς είναι θεωρητικά τεχνικώς δυνατή η υπό όρους άρση της ψευδωνυμοποίησης, για τον λόγο αυτό, από τον ΓΚΠΔ τα ψευδωνυμοποιημένα δεδομένα θεωρούνται πληροφορίες που αφορούν σε φυσικό πρόσωπο που είναι δυνατόν να ταυτοποιηθεί. Συνεπώς, εφαρμόζονται σε αυτά οι αρχές προστασίας δεδομένων του κανονισμού, οι οποίες δεν εφαρμόζονται σε ανώνυμα ή πλήρως ανωνυμοποιημένα δεδομένα.

Πρέπει να αναφερθεί εδώ ότι το Άρθρο 15 του ΓΚΠΔ δίνει στο υποκείμενο της επεξεργασίας την δυνατότητα πρόσβασης στα προσωπικά του δεδομένα τα οποία διατηρεί ο υπεύθυνος επεξεργασίας. Συνεπώς ο υπεύθυνος επεξεργασίας, ακόμα και αν τα δεδομένα ταυτοποίησης δεν είναι απαραίτητα για την επεξεργασία, οφείλει να διατηρεί την δυνατότητα ανάκτησης (recovery) των δεδομένων ταυτοποίησης και της ταυτότητας του υποκειμένου. Με αυτό τον τρόπο, αν αυτό ζητηθεί από το

υποκείμενο, θα είναι σε θέση να εντοπίσει τα δεδομένα που φυλάσσει για το συγκεκριμένο υποκείμενο και να του τα γνωστοποιήσει.

Είναι σημαντικό να γίνει αντιληπτή η διάκριση μεταξύ ψευδωνυμοποίησης και ανωνυμοποίησης. Η ανωνυμοποίηση είναι η προσπάθεια να εξασφαλιστεί ότι τα δεδομένα δεν είναι δυνατόν να συσχετιστούν με κάποιο ή κάποια συγκεκριμένα υποκείμενα ακόμα και αν ο επιτιθέμενος διαθέτει συμπληρωματικά δεδομένα από διαφορετικές πηγές, συνεπώς μετά από θεωρητικά επιτυχημένη ανωνυμοποίηση δεν τίθεται θέμα προσωπικών δεδομένων και ταυτοποίησης του υποκειμένου.

Η ψευδωνυμοποίηση πέρα από τον ορισμό της, συναντάται σε διάφορα άρθρα του ΓΚΠΔ:

- Η ψευδωνυμοποίηση αναφέρεται στην παρ. 1 του Άρθρου 25 του ΓΚΠΔ ενδεικτικά ως κατάλληλο τεχνικό μέτρο για την εφαρμογή αρχών προστασίας των δεδομένων ήδη από τον σχεδιασμό (by design) όπως η ελαχιστοποίηση των δεδομένων και για την εξασφάλιση εγγυήσεων όσον αφορά στην προστασία των δικαιωμάτων των υποκειμένων. Για την εκπλήρωση αυτών των απαιτήσεων θα πρέπει η ψευδωνυμοποίηση να καθορίζεται ήδη από τον σχεδιασμό της επεξεργασίας από τον Υπεύθυνο Επεξεργασίας («κατά την στιγμή του καθορισμού των μέσων επεξεργασίας»), ενώ θα πρέπει να διασφαλίζεται ότι οι καθορισμένες διαδικασίες ψευδωνυμοποίησης εφαρμόζονται πρακτικά κατά την διάρκεια της επεξεργασίας.
- Στην περίπτωση α) της παρ. 1 του Άρθρου 32 του ΓΚΠΔ η ψευδωνυμοποίηση αναφέρεται ως ένα από τα τεχνικά και οργανωτικά μέτρα που πρέπει λαμβάνονται, κατά περίπτωση, κατά την διάρκεια της επεξεργασίας για την διασφάλιση κατάλληλου επιπέδου ασφάλειας ενάντια στους κινδύνους που απειλούν τα δικαιώματα και τις ελευθερίες των υποκειμένων της επεξεργασίας.
- Στην παρ. 1 του Άρθρου 89 η χρήση ψευδωνύμων προτείνεται ενδεικτικά ως ένα τεχνικό και οργανωτικό μέτρο που μπορεί να περιλαμβάνεται στα μέτρα που λαμβάνονται για την διασφάλισης της αρχής της ελαχιστοποίησης των δεδομένων, κατά την επεξεργασία για σκοπούς αρχειοθέτησης προς το δημόσιο συμφέρον ή σκοπούς επιστημονικής ή ιστορικής έρευνας ή στατιστικούς σκοπούς. Έτσι τα ψευδωνυμοποιημένα δεδομένα μπορούν να υποβληθούν σε περεταίρω επεξεργασία για την εκπλήρωση αυτών των σκοπών, εφόσον δεν είναι δυνατή αλλά ούτε και απαραίτητη η ταυτοποίηση των υποκειμένων της επεξεργασίας.
- Στην παρ. 4 του Άρθρου 6 «όταν η επεξεργασία για σκοπό άλλο από αυτόν για τον οποίο έχουν συλλεγεί τα δεδομένα προσωπικού χαρακτήρα δεν βασίζεται στη συγκατάθεση του υποκειμένου των δεδομένων ή στο δίκαιο της Ένωσης ή το δίκαιο κράτους μέλους», ο Υπεύθυνος Επεξεργασίας θα πρέπει να εξακριβώσει κατά πόσο η επεξεργασία για άλλο σκοπό είναι συμβατή με τον σκοπό για τον οποίο συλλέχθηκαν αρχικώς τα δεδομένα

προσωπικού χαρακτήρα. Η ψευδωνυμοποίηση προτείνεται ενδεικτικά στην περίπτωση ε) ως μία από τις εγγυήσεις που μπορεί να λάβει υπόψη ο Υπεύθυνος Επεξεργασίας για την εξακρίβωση της συμβατότητας.

Για την επαρκή ψευδωνυμοποίηση, παρότι φαίνεται να είναι ευκολότερα εφικτή από την πλήρη ανωνυμοποίηση, δεν έχει προτυποποιηθεί κάποια συγκεκριμένη βέλτιστη τεχνική για όλες τις περιπτώσεις. Ανάλογα με την φύση των δεδομένων ταυτοποίησης και τον σκοπό της επεξεργασίας λαμβάνεται υπόψη και κριτήρια όπως η ανάγκη ελαχιστοποίησης των δεδομένων (data minimization) ώστε να μην γνωστοποιούνται δεδομένα ταυτοποίησης του υποκείμενου της επεξεργασίας σε οντότητες για τις οποίες δεν είναι αναγκαία η γνώση της ταυτότητας για να επιτελέσουν την λειτουργία τους κατά την διάρκεια της επεξεργασίας. Η παράγραφος 2 του Άρθρου 25 που ορίζει την εξ ορισμού (by default) προστασία των δεδομένων υποδεικνύει αυτά τα κριτήρια, τα οποία είναι απαραίτητο να λαμβάνονται υπόψη για την επιλογή τόσο συγκεκριμένης τεχνικής ψευδωνυμοποίησης, όσο και για την επιλογή του τόπου, του χρόνου και των υπόλοιπων λεπτομερειών της εφαρμογής της διαδικασίας ψευδωνυμοποίησης:

«Ο υπεύθυνος επεξεργασίας εφαρμόζει κατάλληλα τεχνικά και οργανωτικά μέτρα για να διασφαλίζει ότι, εξ ορισμού, υφίστανται επεξεργασία μόνο τα δεδομένα προσωπικού χαρακτήρα που είναι απαραίτητα για τον εκάστοτε σκοπό της επεξεργασίας. Αυτή η υποχρέωση ισχύει για το εύρος των δεδομένων προσωπικού χαρακτήρα που συλλέγονται, τον βαθμό της επεξεργασίας τους, την περίοδο αποθήκευσης και την προσβασιμότητά τους. Ειδικότερα, τα εν λόγω μέτρα διασφαλίζουν ότι, εξ ορισμού, τα δεδομένα προσωπικού χαρακτήρα δεν καθίστανται προσβάσιμα χωρίς την παρέμβαση του φυσικού προσώπου σε αόριστο αριθμό φυσικών προσώπων»

Ο ENISA στο [1] κωδικοποίησε τα παραπάνω τέσσερα κριτήρια, τα οποία θα σχολιάσουμε σε σχέση με την διαδικασία ψευδωνυμοποίησης, ως εξής:

- Κριτήριο 1 Ελαχιστοποίηση των προσωπικών δεδομένων: Η επιλογή μιας επαρκούς διαδικασίας ψευδωνυμοποίησης θα συνέβαλε στην ελαχιστοποίηση των προσωπικών δεδομένων, στον βαθμό που αποτρέπει την ταυτοποίηση του υποκειμένου των δεδομένων, καθιστώντας τα δεδομένα λιγότερο ευαίσθητα.
- Κριτήριο 2 Ελαχιστοποίηση της επεξεργασίας προσωπικών δεδομένων: Η ψευδωνυμοποίηση μέρους των δεδομένων μπορεί να επιτρέψει την πρόσβαση σε κάποιον εκτελώντα επεξεργασία μόνο στα δεδομένα που είναι απαραίτητα για να επιτελέσει τον ρόλο του κατά την διάρκεια της επεξεργασίας, περιορίζοντας έτσι έμμεσα την επεξεργασία που υπερβαίνει τους προκαθορισμένους σκοπούς της, την συγκατάθεση ή τα δικαιώματα του υποκειμένου.

- Κριτήριο 3 Ελαχιστοποίηση του χρόνου αποθήκευσης προσωπικών δεδομένων : Η επιλογή του χρόνου ψευδωνυμοποίησης μπορεί να περιορίσει έμμεσα τον χρόνο αποθήκευσης προσωπικών δεδομένων π.χ. όταν είναι απαραίτητη η ταυτοποίηση ενός χρήστη μόνο για την αρχική εγγραφή του σε μια συγκεκριμένη υπηρεσία, τα δεδομένα ταυτοποίησης μπορούν να ψευδωνυμοποιηθούν αμέσως μετά την εγγραφή του
- Κριτήριο 4 Ελάχιστη προσβασιμότητα στα προσωπικά δεδομένα : Η ψευδωνυμοποίηση των δεδομένων ταυτοποίησης είναι δυνατόν να χρησιμεύσει ως μέσο περιορισμού της πρόσβασης μόνο σε όποιον εμπλέκεται στην επεξεργασία και μόνο στα ελάχιστα δεδομένα που είναι απαραίτητα για να επιτελέσει τον ρόλο του σε αυτήν.

Μπορούμε να παρατηρήσουμε ότι η χρήση διαδικασιών ψευδωνυμοποίησης είναι δυνατόν να προστατεύσει τον υπεύθυνο επεξεργασίας, σε περίπτωση διαρροής προσωπικών δεδομένων:

- Αποτελεί μια από τις προτεινόμενες μεθόδους, συνεπώς μπορεί να χρησιμοποιηθεί ως αποδεικτικό στοιχείο της κανονιστικής συμμόρφωσης με τον ΓΚΠΔ περιορίζοντας τις ευθύνες που συνεπάγεται η επεξεργασία
- Λόγω της ελαχιστοποίησης των προσωπικών δεδομένων, μπορεί να ελαχιστοποιήσει τον αντίκτυπο τυχούσας διαρροής στα δικαιώματα των υποκειμένων.

Πέρα από την κανονιστική συμμόρφωση και τις νομικές επιπτώσεις, η μέριμνα για την προστασία των προσωπικών δεδομένων των πολιτών είναι προαπαιτούμενο που επιτρέπει την εγκαθίδρυση μιας σχέσης εμπιστοσύνης μεταξύ οποιουδήποτε δημόσιου ή ιδιωτικού φορέα που προσφέρει υπηρεσίες και των αποδεκτών των υπηρεσιών αυτών.

1.2 Ψευδωνυμοποίηση και νέες απαιτήσεις προστασίας των προσωπικών δεδομένων

Η εμφάνιση της πανδημίας Covid-19 έδειξε ότι σε καταστάσεις έκτακτης ανάγκης προκύπτουν νέοι σκοποί αλλά και νέες τεχνικές συλλογής και επεξεργασίας πολύ μεγάλου όγκου δεδομένων προσωπικού χαρακτήρα, όπως για παράδειγμα δεδομένων σχετικών με την επιδημιολογική επιτήρηση, την τηλεργασία, την ιχνηλάτηση επαφών, την τηλεκπαίδευση κ.α.

Για πολλούς από αυτούς τους σκοπούς και τις τεχνικές επεξεργασίας και για την εξασφάλιση της παροχής αποτελεσματικών υπηρεσιών, είναι απαραίτητη η συλλογή δεδομένων ταυτοποίησης των χρηστών, καθώς τα αποτελέσματα της επεξεργασίας μπορεί να δημιουργήσουν την ανάγκη επικοινωνίας με το υποκείμενο των δεδομένων, όπως για παράδειγμα η ειδοποίηση ενός χρήστη για κάποια επαφή του με κάποιον ασθενή. Προκύπτουν έτσι παράλληλα νέες ανάγκες για επαρκή

ψευδωνυμοποίηση για την προστασία της ιδιωτικότητας των υποκειμένων, ενώ συγχρόνως είναι απαραίτητη η διατήρηση της δυνατότητας ελεγχόμενης άρσης της ψευδωνυμοποίησης, όσο και η θέσπιση αυστηρά ελεγχόμενων διαδικασιών και μέτρων και για τις δύο περιπτώσεις.

Ο χαρακτήρας του επείγοντος που φέρουν οι καταστάσεις έκτακτης ανάγκης επιταχύνουν την ανάπτυξη υπηρεσιών για τις οποίες προηγουμένως είτε δεν είχε παρουσιαστεί η ανάγκη, είτε δεν ήταν ώριμη η τεχνολογία είτε δεν προϋπήρχαν οι υποδομές και η κατάλληλη οργάνωση για την εφαρμογή τους. Η άμεση εφαρμογή τέτοιων υπηρεσιών καταδεικνύει την εφικτότητα, την αποτελεσματικότητα αλλά και τις αδυναμίες τους, και συγχρόνως διευκολύνει την εφαρμογή τους σε αντίστοιχες καταστάσεις στο μέλλον.

Παρ ’ότι οι υπάρχουσες υποδομές και τεχνολογίες φαίνεται να επιτρέπουν εφαρμογές όπως για παράδειγμα η παρακολούθηση της κινητικότητας ή της κατάστασης της υγείας μεγάλου μέρους του πληθυσμού, η εθελοντική συμμετοχή των πολιτών φαίνεται να παίζει σημαντικό ρόλο. Σε πολλές περιπτώσεις είναι απαραίτητη η συλλογή και επεξεργασία μεγάλου όγκου προσωπικών δεδομένων σε συνδυασμό με την ταυτοποίηση μεγάλου αριθμού υποκειμένων, συνεπώς απαιτείται η ρητή συναίνεση καθενός από τα υποκείμενα.

Είναι προφανές ότι η κανονιστική συμμόρφωση με τον ΓΚΠΔ και την εφαρμοζόμενη νομοθεσία, πέρα από την αποφυγή νομικών κυρώσεων συμβάλλει στην δημιουργία σχέσης εμπιστοσύνης μεταξύ των υποκειμένων και των υπεύθυνων επεξεργασίας. Οι κανονιστικές ρυθμίσεις, δεν έχουν σκοπό να δυσχεράνουν την συλλογή και επεξεργασία των δεδομένων, αλλά αντιθέτως να την καταστήσουν δυνατή. Η παροχή διαβεβαιώσεων για την χρήση αποκλειστικά και μόνο των απαραίτητων προσωπικών δεδομένων αποκλειστικά και μόνο για τους δηλωμένους σκοπούς της επεξεργασίας είναι δυνατόν να επηρεάσει την συναίνεση των υποκειμένων αλλά και την ακρίβεια των δεδομένων, παράγοντας πιο αξιόπιστη πληροφορία για την αποτελεσματικότερη αντιμετώπιση κρίσιμων καταστάσεων.

Στο άμεσο μέλλον, οι πρόσφατες εμπειρίες είναι δυνατόν να οδηγήσουν στην χρήση παρόμοιων μεθόδων σε κρίσιμες καταστάσεις όπως είναι οι υγειονομικές κρίσεις, φυσικές καταστροφές, κυβερνοεπιθέσεις κ.α. Η ψευδωνυμοποίηση είναι ένα σημαντικό εργαλείο στην προσπάθεια διασφάλισης των δικαιωμάτων των υποκειμένων, όμως η επιλογή των συγκεκριμένων τεχνικών και πολιτικών εφαρμογής της εξαρτάται από τα χαρακτηριστικά και τα εμπλεκόμενα μέρη κάθε συγκεκριμένης επεξεργασίας δεδομένων. Οι πρόσφατες εμπειρίες μας δίνουν ρεαλιστικά σενάρια για την μελέτη των παραμέτρων, των κριτηρίων επιλογής αλλά και των αδυναμιών των διαφόρων υποθέσεων εφαρμογής διαδικασιών ψευδωνυμοποίησης.

2 Ψευδωνυμοποίηση

2.1 Σενάρια ψευδωνυμοποίησης

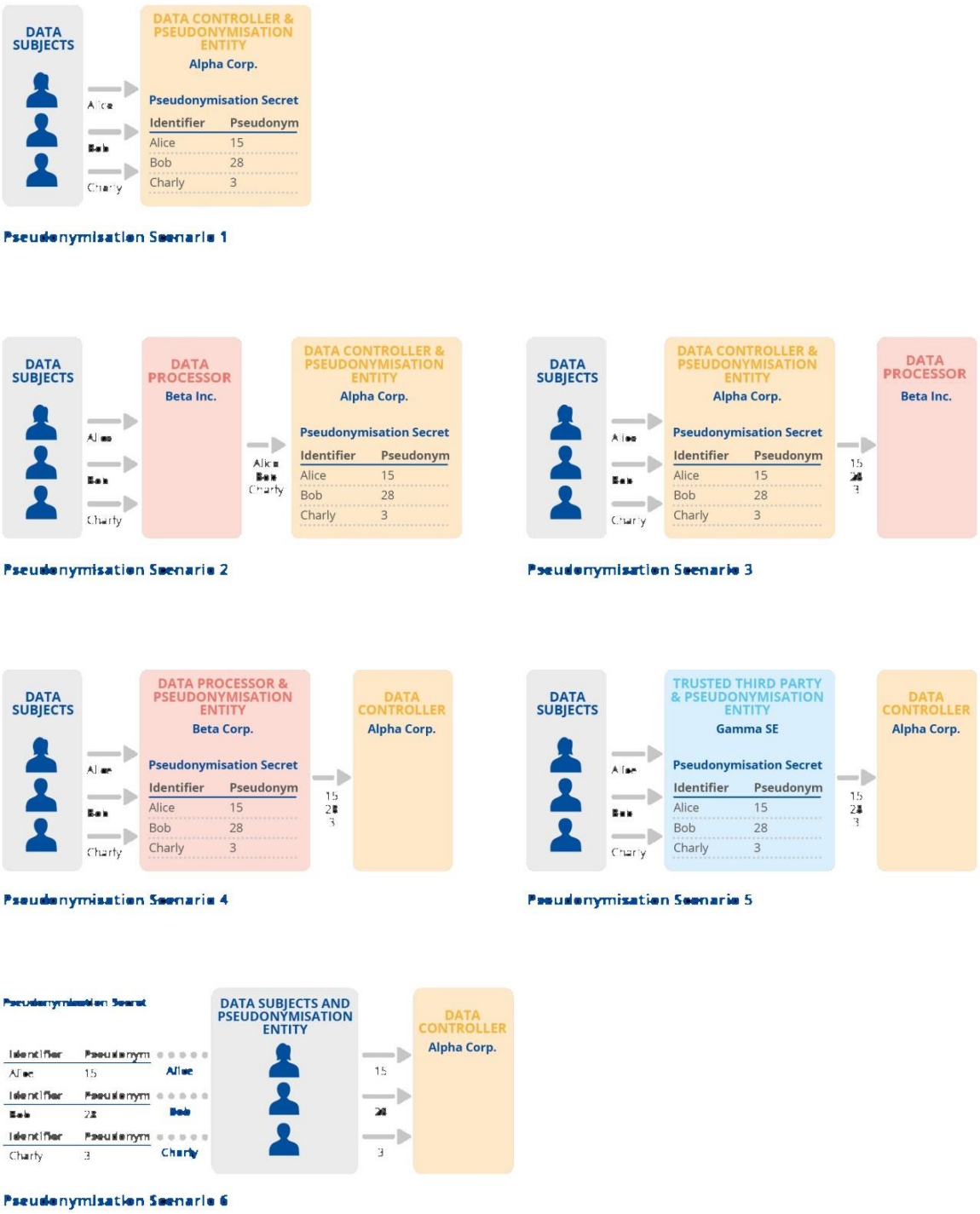
Ο ENISA στο [2] εξέτασε έξι ενδεικτικά σενάρια ψευδωνυμοποίησης, στα οποία τα εμπλεκόμενα στην επεξεργασία μέρη είναι:

- Τα Υποκείμενα των δεδομένων (Data Subjects)
- Ο Υπεύθυνος Επεξεργασίας (Data Controller)
- Ο Εκτελών την Επεξεργασία (Data Processor)
- Έμπιστη Τρίτη Οντότητα (Trusted Third Party)

Οποιοδήποτε από τα εμπλεκόμενα μέρη μπορεί να αναλαμβάνει τον ρόλο της Αρχής Ψευδωνυμοποίησης (Pseudonymisation Authority) που είναι το μέρος που εκτελεί την ψευδωνυμοποίηση.

Τα σενάρια ορίστηκαν ανάλογα με το ποιο εμπλεκόμενο μέρος διενεργεί την ψευδωνυμοποίηση και είναι τα εξής:

1. Ο Υπεύθυνος Επεξεργασίας συλλέγει τα δεδομένα και εκτελεί την ψευδωνυμοποίηση
2. Ο Υπεύθυνος Επεξεργασίας εκτελεί την ψευδωνυμοποίηση, αφού ο Εκτελών την επεξεργασία έχει συλλέξει ή και επεξεργαστεί τα δεδομένα
3. Ο Υπεύθυνος Επεξεργασίας εκτελεί την ψευδωνυμοποίηση και κατόπιν προωθεί τα ψευδωνυμοποιημένα δεδομένα στον Εκτελώντα την Επεξεργασία
4. Ο Εκτελών την Επεξεργασία εκτελεί την ψευδωνυμοποίηση και προωθεί τα ψευδωνυμοποιημένα δεδομένα στον Υπεύθυνο επεξεργασίας
5. Μια Έμπιστη Τρίτη Οντότητα συλλέγει τα δεδομένα και εκτελεί την ψευδωνυμοποίηση. Κατόπιν προωθεί τα δεδομένα στον Υπεύθυνο Επεξεργασίας.
6. Τα ίδια τα Υποκείμενα των δεδομένων αναλαμβάνουν τον ρόλο της Αρχής Ψευδωνυμοποίησης, εκτελούν την ψευδωνυμοποίηση και προωθούν τα δεδομένα στον Υπεύθυνο επεξεργασίας



Εικόνα 1 Basic pseudonymisation scenarios
Πηγή: DATA PSEUDONYMISATION: ADVANCED TECHNIQUES & USE CASES (ENISA 2021)

Η παραπάνω κατηγοριοποίηση είναι υποκειμενική και θα μπορούσε να επεκταθεί με περισσότερα σενάρια ή να συντμηθεί σε λιγότερα, είναι όμως επαρκής και πλήρης για το σκοπό της μελέτης συγκεκριμένων περιπτώσεων ψευδωνυμοποίησης

2.2 Απαιτούμενες ιδιότητες επιλογής τεχνικών ψευδωνυμοποίησης

Κατά την επιλογή μεθόδων εφαρμογής ψευδωνυμοποίησης που μπορεί να περιλαμβάνουν συγκεκριμένες τεχνικές, πολιτικές και οργανωτικά μέτρα, λαμβάνεται υπόψη ο σκοπός της επεξεργασίας των προσωπικών δεδομένων καθώς και το εσωτερικό και εξωτερικό περιβάλλον. Από την ανάλυση αυτή, συχνά προκύπτουν απαιτούμενες ιδιότητες των μεθόδων που θα επιλεγούν, μερικές από τις οποίες μπορεί να περιλαμβάνουν τις:

1. Προστασία Δεδομένων (Data Protection): Η ιδιότητα αυτή είναι πάντα απαιτούμενη, καθώς η διαδικασία της ψευδωνυμοποίησης θα πρέπει πάντα να επιτυγχάνει στον μέγιστο δυνατό βαθμό την μη αποκάλυψη της ταυτότητας των υποκειμένων σε μη εξουσιοδοτημένες οντότητες. Όπως αναφέρθηκε, σε καταστάσεις έκτακτης ανάγκης η αποδείξιμη διαβεβαίωση αυτής της ιδιότητας μπορεί να διευρύνει την συμμετοχή των υποκειμένων και να εξασφαλίσει την επαρκή συλλογή πληροφοριών
2. Χρησιμότητα (Utility): Πολλές φορές προκύπτει η ανάγκη μέρος των αναγνωριστικών να διατηρήσει μέρος ή όλη την χρησιμότητά του μετά την ψευδωνυμοποίηση για τις ανάγκες της επεξεργασίας. Χαρακτηριστικό παράδειγμα είναι η ψευδωνυμοποίηση της IP διεύθυνσης από την οποία συνδέεται ένα υποκείμενο σε μία υπηρεσία. Είναι δυνατόν να μην ψευδωνυμοποιηθεί το μέρος της IP διεύθυνσης που χαρακτηρίζει το υποδίκτυο, εάν υπάρχει η απαίτηση να είναι αυτό γνωστό για λόγους επεξεργασίας, ή να ψευδωνυμοποιηθεί με μια ντετερμινιστική πολιτική αν απλώς κατά την επεξεργασία υπάρχει ανάγκη να κατηγοριοποιηθούν τα υποκείμενα ανάλογα με το υποδίκτυο χωρίς όμως να είναι γνωστό ποιο ψευδώνυμο αντιστοιχεί σε συγκεκριμένο υποδίκτυο. Η απαίτηση της χρησιμότητας, στις περισσότερες περιπτώσεις έχει αρνητικές επιπτώσεις στην προστασία των δεδομένων, καθώς ένας επιτιθέμενος μπορεί να εκμεταλλευτεί την διαθέσιμη χρηστική πληροφορία για να βοηθηθεί στην μερική ταυτοποίηση του υποκειμένου. Στο παραπάνω παράδειγμα, ο επιτιθέμενος μπορεί να έχει πρόσβαση στην πληροφορία ότι το κάθε υποκείμενο ανήκει σε συγκεκριμένο υποδίκτυο, ή στην πληροφορία ότι ένα σύνολο υποκειμένων ανήκουν στο ίδιο υποδίκτυο μεταξύ τους.
3. Επεκτασιμότητα (Scalability): σε πολλές περιπτώσεις απαιτείται η επιλεγμένη μέθοδος ψευδωνυμοποίησης να εφαρμοστεί σε πολύ μεγάλο ή και απροσδιόριστα μεγάλο πλήθος δεδομένων, ενώ παράλληλα μπορεί να είναι σημαντικές παράμετροι της εφαρμογής η εξοικονόμηση αποθηκευτικού χώρου, η ταχύτητα υπολογισμού των ψευδωνύμων και η ταχύτητα ανάκλησης των αναγνωριστικών από τα ψευδώνυμα. Η δυνατότητα να εφαρμοστεί

μια μέθοδος ψευδωνυμοποίησης σε μεγάλο πλήθος δεδομένων ορίζει την ιδιότητα της επεκτασιμότητας, η οποία μπορεί να έχει ιδιαίτερη σημασία σε καταστάσεις έκτακτης ανάγκης, καθώς αυτές είναι δυνατόν να αφορούν σε μεγάλους πληθυσμούς υποκειμένων, από τα οποία ή προς τα οποία θα πρέπει να αντληθούν ή να προωθηθούν αντίστοιχα πληροφορίες

4. Ανάκτηση (Recovery): είναι η ανάκτηση του αναγνωριστικού από το ψευδώνυμο. Σε κάποιες εφαρμογές είναι δυνατόν να μην απαιτείται ανάκτηση, όμως στις περισσότερες περιπτώσεις η δυνατότητα αυτή είναι απαραίτητη, ενώ σε πολλές περιπτώσεις είναι επωφελές να είναι ταχεία και να απαιτεί μικρή υπολογιστική ισχύ

2.3 Πολιτικές ψευδωνυμοποίησης

Οι πολιτικές ψευδωνυμοποίησης που επιλέγονται επηρεάζουν καθοριστικά τις ιδιότητες της μεθόδου ως προς τις παραπάνω απαιτήσεις, αλλά και την επιλογή της συγκεκριμένης τεχνικής ψευδονυμοποίησης που θα χρησιμοποιηθεί. Θα σχολιάσουμε τις πολιτικές ψευδωνυμοποίησης όπως περιγράφονται στο [2]

Θεωρούμε μια βάση δεδομένων ή άλλο έγγραφο που περιλαμβάνει k αναγνωριστικά. Ένα αναγνωριστικό Id εμφανίζεται πολλές φορές σε δύο σύνολα δεδομένων (dataset) A και B. Μετά την εφαρμογή μιας τεχνικής ψευδωνυμοποίησης, το αναγνωριστικό Id αντικαθίσταται από ένα ψευδώνυμο pseudo. Η αντικατάσταση αυτή μπορεί να γίνει με τρείς διαφορετικές πολιτικές ψευδωνυμοποίησης

2.3.1 Ντετερμινιστική ψευδωνυμοποίηση (Deterministic Pseudonymisation)

Από την βάση δεδομένων εξάγουμε όλα τα αναγνωριστικά. Κάθε μοναδικό αναγνωριστικό Id μέσω μιας τεχνικής ψευδωνυμοποίησης αντιστοιχίζεται σε ένα ψευδώνυμο pseudo, δημιουργώντας έτσι έναν πίνακα αντιστοίχισης (mapping table), που περιλαμβάνει το κάθε αναγνωριστικό και το αντίστοιχο ψευδώνυμό του. Ακολούθως, το Id αντικαθίσταται σε όλες τις εμφανίσεις του στην βάση δεδομένων από το pseudo. Το χαρακτηριστικό αυτής της πολιτικής, είναι ότι το αναγνωριστικό αντικαθίσταται σε όλες τις εμφανίσεις του σε όλες τις βάσεις δεδομένων από το ίδιο πάντα ψευδώνυμο.

Η ντετερμινιστική πολιτική είναι πιο απλή στην εφαρμογή της, όμως κάτω από συγκεκριμένες συνθήκες θα μπορούσε να κάνει περισσότερο ρεαλιστικές κάποιες μορφές επιθέσεων στην προστασία των δεδομένων (data protection). Για παράδειγμα, στην περίπτωση που κάποιος

επιτιθέμενος διαθέτει κάποιες πληροφορίες σχετικά με την συχνότητα εμφάνισης ενός αναγνωριστικού σε ένα σύνολο βάσεων δεδομένων θα μπορούσε να εξάγει συμπεράσματα μελετώντας τις συχνότητες εμφάνισης όλων των ψευδωνύμων σε αυτές.

Παρομοίως, η ίδια αδυναμία θα μπορούσε θα ήταν δυνατόν να είναι ταυτόχρονα μια επιθυμητή ιδιότητα που αυξάνει την χρησιμότητα (utility), η οποία θα επέτρεπε σε κάποια, εξουσιοδοτημένη για επεξεργασία των δεδομένων, οντότητα να κάνει συγκρίσεις μεταξύ βάσεων δεδομένων. Για παράδειγμα, σε κάποια εφαρμογή επιδημιολογικής επιτήρησης, θα ήταν δυνατόν μια εξουσιοδοτημένη οντότητα να γνωρίζει ότι κάποιο υποκείμενο δήλωσε ένα θετικό τεστ σε κάποιον ιό και ότι ακολούθως σε σύντομο χρονικό διάστημα επιβεβαίωσε την θετικότητα δηλώνοντας και δεύτερο θετικό τεστ επιβεβαίωσης, χωρίς η εξουσιοδοτημένη οντότητα να γνωρίζει το αναγνωριστικό και χωρίς να έχει πρόσβαση στην ταυτότητα του υποκειμένου.

2.3.2 Ψευδωνυμοποίηση τυχαιοποιημένη ανά έγγραφο (Document-randomized pseudonymization)

Σε αυτή την πολιτική, κάθε εμφάνιση ενός συγκεκριμένου αναγνωριστικού Id σε μια βάση δεδομένων ή ένα έγγραφο, αντικαθίσταται από ένα διαφορετικό ψευδώνυμο pseudo1, pseudo2, pseudo3 ….. κ.ο.κ. Όμως το αναγνωριστικό Id αντιστοιχίζεται πάντα σε όλες τις βάσεις δεδομένων και τα έγγραφα στο ίδιο σύνολο ψευδωνύμων (pseudo1, pseudo2, pseudo3 …..), δηλαδή σε οποιαδήποτε άλλη βάση ή έγγραφο, το Id θα αντικαθίσταται πάλι από τα (pseudo1, pseudo2, pseudo3 …..) .

Η πολιτική τυχαιοποιημένης ανά έγγραφο ψευδωνυμοποίησης προσφέρει καλύτερη προστασία δεδομένων, καθώς είναι πολύ πιο περίπλοκο για κάποιον επιτιθέμενο να ανακτήσει το αναγνωριστικό Id από ένα από τα ψευδώνυμα με τα οποία έχει αντικατασταθεί, ειδικά στην περίπτωση που ο επιτιθέμενος έχει πρόσβαση μόνο σε μία βάση δεδομένων ή ένα έγγραφο, καθώς δεν είναι εύκολο να εξάγει οποιοδήποτε συμπέρασμα ακόμα και αν γνωρίζει ή μαντεύει τις συχνότητες εμφάνισης των αναγνωριστικών.

Επίσης είναι δυνατόν να διατηρήσει σε κάποιο βαθμό την χρησιμότητα της ντετερμινιστικής πολιτικής όπως παρουσιάστηκε στην προηγούμενη παράγραφο. Μια εξουσιοδοτημένη οντότητα, είναι δυνατόν να προχωρήσει σε επεξεργασία που περιλαμβάνει τέτοιες συγκρίσεις μεταξύ βάσεων δεδομένων, αν της δοθεί το σύνολο ψευδωνύμων (pseudo1, pseudo2…..) για το αναγνωριστικό Id, χωρίς να της δοθεί το ίδιο το αναγνωριστικό, οπότε δεν έχει πρόσβαση στην ταυτότητα του υποκειμένου.

Αναλόγως και της τεχνικής ψευδωνυμοποίησης που θα χρησιμοποιηθεί, η πολιτική αυτή μπορεί να επηρεάσει αρνητικά την επεκτασιμότητα (scalability) σε σχέση με την ντετερμινιστική, καθώς για παράδειγμα είναι δυνατόν να απαιτήσει μεγαλύτερο πίνακα συσχέτισης (mapping table). Ένα αναγνωριστικό Id απαιτεί μόνο μία εγγραφή σε πίνακα συσχέτισης ντετερμινιστικής πολιτικής, αλλά απαιτεί αριθμό εγγραφών ίσο με τον αριθμό των εμφανίσεών του στο έγγραφο για την πολιτική ψευδωνυμοποίησης τυχαιοποιημένης ανά έγγραφο. Συνεπώς αν το πλήθος δεδομένων είναι μεγάλο και κατά την διάρκεια της ψευδωνυμοποίησης ή της ανάκτησης απαιτείται αναζήτηση στον πίνακα αντιστοίχισης (για παράδειγμα αν πρέπει να βεβαιωθούμε ότι δεν θα αντιστοιχίσουμε το ίδιο ψευδώνυμο σε δύο αναγνωριστικά) τότε το μέγεθος του πίνακα συσχέτισης μπορεί να θέσει περιορισμούς στην ταχύτητα αυτών των εργασιών.

2.3.3 Πλήρως τυχαιοποιημένη ψευδωνυμοποίηση (Fully-randomized pseudonymisation)

Σε αυτή την πολιτική το αναγνωριστικό Id σε κάθε εμφάνισή του σε οποιαδήποτε βάση αντιστοιχίζεται πάντα από ένα διαφορετικό ψευδώνυμο. Σε αυτή την περίπτωση, η προστασία των δεδομένων φαίνεται να είναι ισχυρότερη, καθώς ο επιτιθέμενος δεν είναι δυνατόν να αναγνωρίσει την συχνότητα εμφάνισης ούτε την θέση εμφάνισης οποιουδήποτε αναγνωριστικού. Κανένα ψευδώνυμο δεν εμφανίζεται ποτέ δύο φορές, αφού ποτέ ένα αναγνωριστικό δεν αντικαθίσταται σε δύο εμφανίσεις από το ίδιο ψευδώνυμο.

Είναι όμως αναλόγως δυσκολότερο να διατηρηθεί οποιαδήποτε χρησιμότητα όπως αυτή στα παραδείγματα των δύο προηγούμενων πολιτικών, καθώς όπως είναι προφανές κάποια οντότητα που δεν έχει δυνατότητα ανάκτησης ενός αναγνωριστικού δεν μπορεί να πραγματοποιήσει επεξεργασία που απαιτεί συγκρίσεις μεταξύ βάσεων δεδομένων ή άλλων εγγράφων. Επίσης, η αντικατάσταση κάθε εμφάνισης ενός αναγνωριστικού με ένα διαφορετικό για κάθε εμφάνιση ψευδώνυμο όπως είναι επόμενο θα απαιτούσε έναν μεγαλύτερο και πιο δύσχρηστο πίνακα αντιστοίχισης, πράγμα που όπως είδαμε στην προηγούμενη παράγραφο σε κάποιες περιπτώσεις μπορεί να επηρεάζει την επεκτασιμότητα του συστήματος για μεγάλο πλήθος δεδομένων, δημιουργώντας μεγαλύτερες ανάγκες σε επεξεργαστική ισχύ και αποθηκευτικό χώρο.

Σε πολλές εφαρμογές όπως αυτές που θα μελετήσουμε μπορεί να απαιτείται η επιλογή μιας ισορροπημένης πολιτικής που να επιτρέπει την ζητούμενη χρησιμότητα, χωρίς να τίθεται σε κίνδυνο η προστασία των δεδομένων και μέσα στα όρια που θέτουν οι τεχνικοί περιορισμοί και οι απαιτήσεις που αφορούν στην ζητούμενη ταχύτητα επεξεργασίας

2.4 Τεχνικές Ψευδωνυμοποίησης

Μια συνάρτηση ψευδωνυμοποίησης αντιστοιχίζει ένα αναγνωριστικό Id σε ένα ψευδώνυμο pseudo, το οποίο θα πρέπει αν είναι μοναδικό, δηλαδή για δύο αναγνωριστικά Id1 και Id2 τα αντίστοιχα ψευδώνυμα pseudo1 και pseudo2 πρέπει να είναι διαφορετικό το ένα από το άλλο, έτσι ώστε να είναι δυνατή η ανάκτηση από το αναγνωριστικό. Αντίθετα ένα αναγνωριστικό Id είναι δυνατόν να αντιστοιχίζεται σε περισσότερα του ενός ψευδώνυμα pseudo1, pseudo2 …. Κλπ. όπως είδαμε και στην περιγραφή των μη ντετερμινιστικών πολιτικών. Η συνάρτηση ψευδωνυμοποίησης θα πρέπει να κάνει χρήση ενός *μυστικού ψευδωνυμοποίησης* (pseudonymisation secret) έτσι ώστε να μην είναι δυνατή η ανάκτηση των αναγνωριστικών χωρίς την γνώση του μυστικού ψευδωνυμοποίησης.

2.4.1 Βασικές τεχνικές ψευδωνυμοποίησης

Οι βασικές τεχνικές ψευδωνυμοποίησης όπως περιγράφονται στο [2] περιλαμβάνουν τις εξής:

- *Μετρητής (Counter)*: Απλή τεχνική όπου σαν ψευδώνυμο αποδίδεται απλά ένας σειριακός μετρητής που αυξάνεται με κάθε νέα είσοδο στην συνάρτηση ψευδωνυμοποίησης. Σε αυτή την περίπτωση το μυστικό ψευδωνυμοποίησης είναι ο ίδιος ο πίνακας αντιστοίχισης, ο οποίος θα πρέπει να αποθηκεύεται ολόκληρος. Αυτό θα μπορούσε να επηρεάσει αρνητικά την επεκτασιμότητα σε μεγάλες και πολύπλοκες βάσεις δεδομένων όπου πρέπει να ψευδωνυμοποιηθούν περισσότερα του ενός αναγνωριστικά, δημιουργώντας μεγαλύτερες απαιτήσεις σε αποθηκευτικό χώρο και ταχύτητα επεξεργασίας.
  Δεν υπάρχει καμία σύνδεση των ψευδωνύμων με τα αναγνωριστικά, όμως αν κάποιος επιτιθέμενος διαθέτει πληροφορίες για τα αναγνωριστικά και την σειρά με την οποία αυτά ψευδωνυμοποιήθηκαν, θα ήταν δυνατόν να διευκολυνθεί στην ταυτοποίηση
- *Γεννήτρια τυχαίων Αριθμών*: Σε αυτή την περίπτωση ως ψευδώνυμο αποδίδεται ένας αριθμός που παράγεται από μια συνάρτηση τυχαία, δηλαδή υπάρχει ομοιόμορφη πιθανότητα για όλες τις τιμές του πεδίου τιμών να παραχθούν. Η κατηγορία αυτή περιλαμβάνει και την χρήση γεννήτριας ψευδοτυχαίων αριθμών, όμως σε όλες τις περιπτώσεις είναι δυνατή σπανίως η ύπαρξη *συγκρούσεων (collision)*, δηλαδή η παραγωγή του ίδιου ψευδώνυμου για δεύτερη φορά. Αυτό δημιουργεί την ανάγκη αποθήκευσης του πίνακα αντιστοίχισης και αναζήτησης σε αυτόν κάθε φορά που παράγεται ένα νέο ψευδώνυμο για να βεβαιωθούμε ότι δεν έχει αποδοθεί σε άλλο αναγνωριστικό, συνεπώς έχει επιπτώσεις στην επεκτασιμότητα σε μεγάλες βάσεις δεδομένων

- *Συνάρτηση Κατακερματισμού (Hash function)*: Οι συναρτήσεις κατακερματισμού είναι εξ ’ορισμού κατασκευασμένες έτσι ώστε να είναι μονόδρομες (one-way), συνεπώς είναι εξαιρετικά δύσκολο να υπολογιστεί το αναγνωριστικό (είσοδος της hash function) αν κάποιος γνωρίζει μόνο το ψευδώνυμο που παράχθηκε από αυτό (έξοδος της hash function). Επίσης εξ’ ορισμού είναι ζητούμενο η αποφυγή συγκρούσεων (collision). Παρ’ όλα αυτά, δεν εξασφαλίζουν μεγάλη προστασία δεδομένων λόγω της *επίθεσης λεξικού (dictionary attack)*: αν ένας επιτιθέμενος έχει πρόσβαση ή μπορεί να μαντέψει το πεδίο τιμών ενός αναγνωριστικού, μπορεί να υπολογίσει τα ψευδώνυμα (hash) όλων και να τα συγκρίνει με τα διαθέσιμα ψευδώνυμα
- *Κωδικός Αυθεντικοποίησης Μηνύματος (Message Authentication Code* – MAC): Εστιάζουμε στην πλέον συνηθισμένη περίπτωση αυτής της κατηγορίας τεχνικών, που είναι η χρήση συνάρτησης κατακερματισμού (hash function) με την χρήση κρυπτογραφικού κλειδιού. Η τεχνική αυτή ονομάζεται keyed-hash ή HMAC. Η χρήση κρυπτογραφικού κλειδιού προσδίδει μεγάλη ανθεκτικότητα σε επιθέσεις λεξικού, εφόσον ο επιτιθέμενος δεν έχει πρόσβαση στο κλειδί και δεν μπορεί να υπολογίσει τα ψευδώνυμα. Οι τεχνικές HMAC προσφέρουν σημαντική προστασία των δεδομένων, ενώ υπάρχει μεγάλη εμπειρία σχετικά με την ασφάλεια και την απόδοσή τους από την χρήση τους για διαφορετικούς σκοπούς σε τεχνολογίες διαδικτύου.
- *Συμμετρική Κρυπτογράφηση (Symmetric Encryption)*: Στην συμμετρική κρυπτογράφηση χρησιμοποιείται ένα κρυπτογραφικό κλειδί, το ίδιο για την κρυπτογράφηση και την αποκρυπτογράφηση εξ ’αρχής ορισμένου μεγέθους τμημάτων (block), δηλαδή σειρών δυαδικών ψηφίων. Η είσοδος του κρυπτογραφικού αλγόριθμου είναι το αναγνωριστικό, το οποίο μπορεί να έχει μέγεθος μικρότερο ή μεγαλύτερο από τον αριθμό ψηφίων που έχει οριστεί σαν μέγεθος του block. Πολλοί συμμετρικοί κρυπτογραφικοί αλγόριθμοι χρησιμοποιούνται σε διάφορες εφαρμογές, και να προβλήματα μεγέθους της εισόδου έχουν ήδη αντιμετωπιστεί, είτε με padding, δηλαδή την με προκαθορισμένο τρόπο συμπλήρωση ψηφίων στα ψηφία του αναγνωριστικού ώστε να φθάσει στο ορισμένο μέγεθος του block όταν είναι μικρότερο, είτε με την χρήση τεχνικών όπως συμπίεση (compression), λειτουργία μετρητή (counter mode), κρυπτογράφηση αλυσίδας μπλοκ (Cipher Block Chaining - CBC). Η απόδοση των αλγορίθμων συμμετρικής κρυπτογράφησης είναι επαρκής σε ρεαλιστικά σενάρια τόσο παραγωγής ψευδωνύμων όσο και ανάκτησης αναγνωριστικών, όπως και η προστασία των δεδομένων.

2.4.2 Προχωρημένες τεχνικές ψευδωνυμοποίησης

Οι προχωρημένες τεχνικές ψευδωνυμοποίησης όπως περιγράφονται στο [3] περιλαμβάνουν τις εξής:

- *Ασύμμετρη κρυπτογραφία:* Ένας χρήστης ασύμμετρου κρυπτογραφικού αλγορίθμου διαθέτει δημόσια το δικό του Δημόσιο Κλειδί (Public Key -PK), και κρατά μυστικό το Μυστικό Κλειδί (Secret Key – SK). Δεδομένα που κρυπτογραφούνται με το δημόσιο κλειδί, μπορούν αποκρυπτογραφηθούν με το μυστικό κλειδί. Συνεπώς οποιοσδήποτε μπορεί να κρυπτογραφήσει δεδομένα με το δημόσιο κλειδί του χρήστη, αλλά μόνο ο ίδιος ο χρήστης μπορεί να τα αποκρυπτογραφήσει. Υπάρχει μεγάλη εμπειρία από την χρήση ασύμμετρων κρυπτογραφικών αλγορίθμων τόσο για κρυπτογράφηση όσο και για την δημιουργία ψηφιακών υπογραφών σε εφαρμογές.

  Η χρήση ασύμμετρης κρυπτογραφίας για λόγους ψευδωνυμοποίησης θα μπορούσε να επιτρέψει την ανάθεση της ψευδωνυμοποίησης. Για παράδειγμα, ένας υπεύθυνος επεξεργασίας, αφού παράγει το ζεύγος κρυπτογραφικών κλειδιών, γνωστοποιεί το δημόσιο κλειδί του στην οντότητα ψευδωνυμοποίησης , η οποία όπως είδαμε στην παράγραφο 2.1 μπορεί να είναι οποιοδήποτε από τα εμπλεκόμενα μέρη έχει αναλάβει την πραγματοποίηση της διαδικασίας ψευδωνυμοποίησης. Η οντότητα ψευδωνυμοποίησης μπορεί να κρυπτογραφήσει τα αναγνωριστικά με το δημόσιο κλειδί του υπεύθυνου επεξεργασίας παράγοντας έτσι τα ψευδώνυμα και να παραδώσει τα ψευδωνυμοποιημένα δεδομένα στον υπεύθυνο επεξεργασίας. Έτσι, αργότερα μόνο ο υπεύθυνος επεξεργασίας θα είναι σε θέση να αποκρυπτογραφήσει τα ψευδώνυμα χρησιμοποιώντας το μυστικό κλειδί του και να ανακτήσει τα αναγνωριστικά. Σε αυτή την περίπτωση κατάλληλα διοικητικά, οργανωτικά και τεχνολογικά μέτρα θα πρέπει να λαμβάνονται ώστε η οντότητα ψευδωνυμοποίησης να μην αποθηκεύει τα αναγνωριστικά που έχει συλλέξει, εάν αυτό δεν απαιτείται ή δεν είναι επιθυμητό.

  Η παραπάνω μέθοδος στην γενική περίπτωση δεν προσφέρει πάντα επαρκή προστασία των δεδομένων αν δεν εφαρμοστούν μέτρα που ενισχύουν την ασφάλεια. Υποθέτουμε ότι ένας επιτιθέμενος θέλει να αναζητήσει σε μια βάση δεδομένων τα στοιχεία που αφορούν σε συγκεκριμένα υποκείμενα το οποίο συνδέονται με συγκεκριμένα αναγνωριστικά όπως για παράδειγμα το επώνυμο. Αν το δημόσιο κλειδί του υπεύθυνου επεξεργασίας έχει δημοσιοποιηθεί, μπορεί να το χρησιμοποιήσει και να κρυπτογραφήσει τα αναγνωριστικά που αναζητά, παράγοντας τα ψευδώνυμά τους. Ακολούθως, αναζητώντας στην βάση τα ψευδώνυμα που βρήκε μπορεί τα ταυτοποιήσει κάποια υποκείμενα και να εντοπίσει τα

δεδομένα που συνδέονται με αυτά. Με τέτοιες προσεγγίσεις, είναι δυνατόν να επιχειρήσει κάποιο είδος εξαντλητικής αναζήτησης (exhaustive search) ή επίθεσης λεξικού. Για παράδειγμα αν γνωρίζει το πεδίο τιμών του αναγνωριστικού μπορεί να συντάξει ένα «λεξικό» που θα περιλαμβάνει όλα τα πιθανά αναγνωριστικά (όπως για παράδειγμα όλα τα πιθανά επώνυμα ή όλους τους πιθανούς αριθμούς κοινωνικής ασφάλισης), να τα κρυπτογραφήσει υπολογίζοντας έτσι όλα τα αντίστοιχα ψευδώνυμα και τέλος να συγκρίνει τα ψευδώνυμα αυτά με αυτά που περιλαμβάνονται στην βάση δεδομένων, ταυτοποιώντας έτσι μεγάλο αριθμό υποκειμένων.

Χρήσιμα μέτρα ενίσχυσης της προστασίας δεδομένων έναντι επιθέσεων εξαντλητικής αναζήτησης, είναι:

Η διάθεση του δημόσιου κλειδιού μόνο στην οντότητα ψευδωνυμοποίησης (ή στις οντότητες ψευδωνυμοποίησης). Στην γενική χρήση των ασύμμετρων κρυπτογραφικών αλγορίθμων η δυνατότητα δημοσιοποίησης του δημόσιου κλειδιού είναι ένα σημαντικό πλεονέκτημα στην διαχείριση των κλειδιών, όμως στην χρήση για λόγους ψευδωνυμοποίησης σε κάποιες περιπτώσεις δεν είναι απαραίτητη, υποθέτοντας ότι οι οντότητες ψευδωνυμοποίησης είναι λίγες σε αριθμό και είναι δυνατή η αποστολή του δημόσιου κλειδιού μέσω ασφαλούς καναλιού. Όμως, η προϋπόθεση αυτή δεν ισχύει όταν οι οντότητες ψευδωνυμοποίησης είναι τα ίδια τα υποκείμενα των δεδομένων, όπως συμβαίνει στο σενάριο ψευδωνυμοποίησης 6 της παρ. 2.1 και σε υποθέσεις εργασίας που θα μελετήσουμε. Παρ' όλα αυτά, το μέτρο αυτό δεν είναι ούτε επαρκές αλλά ούτε και απαραίτητο για την ενίσχυση της προστασίας δεδομένων

Η χρήση ασύμμετρης κρυπτογράφησης στο πλαίσιο μιας πολιτικής τυχαιοποιημένης ψευδωνυμοποίησης γίνεται με την χρήση μιας τυχαίας (ή ψευδοτυχαίας) παραμέτρου nonce (number used once), ενός αριθμού που είναι διαφορετικός για κάθε αναγνωριστικό που κρυπτογραφείται. Αυτό μπορεί να γίνει με την χρήση ενός μη-ντετερμινιστικού ασύμμετρου αλγορίθμου όπως είναι ο αλγόριθμος El Gamal, ή με την χρήση nonce ως συμπλήρωμα (padding) του αναγνωριστικού σε έναν ντετερμινιστικό αλγόριθμο όπως ο RSA. Η χρήση nonce ως είσοδο στον κρυπτογραφικό αλγόριθμο θα αλλάξει καθοριστικά την έξοδο (κρυπτόγραμμα) δηλαδή το παραγόμενο ψευδώνυμο, λόγω της ιδιότητας της *πληρότητας (completeness)* που απαιτείται στον σχεδιασμό κρυπτογραφικών αλγορίθμων. Καθώς ο επιτιθέμενος είναι αδύνατο να υπολογίσει μαζικά όλα τα κρυπτογράμματα για όλα τα πιθανά nonce, παρέχεται επαρκής προστασία από επιθέσεις εξαντλητικής αναζήτησης.

Διάφοροι τρόποι χρήσης της ασύμμετρης κρυπτογραφίας για σκοπούς ψευδωνυμοποίησης έχουν προταθεί, όπως είναι η χρήση διαφορετικών τυχαιοποιημένων ψευδωνύμων του ίδιου αναγνωριστικού σε διαφορετικά πεδία, η κατανεμημένη ψευδωνυμοποίηση χωρίς γνώση του αναγνωριστικού και χρήση συνδέσιμων ψευδωνύμων συναλλαγών (linkable transaction

pseudonyms)

- *Υπογραφές Δακτυλίου (Ring signatures) και ομαδικά ψευδώνυμα (group pseudonyms):* Η ασύμμετρη κρυπτογραφία δίνει την δυνατότητα της αυθεντικοποίησης ενός χρήστη, καθώς αυτός μπορεί να κρυπτογραφήσει ένα κείμενο (plaintext) με το προσωπικό του μυστικό κλειδί SK. Κάποιος που θα λάβει το κείμενο και το κρυπτόγραμμα μπορεί να αποκρυπτογραφήσει το κρυπτόγραμμα με το δημόσιο κλειδί PK του χρήστη. Αν η αποκρυπτογράφηση δώσει σαν αποτέλεσμα το αρχικό κείμενο, μπορεί να είναι σίγουρος ότι κρυπτογραφήθηκε από τον συγκεκριμένο χρήστη με το προσωπικό του μυστικό κλειδί. Συνεπώς το δημόσιο κλειδί ενός χρήστη δίνει την δυνατότητα ταυτοποίησης του. Η ιδιότητα αυτή είναι πολύ χρήσιμη για την δημιουργία προσωπικών ψηφιακών υπογραφών, αλλά δεν είναι επιθυμητή στις τεχνικές ψευδωνυμοποίησης.

  Οι υπογραφές δακτυλίου (ring signatures) επιτρέπουν σε ένα μέλος μιας ομάδα χρηστών που διαθέτουν ο καθένας ένα προσωπικό ζεύγος ιδιωτικού και δημόσιου κλειδιού, να υπογράφει (κρυπτογραφεί) ένα κείμενο έτσι ώστε να αποφανθεί με βεβαιότητα ότι το κείμενο υπογράφηκε από ένα μέλος της ομάδας, αλλά δεν μπορεί να ταυτοποιήσει ποιο είναι το μέλος αυτό, διακρίνοντάς το από τους υπόλοιπους.

  Οι τεχνολογίες που σχετίζονται με τις υπογραφές δακτυλίου χρησιμοποιούνται σε συστήματα «ανώνυμων» κρυπτονομισμάτων, με την πραγματοποίηση συναλλαγών για τις οποίες μπορεί να αποδειχθεί ότι μια συναλλαγή πραγματοποιήθηκε από ένα μέλος μιας ομάδας, χωρις να μπορεί να προσδιοριστεί ποιο μέλος συγκεκριμένα. Ο ENISA αναφέρει ότι «παρόλο που οι υπογραφές δακτυλίου αναφέρονται στην βιβλιογραφία ως ανώνυμες, στην πραγματικότητα αποτελούν ψευδωνυμοποιημένα δεδομένα».

  Η ιδιότητα αυτή κάνει τέτοιες τεχνολογίες κατάλληλες για την κατασκευή ομαδικών ψευδωνύμων (group pseudonyms) για χρήση σε πρωτόκολλα αυτόματης ιχνηλάτησης επαφών, όπως το Pronto-C2 στο οποίο θα αναφερθούμε σε επόμενο κεφάλαιο.

3 Σενάρια χρήσης σε εφαρμογές Υγείας

Θα μελετήσουμε υποθετικά σενάρια χρήσης σε εφαρμογές που μπορεί να είναι χρήσιμες σε περιστάσεις έκτακτης ανάγκης, εστιάζοντας περισσότερο σε αυτές που πιθανώς απαιτούν ευρεία εφαρμογή και συλλογή προσωπικών δεδομένων μεγάλου αριθμού υποκειμένων. Για τον σκοπό αυτό θα χρειαστεί να κάνουμε μια σειρά από υποθέσεις, οι οποίες θα προσπαθήσουμε να είναι όσο το δυνατόν ρεαλιστικές. Οι υποθέσεις αυτές αφορούν τους σκοπούς της επεξεργασίας και τα εμπλεκόμενα μέρη που με την σειρά τους θα επηρεάσουν τις υποθέσεις που θα κάνουμε για τους αντιπάλους (adversaries), τα πιθανά διανύσματα επιθέσεων (attack vector) αλλά και τις επιθυμητές ιδιότητες της μεθόδου, όπως αυτές που αναφέρθηκαν στην παράγραφο 2.2:

- Το επιθυμητό επίπεδο προστασίας δεδομένων
- Την τυχούσα επιθυμητή διατήρηση της χρησιμότητας (utility) των δεδομένων για περεταίρω επεξεργασία
- Την τυχούσα επιθυμητή επεκτασιμότητα των εφαρμογών σε μεγάλης κλίμακας αριθμό υποκειμένων
- Την τυχούσα επιθυμητή δυνατότητα ανάκτησης των αναγνωριστικών και της ταυτοποίησης υποκειμένων

Μια συνηθισμένη υπόθεση που θα κάνουμε είναι η συμμετοχή ως ένα από τα εμπλεκόμενα μέρη (υπεύθυνος επεξεργασίας, εκτελών την επεξεργασία, έμπιστη τρίτη οντότητα) κάποιου φορέα του δημόσιου τομέα, καθώς στις περισσότερες περιπτώσεις κάποιος δημόσιος φορέας έχει την αρμοδιότητα αντιμετώπισης έκτακτων καταστάσεων, ενώ είναι και πιθανότερο να προβλέπεται από την νομοθεσία η αρμοδιότητα συλλογής και επεξεργασίας δεδομένων.

Επίσης, κοινή υπόθεση στις ενδεικτικές περιπτώσεις χρήσης που εξετάζουμε είναι η εμπλοκή κάποιου ιδιωτικού φορέα. Η μέχρι τώρα εμπειρία έχει δείξει ότι, σε επείγουσες συνθήκες, η συλλογή επεξεργασία μεγάλου όγκου δεδομένων που αφορά πολλά υποκείμενα απαιτεί πόρους όπως οι δικτυακές υποδομές, οι δομές αποθήκευσης, η επεξεργαστική ισχύς, ακόμα και λογισμικό που έχει παραχθεί πιθανώς για διαφορετική από την ζητούμενη χρήση, αλλά με την κατάλληλη παραμετροποίηση είναι επαρκές. Οι πόροι αυτοί είναι απίθανο να υπάρχουν διαθέσιμοι εκ τον προτέρων και να συντηρούνται από δημόσιους οργανισμούς για την ενδεχόμενη χρήση τους σε έκτακτες συνθήκες, ενώ δεν είναι συμφέρουσα η έναρξη της ανάπτυξής τους μετά την εμφάνιση των έκτακτων αναγκών, είτε επειδή δεν υπάρχει ο απαιτούμενος χρόνος είτε επειδή υπάρχουν διαθέσιμοι για άλλες εφαρμογές. Συνεπώς είναι λογικό να διατίθενται ή να μισθώνονται τέτοιοι πόροι από τον ιδιωτικό τομέα.

Η εμπλοκή κάποιου ιδιωτικού φορέα αυξάνει τον αριθμό των εμπλεκόμενων μερών, την πολυπλοκότητα του συστήματος ψευδωνυμοποίησης και τις απαιτήσεις για προστασία των δεδομένων.

3.1 Σενάριο χρήσης σε εφαρμογές δήλωσης αποτελέσματος διαγνωστικού ελέγχου

3.1.1 Σκοπός της επεξεργασίας

Όπως έγινε αντιληπτό κατά την πρόσφατη πανδημία, κατά την διάρκεια παρόμοιων υγειονομικών κρίσεων είναι δυνατόν να δημιουργηθεί η ανάγκη ευρείας διενέργειας διαγνωστικών ελέγχων σε πολύ μεγάλους πληθυσμούς και για μεγάλα χρονικά διαστήματα για λόγους επιδημιολογικής επιτήρησης. Η χρήση των τεχνολογιών πληροφορικής και επικοινωνιών επιτρέπει την ταχύτατη μετάδοση και επεξεργασία των δεδομένων, και συνεπώς την επίκαιρη παρακολούθηση και εκτίμηση της εξέλιξης μιας υγειονομικής κρίσης για την έγκαιρη λήψη αποφάσεων. Ανάλογα με την πρακτική που ακολουθείται, οι έλεγχοι μπορεί να διενεργούνται από εξειδικευμένα εργαστήρια και άλλους ιδιωτικούς φορείς, από σταθερές και κινητές δομές δημοσίων φορέων, αλλά και από τους ίδιους τους πολίτες.

Σε όλες τις περιπτώσεις γίνεται ταυτοποίηση του ελεγχόμενου, οπότε είναι η απαραίτητη η καταχώριση των ατομικών αναγνωριστικών του κάθε υποκειμένου, ενώ υπάρχει η απαίτηση της δυνατότητας ενημέρωσης του υποκειμένου σχετικά με τυχόν θετικό αποτέλεσμα του διαγνωστικού ελέγχου και με τις ενέργειες που θα πρέπει το υποκείμενο να κάνει σε αυτή την περίπτωση, πράγμα που σημαίνει απαίτηση δυνατότητας ανάκτησης των αναγνωριστικών. Παράλληλα ισχύει η αρχική υπόθεση της απαίτησης περαιτέρω στατιστικής επεξεργασίας των δεδομένων για λόγους επιδημιολογικής επιτήρησης. Συνεπώς για τα σενάρια αυτά φαίνεται η ψευδωνυμοποίηση να είναι κατάλληλη τεχνική, καθώς γίνεται επεξεργασία δεδομένων δυνητικά και υπό όρους ταυτοποιήσιμων υποκειμένων.

Όσον αφορά στους σκοπούς της επεξεργασίας που στην συγκεκριμένη περίπτωση είναι η επιδημιολογική παρακολούθηση, είναι δεδομένη η καταχώριση για στατιστική επεξεργασία του αρνητικού ή θετικού αποτελέσματος ενός διαγνωστικού ελέγχου, όμως είναι δυνατόν να υπάρχει η απαίτηση για περισσότερο εκλεπτυσμένη επεξεργασία με την κατανομή των αποτελεσμάτων σε πληθυσμιακές ομάδες, όπως για παράδειγμα η γεωγραφική κατανομή των αποτελεσμάτων ή η κατανομή τους σε ηλικιακές ομάδες. Σε αυτές τις περιπτώσεις, θεωρούμε ότι δεν είναι ιδιαίτερα σημαντικό το πρόβλημα της ανωνυμοποίησης, καθώς θεωρήσαμε από την αρχή εφαρμογές ευρείας διενέργειας διαγνωστικών ελέγχων σε πολύ μεγάλα δείγματα του πληθυσμού και όχι στοχευμένους σε μικρές ομάδες ελέγχους.

Συνεπώς, η συλλογή και επεξεργασία περισσότερων δεδομένων για τα υποκείμενα δεν θα δημιουργήσει κινδύνους ταυτοποίησης τους χωρίς την χρήση αναγνωριστικών, επειδή τα μη

αναγνωριστικά δεδομένα συλλέγονται και καταχωρίζονται γενικευμένα. Για παράδειγμα, δεν απαιτείται σε τέτοιες εφαρμογές πεδίο «Διεύθυνση κατοικίας» αλλά πεδίο «Ευρύτερη γεωγραφική περιοχή», ούτε πεδίο «Ηλικία» αλλά πεδίο «Ηλικιακή κατηγορία». Ο συνδυασμός γενικευμένων δεδομένων και μεγάλου αριθμού υποκειμένων κάνουν εξαιρετικά δύσκολη την ταυτοποίηση χωρίς την χρήση αναγνωριστικών.

Η απαίτηση της ψευδωνυμοποίησης παραμένει επιτακτική: τα αναγνωριστικά δεν είναι χρήσιμα στην στατιστική επεξεργασία, αλλά όπως ήδη αναφέρθηκε είναι απαραίτητη η συλλογή και αποθήκευσή τους, τόσο για την αρχική ταυτοποίηση των υποκειμένων, όσο και για την διατήρηση της δυνατότητας ανάκτησης. Η ταυτοποίηση είναι στιγμιαία και κατάλληλα διοικητικά μέτρα και διαδικασίες θα μπορούσαν να επιβάλουν την άμεση διαγραφή των δεδομένων ταυτοποίησης αμέσως μετά από αυτήν, αν δεν υπάρχει η απαίτηση της ανάκτησης για την ενημέρωση και την παροχή οδηγιών στο υποκείμενο, σε περίπτωση θετικού διαγνωστικού ελέγχου. Όμως και στην περίπτωση απαίτησης δυνατότητας ανάκτησης, ο χρόνος αποθήκευσης των αναγνωριστικών μπορεί να είναι περιορισμένος, καθώς η φύση της υγειονομικής κρίσης μπορεί να επιβάλει την ταχεία ενημέρωση αλλά συγχρόνως να την καθιστά αναίτια μετά την παρέλευση κάποιου περιορισμένου χρονικού διαστήματος. Μετά την παρέλευση αυτού του χρονικού διαστήματος, είναι δυνατή η διαγραφή των αναγνωριστικών και η διατήρηση των ανωνυμοποιημένων δεδομένων για στατιστική επεξεργασία.

Μέχρι τώρα αναγνωρίσαμε ως σκοπούς της επεξεργασίας την στατιστική επιδημιολογική επιτήρηση και την παροχή οδηγιών σε περίπτωση θετικού ελέγχου, οι οποίοι είναι κοινοί για τις περιπτώσεις εφαρμογών στις οποίες θα αναφερθούμε στην συνέχεια, οι οποίες διαφέρουν στις απαιτήσεις ψευδωνυμοποίησης ανάλογα με την ταυτότητα και τον αριθμό των εμπλεκόμενων μερών.

3.1.2 Δήλωση αποτελεσμάτων διαγνωστικών ελέγχων που διενεργούνται από εξειδικευμένο προσωπικό ιδιωτικών και δημοσίων φορέων

Στην περίπτωση των διαγνωστικών ελέγχων που διενεργούνται από εξειδικευμένο προσωπικό ιδιωτικών ή δημοσίων φορέων, όσον αφορά στην προστασία των δεδομένων, υπάρχει το πλεονέκτημα της παρουσίας τοπικά κάποιου εξειδικευμένου επαγγελματία ή τοπικού φορέα, που με την λήψη των κατάλληλων διοικητικών μέτρων διασφαλίζει την τήρηση όλων των διαδικασιών για την προστασία των δεδομένων των υποκειμένων. Η ταυτοποίηση του ελεγχόμενου γίνεται τοπικά και τα αναγνωριστικά γνωστοποιούνται μόνο στον τοπικό φορέα που την διενεργεί την ταυτοποίηση. Τα εμπλεκόμενα μέρη είναι

- Το υποκείμενο των δεδομένων
- Ο τοπικός φορέας: δημόσια και ιδιωτικά διαγνωστικά κέντρα, ιατροί, φαρμακεία, δημόσιες και ιδιωτικές κλινικές και οποιοσδήποτε φορέας διενεργεί διαγνωστικούς ελέγχους
- Ο κεντρικός φορέας: θεωρούμε κάποιον κεντρικό φορέα που συλλέγει δεδομένα από όλους τους τοπικούς φορείς και τα επεξεργάζεται για την εξαγωγή συμπερασμάτων

Καθώς οι σκοποί της επεξεργασίας θεωρούμε ότι αφορούν στην στατιστική επεξεργασία για λόγους επιδημιολογικής επιτήρησης, η ταυτότητα των υποκειμένων δεν αναγκαιεί για την επεξεργασία οπότε και δεν χρειάζεται να γνωστοποιηθεί ούτε στον κεντρικό φορέα. Όμως πιθανώς να υπάρχει η απαίτηση σε περίπτωση θετικού ελέγχου να υπάρξει επικοινωνία του κεντρικού φορέα με το υποκείμενο των δεδομένων για την διερεύνηση και επιβεβαίωση με νέο διαγνωστικό έλεγχο, την ιχνηλάτηση των κοινωνικών επαφών, την επιβολή κοινωνικής αποστασιοποίησης κ.α.

Πιθανώς είναι λειτουργική η απόδοση στον τοπικό φορέα του ρόλου της οντότητας ψευδωνυμοποίησης που θα εκτελεί την ψευδωνυμοποίηση, θα φυλάσσει τα ψευδωνυμοποιημένα δεδομένα και το μυστικό ψευδωνυμοποίησης και θα εκτελεί την ανάκτηση των αναγνωριστικών όταν χρειάζεται, ή γενικότερα του ρόλου του συνοδού των δεδομένων (data custodian) που επιπρόσθετα θα παρέχει πρόσβαση στα δεδομένα και τα αναγνωριστικά ανάλογα με τις επιθυμίες του ελεγχόμενου και την νομοθεσία.

Εφόσον θεωρήσαμε εφαρμογές που αφορούν σε μεγάλο αριθμό υποκειμένων με σκοπό την επιδημιολογική παρακολούθηση, πρέπει να υποθέσουμε και αντίστοιχα μεγάλο αριθμό τοπικών φορέων. Αυτή η αποκεντρωμένη προσέγγιση, προσθέτει ευθύνες προστασίας των δεδομένων σε μεγάλο αριθμό τοπικών φορέων, που πιθανόν να κάνει λίγο δυσκολότερο τον συντονισμό, την εκπαίδευση και τις ανάγκες σε τεχνολογικό εξοπλισμό σε σχέση με την διαχείριση των αναγνωριστικών και της ψευδωνυμοποίησης από έναν κεντρικό φορέα.

Χαρακτηρίζεται όμως από ένα σημαντικό πλεονέκτημα, καθώς καθένας από τους τοπικούς φορείς έχει πρόσβαση σε αναγνωριστικά συγκριτικά ελάχιστων υποκειμένων σε σχέση με το σύνολο των υποκειμένων. Έτσι, σε περίπτωση διαρροής ή παραβίασης του συστήματος, ο αντίκτυπος (impact) θα είναι ασύγκριτα μικρότερος σε σχέση με ένα κεντρικοποιημένο σύστημα. Επιπρόσθετα, καθώς κάθε τοπικός φορέας χειρίζεται δεδομένα συγκριτικά πολύ μικρού αριθμού υποκειμένων, οι ανάγκες σε αποθηκευτικά μέσα και υπολογιστική ισχύ είναι πολύ μικρότερες και πολύ πιθανό να καλύπτονται από τον ήδη υπάρχοντα τεχνολογικό εξοπλισμό. Η οποιαδήποτε δυσχέρεια φαίνεται να περιορίζεται στην τήρηση των διαδικασιών για την διαφύλαξη του μυστικού ψευδωνυμοποίησης, διαδικασίες που μπορούν να διευκολύνονται από το λογισμικό του τοπικού φορέα.

Μια ντετερμινιστική πολιτική ψευδωνυμοποίησης θα αντικαθιστά ένα αναγνωριστικό του υποκειμένου σε όλες τις περιστάσεις με το ίδιο ψευδώνυμο. Αυτό κάνει πολύ εύκολη την αναζήτηση, διατηρώντας κάποια χρησιμότητα (utility). Ένα παράδειγμα θα ήταν η προσέλευση ενός υποκειμένου με αρνητικό αποτέλεσμα διαγνωστικού ελέγχου, που ζητά να διαπιστωθεί η ανάρρωση. Η αναζήτηση προηγούμενου θετικού αποτελέσματος θα είναι πολύ εύκολη, καθώς η αναζήτηση στην βάση δεδομένων θα γίνεται χρησιμοποιώντας το ψευδώνυμο του αναγνωριστικού που είναι σε όλες τις περιστάσεις το ίδιο.

Η ντετερμινιστική πολιτική μπορεί να εμφανίζει κάποιες αδυναμίες σχετικές με την προστασία των δεδομένων. Αν υποθέσουμε ότι ένας επιτιθέμενος αποκτά πρόσβαση στην ψευδωνυμοποιημένη βάση δεδομένων αλλά όχι στο μυστικό ψευδωνυμοποίησης, και συγχρόνως γνωρίζει ότι ένα υποκείμενο υποβλήθηκε σε διαγνωστικούς ελέγχους κάποιες συγκεκριμένες ημερομηνίες. Πιθανότατα η γνώση ενός αριθμού ημερομηνιών να του επιτρέψει να ταυτοποιήσει το υποκείμενο και να πληροφορηθεί τα αποτελέσματα όλων των ελέγχων του υποκειμένου.

Μια ανά έγγραφο τυχαιοποιημένη πολιτική ψευδωνυμοποίησης μπορεί να προσφέρει καλύτερη προστασία των δεδομένων, διατηρώντας παράλληλα δυνατότητες αναζητήσεων. Ως έγγραφο στην συγκεκριμένη περίπτωση θεωρούμε την βάση δεδομένων του κάθε τοπικού φορέα. Με την γνωστοποίηση στον κεντρικό φορέα των ψευδώνυμων ενός αναγνωριστικού μπορούν να γίνουν αναζητήσεις στο σύνολο των βάσεων.

Μια πλήρως τυχαιοποιημένη πολιτική, θα προσφέρει το καλύτερο επίπεδο ασφάλειας έναντι τέτοιων επιθέσεων, χωρίς όμως να δίνει την δυνατότητα αναζητήσεων. Στην περίπτωση που ότι οι σκοποί της επεξεργασίας περιλαμβάνουν μόνο απλή στατιστική επεξεργασία όπως για παράδειγμα η παρακολούθηση των ημερήσιων θετικών διαγνωστικών ελέγχων, η πλήρως τυχαιοποιημένη πολιτική είναι μια επιλογή που διασφαλίζει καλύτερη προστασία των δεδομένων.

Η τεχνική ψευδωνυμοποίησης που θα χρησιμοποιηθεί εξαρτάται κι αυτή από διάφορους παράγοντες. Η χρήση Hash function θα προσφέρει μια πολύ απλοποιημένη εφαρμογή που μπορεί να εκτελεστεί εύκολα από το λογισμικό του τοπικού φορέα, παρουσιάζει όμως κάποια αδυναμία στην προστασία των δεδομένων. Η επιλογή κάποιας τεχνικής που κάνει χρήση κρυπτογραφικού κλειδιού όπως για παράδειγμα η HMAC είναι πιο ασφαλής, επιβάλει όμως την ασφαλή φύλαξη του κρυπτογραφικού κλειδιού από τον τοπικό φορέα. Καθώς οι προηγούμενες είναι μη-αντιστρεπτές διαδικασίες, απαιτείται η αποθήκευση του πίνακα αντιστοίχισης των αναγνωριστικών με τα ψευδώνυμα, όμως αυτό σε πολλές περιπτώσεις θα κάνει τις αναζητήσεις στην βάση δεδομένων ταχύτερη. Αν ο υπάρχων τεχνολογικός εξοπλισμός σε συνδυασμό με τον αριθμό των ελέγχων επιβάλλουν περιορισμούς στον αποθηκευτικό χώρο, η κρυπτογράφηση είναι αντιστρεπτή διαδικασία που επιτρέπει, αν κάποιος γνωρίζει το κλειδί κρυπτογράφησης, την ανάκτηση των αναγνωριστικών χωρίς να χρειάζεται η

αποθήκευση του πίνακα αντιστοιχίσεων.

Στην απλή περίπτωση της αρχικής υπόθεσης όπου ο κεντρικός φορέας επεξεργάζεται μόνο τα ημερήσια στοιχεία για την στατιστική παρακολούθηση των ελέγχων, είναι δυνατόν να χρησιμοποιηθεί κάποια πλήρως τυχαιοποιημένη τεχνική ψευδωνυμοποίησης με την χρήση κρυπτογραφικού κλειδιού, διαφορετικού για κάθε τοπικό φορέα. Στην περίπτωση που αναγκαιεί ανάκτηση των αναγνωριστικών, μόνο ο τοπικός φορέας θα είναι σε θέση να την εκτελέσει.

Η πιθανότητα απαίτησης για την έκδοση ηλεκτρονικού πιστοποιητικού του οποίου η διακρίβωση της αυθεντικότητας θα γίνεται ηλεκτρονικά θα δημιουργούσε την ανάγκη της εμπλοκής ενός κεντρικού φορέα, όμως υπάρχουν τρόποι αυτό να γίνει χωρίς την πρόσβαση αυτού στα στοιχεία ταυτοποίησης. Υποθέτουμε την έκδοση πιστοποιητικού διαγνωστικού ελέγχου από τον τοπικό φορέα, το οποίο υπογράφει με το ιδιωτικό του κλειδί, στα πλαίσια κάποιου από τα ήδη χρησιμοποιούμενα σχήματα ψηφιακών υπογραφών. Ο τοπικός φορέας έχει κάνει διαθέσιμο το δημόσιο κλειδί του μέσω ενός έμπιστου τρίτου μέρους. Η αποθήκευση και ο έλεγχος των πιστοποιητικών συνήθως θα γίνεται μέσω φορητών συσκευών. Το λογισμικό του ελεγχόμενου μπορεί να στέλνει το πιστοποιητικό στην συσκευή του ελεγκτή, η οποία και θα πραγματοποιεί τον έλεγχο της γνησιότητας.

Στο παραπάνω σενάριο ακολουθώντας μια προσέγγιση με βάση τον κίνδυνο (risk based approach) θεωρήσαμε ότι για τον κεντρικό φορέα δεν είναι απαραίτητη η γνώση των αναγνωριστικών για την εκπλήρωση των σκοπών της επεξεργασίας, συνεπώς θεωρήσαμε ότι μπορεί σε αυτόν να υπάρξει κάποιος εσωτερικός αντίπαλος (insider adversary) επιπρόσθετα από τυχόντες εξωτερικούς αντιπάλους.

Υπάρχει η περίπτωση ο κεντρικός φορέας να είναι κάποια οντότητα η οποία από τον νόμο ή με την συγκατάθεση του υποκειμένου των δεδομένων να έχει το δικαίωμα πλήρους πρόσβασης στα δεδομένα ταυτοποίησης και περίπλοκων στατιστικών αναζητήσεων στο σύνολο των δεδομένων για τους σκοπούς της επεξεργασίας. Σε αυτή την περίπτωση, είναι δυνατόν να γίνει ανάθεση (delegation) της διαδικασίας ψευδωνυμοποίησης στους τοπικούς φορείς με την χρήση ασύμμετρων κρυπτογραφικών αλγορίθμων, όπως περιγράφεται στο [3].

Σε αυτό το σενάριο, ο κεντρικός φορέας παράγει το ζεύγος του ιδιωτικού και του δημόσιου κλειδιού του και ακολούθως γνωστοποιεί το δημόσιο κλειδί του στους τοπικούς φορείς. Δεν μπορούμε να αναγνωρίσουμε καμία απαίτηση να γνωστοποιηθεί το δημόσιο κλειδί σε οποιονδήποτε άλλον ενώ το να γνωστοποιηθεί μόνο στους τοπικούς φορείς συμβάλει στην προστασία των δεδομένων, συνεπώς είναι σημαντικό η διανομή να γίνει μέσω ενός εύχρηστου αλλά ασφαλούς καναλιού. Οι τοπικοί φορείς κρυπτογραφούν τα αναγνωριστικά με το δημόσιο κλειδί του κεντρικού φορέα χρησιμοποιώντας τυχαιοποιημένη ασύμμετρη κρυπτογράφηση παράγοντας έτσι τα ψευδώνυμα και αποστέλλουν σε αυτόν τα ψευδωνυμοποιημένα δεδομένα. Παρά την υπόθεση ότι η αποστολή γίνεται

με κάποιο κρυπτογραφημένο κανάλι, η ψευδωνυμοποίηση των αναγνωριστικών προσθέτει ένα ακόμα επίπεδο ασφάλειας.

Με αυτό τον τρόπο, ο κεντρικός φορέας λαμβάνει και επεξεργάζεται ψευδωνυμοποιημένα δεδομένα, διατηρώντας την δυνατότητα ανάκτησης των αναγνωριστικών αν αυτή απαιτείται, αποκρυπτογραφώντας τα ψευδώνυμα με το ιδιωτικό του κλειδί. Επίσης ο κεντρικός φορέας είναι σε θέση να λάβει κατάλληλα διοικητικά και οργανωτικά μέτρα σχετικά με την πρόσβαση στο ιδιωτικό του κλειδί, άρα και στην δυνατότητα ανάκτησης. Η οντότητα/ες που θα έχουν πρόσβαση στο ιδιωτικό κλειδί μπορεί να είναι μέρος του κεντρικού φορέα ή κάποια τρίτη οντότητα, που θα προβαίνει στην ανάκτηση όταν απαιτείται.

3.1.3 Δήλωση αποτελεσμάτων αυτοδιαγνωστικών ελέγχων

Είναι δυνατόν να υπάρχει η απαίτηση για την ηλεκτρονική υποβολή αποτελεσμάτων αυτοδιαγνωστικών ελέγχων από τα ίδια τα υποκείμενα των δεδομένων. Σε αυτή την περίπτωση δεν υπάρχει κάποιος τοπικός φορέας που θα μπορούσε να διασφαλίσει την αξιοπιστία της διαδικασίας και να βοηθήσει τον ελεγχόμενο, ο οποίος θα πρέπει να κάνει την δήλωση ο ίδιος. Η υπόθεση είναι ότι ο αυτοδιαγνωστικός έλεγχος γίνεται διαθέσιμος σε πολύ μεγάλα τμήματα του πληθυσμού, με χαμηλό κόστος ή χωρίς καθόλου κόστος, εθελοντικός ή υποχρεωτικός, δίνοντας έτσι την δυνατότητα συλλογής πολύ μεγάλου πλήθους δεδομένων με σκοπό την ακριβέστερη παρακολούθηση μιας υγειονομικής κρίσης.

Υποθέτουμε ότι ο πλέον εύχρηστος τρόπος που μπορεί να εξυπηρετήσει μια τέτοιας έκτασης εφαρμογή είναι η καταχώριση του αποτελέσματος μέσω μιας διαδικτυακής εφαρμογής ή μιας εφαρμογής φορητής συσκευής όπως ενός κινητού τηλεφώνου. Εφόσον η δήλωση γίνεται από τα ίδια τα υποκείμενα των δεδομένων, δεν υπάρχει το πλεονέκτημα ενός έμπιστου τοπικού διάμεσου που χειρίζεται ευαίσθητα αλλά λίγα σε πλήθος προσωπικά δεδομένα.

Για την βελτίωση της αξιοπιστίας των δηλωθέντων αποτελεσμάτων, προτείνουμε την χρήση μιας μεθόδου παρόμοιας με την μέθοδο εξουσιοδότησης χρήστη που προτείνουν οι Avitabile, Bottay, Iovinoz και Visconti για το σύστημα ιχνηλάτησης επαφών Pronto-C2 [4] : στην συσκευασία του αυτοδιαγνωστικού τεστ μπορεί να περιλαμβάνεται ένας τυχαίος κωδικός αριθμός που θα επιτρέπει στον χρήστη την δήλωση αποτελέσματος, εφόσον είναι έγκυρος. Η μέθοδος αυτή δεν εξασφαλίζει την αληθή δήλωση, αλλά μόνο το γεγονός ότι ο χρήστης έχει πράγματι στην κατοχή του ένα

διαγνωστικό τεστ.

Η συγκέντρωση των αποτελεσμάτων θα πρέπει να γίνεται από έναν κεντρικό φορέα ο οποίος θα πρέπει να λαμβάνει διοικητικά και τεχνικά μέτρα για την ασφάλεια των δεδομένων και την ελεγχόμενη πρόσβαση σε αυτά. Καθώς ο όγκος των δεδομένων σε αυτή την περίπτωση είναι μεγάλος, ο αντίκτυπος μιας διαρροής θα είναι πολύ μεγαλύτερος. Θεωρούμε ότι από την φύση των δεδομένων είναι απαραίτητη και σε αυτή την περίπτωση η ταυτοποίηση του υποκειμένου, για λόγους αξιοπιστίας των δεδομένων.

Μια κατάλληλη αρχιτεκτονική θα ήταν να γίνει διαχωρισμός της διαδικασίας της ταυτοποίησης από αυτήν της επεξεργασίας. Σε μια τέτοια αρχιτεκτονική, υποθέτουμε έναν φορέα ταυτοποίησης ο οποίος θα αναλαμβάνει την εξακρίβωση της ταυτότητας του υποκειμένου. Συνεπώς μπορούμε να θεωρήσουμε ως εμπλεκόμενα μέρη

- Το υποκείμενο των δεδομένων
- Τον φορέα ταυτοποίησης
- Τον κεντρικό φορέα

Σε αυτόν τον απλοποιημένο διαχωρισμό, ο κεντρικός φορέας μπορεί να είναι ο υπεύθυνος επεξεργασίας ή/και ο εκτελών την επεξεργασία, ο οποίος θα λαμβάνει από τον φορέα ταυτοποίησης την πληροφορία ότι συγκεκριμένο υποκείμενο είναι κάποιο ταυτοποιημένο φυσικό πρόσωπο που μπορεί να κάνει χρήση της συγκεκριμένης υπηρεσίας. Η σύνδεση με την υπηρεσία μπορεί να γίνεται με ανακατεύθυνση στον φορέα ταυτοποίησης, ταυτοποίηση με όνομα χρήστη και κωδικό και επιστροφή στην υπηρεσία με κάποιο token (για παράδειγμα ένα One Time Password περιορισμένης χρονικής ισχύος) για καταχώριση του αποτελέσματος. Όπως και στην προηγούμενη παράγραφο, απαιτείται η διατήρηση της δυνατότητας ανάκτησης των αναγνωριστικών για την διερεύνηση και επιβεβαίωση με νέο διαγνωστικό έλεγχο, την ιχνηλάτηση των κοινωνικών επαφών, την επιβολή κοινωνικής αποστασιοποίησης κ.α.

Η διαδικασία της ταυτοποίησης είναι στιγμιαία, και μετά από αυτήν τα αναγνωριστικά του υποκειμένου στην γενική περίπτωση δεν είναι απαραίτητα για την εκτέλεση απλής στατιστικής επεξεργασίας. Ο φορέας ταυτοποίησης είναι δεδομένο ότι αποθηκεύει επεξεργάζεται αναγνωριστικά ταυτοποίησης, αλλά δεν απαιτείται να στέλνει στον κεντρικό φορέα τα αναγνώσιμα αναγνωριστικά του υποκειμένου, αλλά μόνο τα ψευδώνυμά τους που θα χρησιμεύσουν μετά από ανάκτηση στην επαναταυτοποίηση του υποκειμένου όταν χρειαστεί. Έτσι είναι δυνατόν ο φορέας ταυτοποίησης να αναλαμβάνει ταυτόχρονα και τον ρόλο της οντότητας ψευδωνυμοποίησης.

Ο φορέας ταυτοποίησης θα πρέπει να είναι κάποιος δημόσιος ή ιδιωτικός αξιόπιστος οργανισμός που ήδη έχει θεσμοθετημένη σχέση εμπιστοσύνης με το υποκείμενο. Παραδείγματα φορέων που ήδη έχουν συμμετάσχει σε τέτοια σχήματα θα μπορούσε να είναι κάποιος δημόσιος φορολογικός οργανισμός, οργανισμός κοινωνικής ασφάλισης, ισότοποι συναλλαγών με το κράτος κ.α. καθώς και ιδιωτικοί οργανισμοί όπως για παράδειγμα τράπεζες. Σε μια τέτοια εφαρμογή θα πρέπει να υπάρχει εκ το προτέρων η δυνατότητα σύνδεσης πολύ μεγάλου αριθμού υποκειμένων, και σε πολλές περιπτώσεις το σύνολο του πληθυσμού μιας χώρας.

Στην αρχιτεκτονική που περιγράψαμε, ο φορέας ταυτοποίησης μπορεί να αποστέλλει στον κεντρικό φορέα αποκλειστικά και μόνο ψευδωνυμοποιημένα αναγνωριστικά, ενώ η καταχώριση των αποτελεσμάτων θα γίνεται απ' ευθείας στον κεντρικό φορέα. Ως οντότητα ψευδωνυμοποίησης, ο φορέας ταυτοποίησης αναλαμβάνει και την ανάκτηση των αναγνωριστικών, μετά από αίτημα του κεντρικού φορέα.

Η επιλογή πολιτικής ψευδωνυμοποίησης εξαρτάται όπως και στην παράγραφο 3.1 από τις απαιτήσεις.

Είναι δυνατόν να υπάρχουν απαιτήσεις αναζητήσεων σε παρελθοντικές δηλώσεις, είτε για λόγους παρακολούθησης της υγειονομικής κρίσης, είτε για λόγους ανίχνευσης ανακριβών δηλώσεων, πράγμα που θα υποδείξει μια ντετερμινιστική πολιτική, η οποία πιθανώς να έχει και το επιπλέον πλεονέκτημα του μικρότερου μεγέθους πίνακα αντιστοιχίσεων, θα πρέπει όμως να ληφθούν παράλληλα από τον κεντρικό φορέα κατάλληλα οργανωτικά και τεχνικά μέτρα ασφάλειας και ελέγχου πρόσβασης στα ψευδωνυμοποιημένα δεδομένα. Η τεχνική ψευδωνυμοποίησης που θα επιλεγεί μπορεί εδώ να είναι περισσότερο ασφαλής, αν θεωρήσουμε ότι ο φορέας ταυτοποίησης διαθέτει επαρκή τεχνολογικό εξοπλισμό.

Όπως και στην παράγραφο 3.1, μέχρι εδώ υποθέσαμε ότι ο σκοπός της επεξεργασίας του κεντρικού φορέα είναι η απλή στατιστική επεξεργασία ημερήσιων στοιχείων για λόγους παρακολούθησης μιας υγειονομικής κρίσης. Υποθέσαμε και πάλι ότι μπορεί να λειτουργήσει ως εσωτερικός αντίπαλος, ενώ παράλληλα τα αναγνωριστικά δεν αναγκαιούν για τους σκοπούς την επεξεργασία.

Είναι όμως ρεαλιστικό να υποθέσουμε την περίπτωση υποχρεωτικής από την νομοθεσία δήλωσης αυτοδιαγνωστικού ελέγχου, συνεπώς και έκδοσης βεβαίωσης για την σχετική δήλωση. Σε αυτή την περίπτωση απαιτείται πρόσβαση του κεντρικού φορέα στα δεδομένα ταυτοποίησης και συνεπώς ο φορέας ταυτοποίησης θα πρέπει να αποστέλει στον κεντρικό αμέσως μετά την ταυτοποίηση. Ομοίως θα μπορούσε κι εδώ να χρησιμοποιηθεί η ασύμμετρη κρυπτογραφία ως εργαλείο για να γίνει δυνατή η ανάθεση της ψευδωνυμοποίησης, ειδικά στην περίπτωση που οι φορείς ταυτοποίησης είναι περισσότεροι από έναν. Οι φορείς ταυτοποίησης θα κρυπτογραφούν τα αναγνωριστικά με το δημόσιο κλειδί του κεντρικού φορέα, ο οποίος μπορεί να τα αποκρυπτογραφήσει με το ιδιωτικό του κλειδί.

3.2 Σενάριο χρήσης σε εφαρμογές ιχνηλάτησης επαφών

Κατά την διάρκεια της πρόσφατης πανδημίας, εφαρμόστηκαν στην πράξη συστήματα Αυτόματης Ιχνηλάτησης Επαφών (Automatic Contact Tracing -ACT). Τα περισσότερα από αυτά έχουν σαν κοινό χαρακτηριστικό την εγκατάσταση από τον χρήστη μιας εφαρμογής σε φορητή συσκευή που συνήθως φέρει μαζί του, η οποία με την χρήση τεχνολογίας Bluetooth Low Energy (Bluetooth LE) καταγράφει όλες τις κοντινές επαφές του χρήστη με άλλους χρήστες της ίδιας εφαρμογής [5].

Μέχρι τότε, η ιχνηλάτηση επαφών γινόταν μόνο από τις υγειονομικές αρχές μετά από συνέντευξη με τον ασθενή, ο οποίος χρειαζόταν να ανακαλέσει στην μνήμη του τις πρόσφατες επαφές του. Η τεχνολογία Bluetooth LE και κυρίως η χρήση της σε φορητές συσκευές που ένα εξαιρετικά μεγάλο ποσοστό του πληθυσμού φέρει μαζί του το μεγαλύτερο μέρος της ημέρας όπως είναι τα smartphone, έδωσε την δυνατότητα της αυτόματης και περισσότερο εξαντλητικής καταγραφής τους.

Όταν ένας χρήστης εγκαταστήσει σε φορητή συσκευή μια τέτοια εφαρμογή, αυτή εκπέμπει περιοδικά ένα σήμα Bluetooth LE στο οποίο συμπεριλαμβάνεται και η μέτρηση της ισχύος του ίδιου του εκπεμπόμενου σήματος. Στην περίπτωση που ένας άλλος χρήστης που έχει εγκαταστήσει την εφαρμογή βρίσκεται σε κοντινή απόσταση, η συσκευή του θα λάβει το σήμα Bluetooth LE, και θα είναι σε θέση να υπολογίσει την απόσταση, συγκρίνοντας την μέτρηση της ισχύος εκπομπής που περιέχεται στο σήμα, με την μέτρηση της ισχύος του σήματος κατά την λήψη του. Οι επαφές που είναι πιο κοντινές σε απόσταση από ένα όριο που τίθεται, θεωρούνται περισσότερο επικίνδυνες.

Με αυτό τον τρόπο, όταν κάποιος χρήστης δηλώσει θετική διάγνωση, μπορεί να γίνει ανάκληση των επαφών έτσι ώστε όλοι χρήστες που ήρθαν σε κοντινή επαφή μαζί του να ειδοποιηθούν, και επομένως να λάβουν κατάλληλα μέτρα για την διάγνωση πιθανής μόλυνσης και την αποφυγή περεταίρω μετάδοσης.

Η χρήση μιας τέτοιου είδους εφαρμογής για την ιχνηλάτηση των κοινωνικών επαφών ενός χρήστη φαίνεται να είναι κατάλληλη για μια περίοδο πανδημίας μεταδοτικής ασθένειας, όμως θεωρητικά θα μπορούσε να χρησιμοποιηθεί στο μέλλον και σε όμοιες ή και διαφορετικές καταστάσεις. Η καταγραφή των κοινωνικών επαφών των χρηστών είναι μια διαδικασία που θα ήταν δυνατόν να παράγει πρόσθετες πληροφορίες, όπως για παράδειγμα σχετικές με την συγκέντρωση προσώπων σε συγκεκριμένους χώρους ή την αναδρομή της εκτέλεσης συγκεκριμένων διαδικασιών από αρμόδιους κατά την διάρκεια καταστάσεων έκτακτης ανάγκης. Το πλήθος ,το εύρος και η φύση των προσωπικών δεδομένων που μπορεί να εκθέσει μια τέτοια εφαρμογή είναι η βάση επιχειρημάτων υπέρ της χρησιμότητάς της, αλλά ταυτόχρονα είναι και πηγή κινδύνων για τα προσωπικά δεδομένα σε περίπτωση

που θα γίνει χρήση της για σκοπούς διαφορετικούς από τους δηλωθέντες. Καθώς, στις μέχρι τώρα πρακτικές εφαρμογές, η χρήση τέτοιων εφαρμογών είναι προαιρετική, η εξασφάλιση της εμπιστοσύνης των χρηστών μπορεί να επιδράσει θετικά στον αριθμό των συμμετεχόντων χρηστών.

Με αυτό το σκεπτικό, κατά την διάρκεια της πανδημίας προτάθηκε ένας αριθμός από πρωτόκολλα, για τα οποία σημαντικός στόχος μεταξύ άλλων είναι η εξυπηρέτηση απαιτήσεων προστασίας της ταυτότητας των χρηστών. Τα περισσότερα κάνουν χρήση τεχνικών ψευδωνυμοποίησης για την προστασία της ταυτότητας του χρήστη, σε πολλές περιπτώσεις διατηρώντας παράλληλα την δυνατότητα ανάκλησης των στοιχείων ταυτοποίησης, εφόσον είναι απαραίτητο να ειδοποιηθούν οι χρήστες σε περίπτωση που είχαν κάποια επαφή με χρήστη που δήλωσε θετική διάγνωση κατά το χρονικό διάστημα που προηγείται της δήλωσης.

Ένα από τα σημαντικότερα θέματα που αφορά στην προστασία της ταυτότητας των χρηστών κατά την χρήση τέτοιων πρωτοκόλλων, είναι η διάκριση σε δύο είδη, ανάλογα με την συγκεντρωτική ή την αποκεντρωμένη επεξεργασία των αναφορών θετικής διάγνωσης. Η διάκριση αφορά στην διαχείριση της υποδομής αποθήκευσης και επεξεργασίας των αρχείων καταγραφής επαφών που συγκροτούνται για κάθε χρήστη.

Στα συγκεντρωτικά (centralized) πρωτόκολλα, υπάρχει ένας κεντρικός διακομιστής τον οποίο διαχειρίζεται ένας κεντρικός φορέας. Τα αρχεία καταγραφής συγκροτούνται από καταγραφές επαφών που συλλέγει η συσκευή του χρήστη κατά τις καθημερινές του δραστηριότητες. Όταν διαπιστωθεί θετική διάγνωση, ο χρήστης πρέπει να μεταφορτώσει το αρχείο καταγραφής όλων των επαφών του για κάποιο συγκεκριμένο χρονικό διάστημα. Η επεξεργασία γίνεται από τον κεντρικό διακομιστή, ο οποίος αποφαίνεται για τους χρήστες που είχαν κάποια επαφή με τον χρήστη που παρουσίασε θετική διάγνωση, και τους ειδοποιεί. Έτσι ο διαχειριστής του κεντρικού διακομιστή, έχει πρόσβαση στις καταγεγραμμένες επαφές χρηστών για συγκεκριμένα χρονικά διαστήματα, και δυνατότητα επικοινωνίας με αυτούς.

Στα αποκεντρωμένα (decentralized) πρωτόκολλα, η επεξεργασία των αρχείων καταγραφής δεν γίνεται από τον κεντρικό διακομιστή αλλά από το δίκτυο των συσκευών των χρηστών. Συνεπώς δεν υπάρχει ανάγκη μεταφόρτωσης των αρχείων καταγραφής των επαφών και δυνατότητας πρόσβασης του κεντρικού φορέα σε αυτές.

Στην συνέχεια θα αναφερθούμε σε τρία πρωτόκολλα που προτάθηκαν κατά την διάρκεια της πρόσφατης πανδημίας με προσανατολισμό την χρήση τους για την ιχνηλάτηση επαφών για επιδημιολογικούς σκοπούς, το Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT/PEPP), το Decentralized Privacy-Preserving Proximity Tracing (DP-3T) και το Pronto-C2.

3.2.1 Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT/PEPP)

Το Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT/PEPP) είναι ένα ανοιχτό πρωτόκολλο που παρουσιάστηκε τον Απρίλιο του 2020. Η ανάπτυξή του υποστηρίχθηκε από μια ομάδα Ευρωπαϊκών ακαδημαϊκών και ερευνητικών φορέων με προσανατολισμό στην χρήση του για σκοπούς αντιμετώπισης της πανδημίας του COVID-19, που είχε αρχίσει να εξελίσσεται.

Σε αυτό το σχήμα όπως περιγράφεται στην τεκμηρίωση της Γερμανικής υλοποίησης [6], ο χρήστης εγκαθιστά μια εφαρμογή στο κινητό του τηλέφωνο και την ενεργοποιεί ανώνυμα μέσα από μια διαδικασία που συνδυάζει ένα πρωτόκολλο Proof of Work (PoW) με captcha, για την αποφυγή αυτοματοποιημένης δημιουργίας ψεύτικων λογαριασμών. Η διαδικασία γίνεται μέσω της επικοινωνίας με έναν κεντρικό διακομιστή, τον οποίο διαχειρίζεται ένας κεντρικός φορέας, όπως για παράδειγμα οι υγειονομικές αρχές μιας χώρας.

Ο διακομιστής αποδίδει στην συσκευή ένα μόνιμο αναγνωριστικό ταυτοποίησης που ονομάζεται PUID. Το PUID είναι αναγνωριστικό ταυτοποίησης που επιτρέπει στον διακομιστή να διακρίνει τον χρήστη και να στέλνει ειδοποιήσεις στο κινητό τηλέφωνο του σε περίπτωση που διαπιστωθεί ότι κάποια από τις επαφές του αναφέρει θετική διάγνωση. Το PUID δεν μεταβάλλεται και δεν μπορεί να εκπέμπεται στις συσκευές των άλλων χρηστών, καθώς κάποιος θα μπορούσε να το καταγράψει και στο εξής να αναγνωρίζει την συσκευή του χρήστη σε κάθε προσέγγιση, αποκαλύπτοντας έτσι την ταυτότητα του χρήστη. Προκύπτει έτσι η ανάγκη ψευδωνυμοποίησης του αναγνωριστικού ταυτοποίησης PUID.

Η τεχνική ψευδωνυμοποίησης που χρησιμοποιείται βασίζεται στην συμμετρική κρυπτογράφηση με τον αλγόριθμο Advanced Encryption Standard (AES). Ο κεντρικός διακομιστής, παράγει περιοδικά ένα καθολικό μυστικό κλειδί BK_t το οποίο έχει ισχύ για ένα μικρό χρονικό διάστημα (π.χ. μία ώρα) όπου t είναι το συγκεκριμένο χρονικό διάστημα για το οποίο ισχύει το συγκεκριμένο μυστικό κλειδί για όλους τους χρήστες. Χρησιμοποιώντας αυτό το μυστικό κλειδί παράγει για όλους τους χρήστες ένα εφήμερο αναγνωριστικό ταυτοποίησης Ephemeral Bluetooth ID (EBID) κρυπτογραφώντας το μόνιμο PUID με τον αλγόριθμο AES

$$EBID_t(PUID) = AES(BK_t, PUID)$$

και ακολούθως αποστέλλει σε κάθε χρήστη τα αντίστοιχα $EBID_t$ που παρήγαγε από το PUID του

χρήστη με τα καθολικά κλειδιά BK_t.[1]

Η PEPP-PT εφαρμογή του χρήστη εκπέμπει διαρκώς περιοδικά σήμα που περιλαμβάνει το εφήμερο αναγνωριστικό ταυτοποίησης EBID και λαμβάνει διαρκώς τα σήματα της ίδιας εφαρμογής που τυχαίνει να έχουν εγκαταστήσει οι χρήστες με τους οποίους έρχεται σε κοντινή επαφή. Έτσι η εφαρμογή του χρήστη καταγράφει και αποθηκεύει στην συσκευή του χρήστη όλα τα EBID των άλλων χρηστών με τους οποίους έρχεται σε κοντινή επαφή, μαζί με δεδομένα που επιτρέπουν τον υπολογισμό της απόστασης, του χρόνου και της διάρκειας της συγκεκριμένης επαφής (Contact/Time data - CTD) έτσι ώστε εάν προκύψει η ανάγκη να είναι δυνατόν να εκτιμηθεί η επικινδυνότητα της επαφής.

Όταν διαπιστωθεί και δηλωθεί στον κεντρικό φορέα θετική διάγνωση ενός χρήστη, τότε η εφαρμογή του συγκεκριμένου χρήστη μεταφορτώνει στον κεντρικό διακομιστή όλα τα CTD των καταγεγραμμένων στην συσκευή του επαφών. Το υπολογιστικό σύστημα του κεντρικού φορέα, είναι σε θέση να διακρίνει για κάθε χρονικό διάστημα t το αντίστοιχο μυστικό κλειδί BK_t που χρησιμοποιήθηκε για την κρυπτογράφηση των PUID και την παραγωγή των EBID, συνεπώς είναι σε θέση για κάθε καταγεγραμμένη επαφή να ανακτήσει το αναγνωριστικό ταυτοποίησης ως το

$$PUID' = AES^{-1}(BK_t, EBID)$$

αποκρυπτογραφώντας το EBID. Στην συνέχεια εκτιμά την επικινδυνότητα της κάθε επαφής, και αποστέλλει μαζικά μια επείγουσα ειδοποίηση (push notification) ενημερώνοντας τους αντίστοιχους χρήστες.

Στο συγκεκριμένο πρωτόκολλο, το υποκείμενο των δεδομένων είναι ο χρήστης, αναγνωριστικό ταυτοποίησης θεωρείται το PUID[2], ψευδώνυμο το EBID και μυστικό ψευδωνυμοποίησης το BK_t . Ένας κεντρικός φορέας όπως για παράδειγμα οι υγειονομικές αρχές μιας χώρας, θεωρείται ο Υπεύθυνος Επεξεργασίας και συγχρόνως ο φορέας ψευδωνυμοποίησης. Ο κεντρικός φορέας είναι απαραίτητο

[1] Μπορούμε να παρατηρήσουμε εδώ, ότι αυτό το σχήμα απαιτεί τακτικές επικοινωνίες δεδομένων μεταξύ του κεντρικού διακομιστή και του τηλεφώνου του χρήστη, και η χρήση αυτής της εφαρμογής δεν θα ήταν ρεαλιστική παλαιότερα όταν οι συνδέσεις δεδομένων κόστιζαν περισσότερο και η χρήση τους ήταν πιο σπάνια. Στο PEPP-PT παρ' ότι οι συνδέσεις δεδομένων στην εποχή μας είναι κάτι που διαθέτουν πολλοί χρήστες, δεν είναι δεδομένο ότι θα έχουν συνεχώς δυνατότητα σύνδεσης δεδομένων, οπότε ο κεντρικός διακομιστής πρακτικά θα παράγει για κάθε χρήστη έναν αριθμό από εφήμερα αναγνωριστικά EBIDt για ένα μεγαλύτερο χρονικό διάστημα (π.χ. για δύο ημέρες) και τα αποστέλλει στην αρχή του χρονικού διαστήματος

[2] σύμφωνα με τον ορισμό του άρθρου 4 του Γενικού Κανονισμού Προστασίας Δεδομένων: «δεδομένα προσωπικού χαρακτήρα»: κάθε πληροφορία που αφορά ταυτοποιημένο ή ταυτοποιήσιμο φυσικό πρόσωπο («υποκείμενο των δεδομένων»)· το ταυτοποιήσιμο φυσικό πρόσωπο είναι εκείνο του οποίου η ταυτότητα μπορεί να εξακριβωθεί, άμεσα ή έμμεσα, ιδίως μέσω αναφοράς σε αναγνωριστικό στοιχείο ταυτότητας, όπως όνομα, σε αριθμό ταυτότητας, σε δεδομένα θέσης, σε επιγραμμικό αναγνωριστικό ταυτότητας ή σε έναν ή περισσότερους παράγοντες που προσιδιάζουν στη σωματική, φυσιολογική, γενετική, ψυχολογική, οικονομική, πολιτιστική ή κοινωνική ταυτότητα του εν λόγω φυσικού προσώπου.

να έχει την δυνατότητα ανάκτησης του αναγνωριστικού ταυτοποίησης από το ψευδώνυμο ενός χρήστη, όμως δεν πρέπει να έχουν αυτήν την δυνατότητα τρίτοι όπως οι υπόλοιποι χρήστες της εφαρμογής.

Το PEPP-PT είναι ένα συγκεντρωτικό σύστημα, στο οποίο ο κεντρικός φορέας θεωρείται έμπιστη οντότητα που προστατεύει τα προσωπικά δεδομένα των υποκειμένων και τα επεξεργάζεται ασφαλώς και σύμφωνα με τους δηλωμένους σκοπούς της επεξεργασίας. Αυτό είναι το σημαντικότερο χαρακτηριστικό για το οποίο διατυπώθηκαν αρνητικές κριτικές για το πρωτόκολλο αυτό, καθώς από τον αρχικό σχεδιασμό του δεν λαμβάνεται υπόψη η κατάχρηση από την πλευρά του κεντρικού φορέα της επεξεργασίας των δεδομένων για σκοπούς που δεν έχουν δηλωθεί, ούτε η ύπαρξη εσωτερικών επιτιθέμενων στην διεργασία της επεξεργασίας.

Πολύ σύντομα μετά την παρουσίαση του PEPP-PT, η ομάδα DP-3T διατύπωσε μια σειρά από προβληματισμούς [7] σχετικά με την επάρκεια του σχεδιασμού του PEPP-PT για την προστασία των δεδομένων των χρηστών:

- Ο κεντρικός φορέας μπορεί να ανακτήσει το μόνιμο αναγνωριστικό ταυτοποίησης από οποιοδήποτε εφήμερο EBID. Εάν είναι σε θέση να συνδέσει οποιοδήποτε EBID με κάποιο συγκεκριμένο υποκείμενο, είναι δυνατή η παρακολούθηση των κινήσεων του συγκεκριμένου υποκειμένου, είτε από τις επαφές άλλων υποκειμένων που διαγνώστηκαν είτε με κάποιο δίκτυο αισθητήρων Bluetooth, όπως για παράδειγμα κινητά τηλέφωνα με τροποποιημένη εφαρμογή σε κοντινή με το υποκείμενο απόσταση.
- Ο κεντρικός φορέας έχει την δυνατότητα να αποδίδει οποιεσδήποτε επιλεγμένες από αυτόν EBID σε οποιονδήποτε χρήστη χωρίς αυτές να αποτελούν κρυπτογράφηση της PUID με το αντίστοιχο του χρονικού διαστήματος μυστικό κλειδί. Γνωστοποιώντας κάποιο ή όλα τα EBID κάποιου υποκειμένου σε τρίτους, όπως για παράδειγμα οι διωκτικές αρχές, μπορεί να κάνει δυνατή για αυτούς την βραχυχρόνια ή μακροχρόνια παρακολούθησή τους, χωρίς την ανάγκη περαιτέρω επικοινωνίας με τον κεντρικό διακομιστή.
- Σε περίπτωση υποκειμένων με θετική διάγνωση όλες οι επαφές τους των τελευταίων εβδομάδων μεταφορτώνονται στον κεντρικό διακομιστή. Με αυτή την πληροφορία, είναι δυνατόν να χαρτογραφηθούν οι επαφές τους με την κατασκευή κάποιου κοινωνικού γραφήματος (social graph) το οποίο είναι πιθανό να επεκτείνεται συνεχώς, λόγω του ότι κάποιες από τις κοντινές επαφές μολυσμένων χρηστών αναμένεται να έχουν μολυνθεί και με την σειρά τους να μεταφορτώσουν όλες τις τελευταίες επαφές τους στον κεντρικό διακομιστή. Καθώς η μετάδοση του ιού θα επεκτείνεται, οι πληροφορίες που θα περιλαμβάνουν τα τυχόντα παραγόμενα κοινωνικά γραφήματα θα αυξάνονται γεωμετρικά και θα αφορούν και σε υποκείμενα που δεν

έχουν μολυνθεί, όπως είναι οι κοινές κοινωνικές επαφές των μολυσμένων. Η επαρκής χαρτογράφηση των κοινωνικών επαφών ενός υποκειμένου, μπορεί επίσης να επιτρέψει την αναγνώριση και ταυτοποίησή του στο μέλλον, όπως για παράδειγμα με την καταγραφή και ανάλυση των στοιχείων σύνδεσης των τηλεφωνικών του επικοινωνιών.

- Περιγράφεται μέθοδος επίθεσης κατά την οποία κάποιος που ειδοποιείται ότι έχει έρθει σε επαφή με θετικό κρούσμα, μπορεί να αναγνωρίσει ποιος από τις επαφές του είναι μολυσμένος, δημιουργώντας έναν αριθμό διαφορετικών λογαριασμών και χρησιμοποιώντας τον καθένα από αυτούς για ένα συγκεκριμένο χρονικό διάστημα, το ίδιο κάθε μέρα. Παράλληλα θα πρέπει να καταγράφει με ποιους έρχεται σε επαφή και την ακριβή ώρα της επαφής. Έτσι, αν λάβει μήνυμα σε συγκεκριμένο λογαριασμό ότι ήρθε σε επαφή με κρούσμα, ο λογαριασμός αυτός θα αντιστοιχεί στο συγκεκριμένο χρονικό διάστημα στο οποίο χρησιμοποιήθηκε, και ο μολυσμένος μπορεί να ταυτοποιηθεί ελέγχοντας το ιδιωτικό του αρχείο καταγραφών. Η επίθεση αυτή δεν απαιτεί καμία πληροφορία από τον κεντρικό διακομιστή, και οδηγεί στην μη-εξουσιοδοτημένη άρση της ψευδωνυμοποίησης.

Στην ίδια ανάλυση [7], η ομάδα DP-3T αναγνώρισε δύο ακόμα πιθανά σημεία επίθεσης σε ένα συγκεντρωτικό σύστημα όπως το PEPP-PT:

- Κάποιος που δηλώνει μολυσμένος και μεταφορτώνει τις επαφές του μέσω της εφαρμογής στον κεντρικό διακομιστή, μπορεί να εισάγει στην λίστα κάποιο συγκεκριμένο EBID που έχει καταγράψει, αναγκάζοντας το σύστημα να τον ειδοποιήσει με την ψευδή υπόθεση ότι μπορεί να κινδυνεύει
- Λόγω κανονιστικών αλλαγών από τους οργανισμούς που παράγουν τα περισσότερο χρησιμοποιούμενα λειτουργικά συστήματα φορητών συσκευών, η εφαρμογή δεν μπορεί να λειτουργεί σε λειτουργία background, γεγονός που μπορεί να εκθέσει προσωπικά δεδομένα σε κίνδυνο αν κάποιος κακόβουλος αποκτήσει πρόσβαση σε μια ξεκλείδωτη συσκευή

Παρά την κριτική που διατυπώθηκε από ειδικούς στην ασφάλεια, το PEPP-PT αποτέλεσε ένα σημαντικό βήμα της τεχνολογίας που προσέφερε μια λύση καλύτερη από τις μέχρι τότε διαθέσιμες σε ένα πρόβλημα που χρειαζόταν επείγουσα αντιμετώπιση, και ξεκίνησε μια διαδικασία για την ταχύτατη ανάπτυξη συστημάτων ιχνηλάτησης επαφών με προσανατολισμό στην βελτίωση της προστασίας της ταυτότητας και των δεδομένων των χρηστών.

3.2.2 Decentralized Privacy-Preserving Proximity Tracing (DP-3T)

Το Decentralized Privacy-Preserving Proximity Tracing (DP-3T) παρουσιάστηκε από την ομώνυμη ομάδα ακαδημαϊκών λίγες ημέρες μετά το PEPP-PT. Είναι και αυτό ένα ανοιχτό πρωτόκολλο για την ανάπτυξη συστημάτων αυτόματης ιχνηλάτησης επαφών μέσω εφαρμογών για φορητές συσκευές που κάνει χρήση της τεχνολογίας Bluetooth LE, με σημαντικότερη διαφορά τον αποκεντρωμένο σχεδιασμό του, καθώς τα αρχεία καταγραφής των επαφών του χρήστη φυλάσσονται αποκλειστικά στην συσκευή του και δεν μεταφορτώνονται στον κεντρικό διακομιστή. Με αυτό τον αποκεντρωμένο σχεδιασμό, επιχειρείται να ενισχυθεί η προστασία του χρήστη από την κακόβουλη ή την μη-εξουσιοδοτημένη άρση της ψευδωνυμοποίησης.

Στο white paper του DP-3T [8] περιγράφονται δύο εναλλακτικοί σχεδιασμοί του συστήματος, ένας «χαμηλού κόστους» και ενός που επιχειρεί ενίσχυση της προστασίας των προσωπικών δεδομένων με κόστος μικρή επιβάρυνση στον όγκο των δεδομένων που διακινούνται μέσω του δικτύου δεδομένων.

Όπως και στο PEPP-PT, το κινητό του χρήστη εκπέμπει μέσω Bluetooth LE ένα εφήμερο αναγνωριστικό ταυτοποίησης EphID το οποίο αλλάζει σε τακτά χρονικά διαστήματα, με την σημαντική διαφοροποίηση στο ότι το EphID παράγεται από την εφαρμογή του χρήστη και όχι από τον κεντρικό διακομιστή.

Κάθε ημέρα t η εφαρμογή του χρήστη παράγει ένα μυστικό κλειδί SK_t εφαρμόζοντας μια hash function στο μυστικό κλειδί της προηγούμενης ημέρας. Το πρώτο SK_t παράγεται τυχαία κατά την αρχικοποίηση της εφαρμογής. Ακολούθως η εφαρμογή του χρήστη παράγει n=24*60 εφήμερα αναγνωριστικά EphIDi , τα οποία θα εκπέμπει ανά 1 λεπτό, ως εξής:

$$EphID1 || \dots || EphIDn = PRG(PRF(SK_t , \text{“broadcast key”}))$$

όπου PRF είναι μια ψευδο-τυχαία συνάρτηση (π.χ. HMAC-SHA256), το «broadcast key» είναι σταθερή και δημόσια συμβολοσειρά, και το PRG είναι μια κρυπτογραφική συνάρτηση ροής (stream cipher) που παράγει n*16 byte, τα οποία χωρίζουμε σε τμήματα των 16 byte για να παραχθούν τα n εφήμερα αναγνωριστικά EphIDi της ημέρας. Μπορούμε να παρατηρήσουμε εδώ ότι κάποιος μπορεί να υπολογίσει τα EphIDi της ημέρας μόνο αν γνωρίζει την ημέρα t και το αντίστοιχο μυστικό κλειδί SKt της ημέρας t.

Η συσκευή του χρήστη εκπέμπει συνεχώς τα EphID και καταγράφει όλες τις κοντινές επαφές του

χρήστη με τα αντίστοιχα EphID των άλλων χρηστών. Τα αρχεία καταγραφής των επαφών παραμένουν αποθηκευμένα στην συσκευή του χρήστη και δεν μεταφορτώνονται στον κεντρικό διακομιστή.

Στην περίπτωση που ένας χρήστης διαπιστωθεί από τις υγειονομικές αρχές ότι μολύνθηκε, ειδοποιείται έτσι ώστε μέσω της εφαρμογής, να μεταφορτώσει το δικό του μυστικό κλειδί SK_t και την ημέρα t η οποία είναι η πρώτη ημέρα για την οποία θεωρείται από τις υγειονομικές αρχές ότι ήταν μολυσματικός. Ο ρόλος του κεντρικού διακομιστή περιορίζεται στο να διαθέτει στην εφαρμογή των χρηστών τα ζεύγη (SK_t , t) των χρηστών που δήλωσαν ότι έχουν μολυνθεί[3]. Όλη η επεξεργασία γίνεται από την συσκευή του χρήστη, η οποία λαμβάνει τα ζεύγη (SK_t , t) των μολυσμένων χρηστών από τα οποία παράγει τα αντίστοιχα EphID για τις αντίστοιχες ημέρες. Ακολούθως τις αναζητά στα δικά του αρχεία καταγραφής επαφών και αν διαπιστώσει ότι έχει έρθει σε επαφή υπολογίζει τον αντίστοιχο κίνδυνο.

Όπως αναφέρθηκε, στο ίδιο έγγραφο [8] περιγράφεται και ένας εναλλακτικός σχεδιασμός, στον οποίο τα μυστικά κλειδιά των μολυσμένων χρηστών δεν μεταφορτώνονται στον κεντρικό διακομιστή, αλλά τα EphID περνούν από μία συνάρτηση hash και αποθηκεύονται σε ένα φίλτρο (Cuckoo filter) το οποίο διανέμεται στις εφαρμογές των χρηστών, επιτρέποντάς τους να διαπιστώνουν κοντινή επαφή με κρούσμα χωρίς να εκτίθεται σε τρίτους κανένα ψευδώνυμο των μολυσμένων χρηστών.

Και στους δύο εναλλακτικούς σχεδιασμούς η εφαρμογή του χρήστη αναλαμβάνει τον ρόλο της αρχής ψευδωνυμοποίησης, παράγει και αποθηκεύει το μυστικό κλειδί SK_t και τα εφήμερα αναγνωριστικά ταυτοποίησης χωρίς να τα μεταδίδει στον κεντρικό διακομιστή. Ο αποκεντρωμένος σχεδιασμός στοχεύει στην τοπική αποθήκευση και επεξεργασία προσωπικών δεδομένων στην συσκευή του χρήστη, περιορίζοντας σημαντικά την εμπλοκή του κεντρικού διακομιστή, και του όγκου των προσωπικών δεδομένων στα οποία δύναται να έχει πρόσβαση ο κεντρικός φορέας. Αυτό έχει σαν αποτέλεσμα κάποια επιβάρυνση των απαιτήσεων για υπολογιστική ισχύ και αποθηκευτικό χώρο της συσκευής του χρήστη αλλά σημαντικά βελτιωμένη προστασία των δεδομένων σε περίπτωση μη-εξουσιοδοτημένης χρήσης των δεδομένων ή παρουσίας επιτιθέμενου στον κεντρικό φορέα.

Η αποκεντρωμένη αρχιτεκτονική, υποστηρίχθηκε με μια κοινή δήλωση [9] υποστήριξης των τεχνολογιών που στοχεύουν στην ενίσχυση της προστασίας των προσωπικών δεδομένων από εκατοντάδες ακαδημαϊκούς. Ο Serge Vaudenay [10] υπέδειξε έναν αριθμό τεχνικών επίθεσης στο DP-3T καταδεικνύοντας κάποιες ατέλειές του, συμπεραίνοντας ότι η αποκεντρωμένη αρχιτεκτονική δημιουργεί περισσότερα προβλήματα στην προστασία των προσωπικών δεδομένων από όσα λύνει. Αργότερα [11],

[3] Όταν ένας χρήστης δηλώσει την μόλυνση και μεταφορτώσει το μυστικό κλειδί SK_t στον κεντρικό διακομιστή, η εφαρμογή του αρχικοποιείται και παράγει ένα νέο τυχαίο μυστικό κλειδί, καθώς το τρέχον μυστικό κλειδί κοινοποιείται στους υπόλοιπους χρήστες, οι οποίοι μπορούν πλέον να υπολογίσουν όλα τα επόμενα μυστικά κλειδιά

καταδεικνύει τις αδυναμίες των πρωτοκόλλων αποκεντρωμένης αρχιτεκτονικής γενικότερα, συμπεραίνοντας ότι ούτε η αποκεντρωμένη ούτε η συγκεντρωτική αρχιτεκτονική χαρακτηρίζονται από επαρκή προστασία των προσωπικών δεδομένων. Η συγκεντρωτική αρχιτεκτονική φαίνεται να παρουσιάζει κινδύνους στην περίπτωση που ο επιτιθέμενος είναι ο κεντρικός φορέας, ενώ η αποκεντρωμένη αρχιτεκτονική παρουσιάζει κινδύνους όταν ο επιτιθέμενος είναι κάποιος τρίτος.

Υπό το πρίσμα της αποτελεσματικής ψευδωνυμοποίησης για την προσφορά υπηρεσιών στους πολίτες από έναν κεντρικό φορέα, ο οποίος στις περισσότερες εφαρμογές θα είναι κρατικός, η αποκεντρωμένη αρχιτεκτονική συνεχίζει να παρουσιάζει πλεονεκτήματα, καθώς ικανοποιεί πληρέστερα τα κριτήρια επιλογής τεχνικών και οργανωτικών μέτρων που καθορίζει η παράγραφος 2 του άρθρου 25 του ΓΚΠΔ, όπως αυτά αναφέρθηκαν στην παρ. 1.1:

- Κριτήριο 1 Ελαχιστοποίηση των προσωπικών δεδομένων: στην αποκεντρωμένη αρχιτεκτονική, τα ψευδώνυμα καθώς και το μυστικό ψευδωνυμοποίησης αποθηκεύονται αποκλειστικά στην συσκευή του χρήστη, εάν αυτός δεν δηλώσει ότι μολύνθηκε. Στην συγκεντρωτική αρχιτεκτονική, τα ψευδώνυμα και το μυστικό ψευδωνυμοποίησης αποθηκεύονται στον κεντρικό διακομιστή, συνεπώς ο κεντρικός φορέας έχει την ευθύνη της προστασίας και ορθής χρήσης τους.
- Κριτήριο 2 Ελαχιστοποίηση της επεξεργασίας προσωπικών δεδομένων: στην αποκεντρωμένη αρχιτεκτονική, τα ψευδώνυμα καθώς και το μυστικό ψευδωνυμοποίησης επεξεργάζονται στην συσκευή του χρήστη. Στην συγκεντρωτική αρχιτεκτονική, τα ψευδώνυμα και το μυστικό ψευδωνυμοποίησης επεξεργάζονται στον κεντρικό διακομιστή, συνεπώς ο κεντρικός φορέας έχει την ευθύνη της ασφαλούς και εξουσιοδοτημένης επεξεργασίας τους.
- Κριτήριο 3 Ελαχιστοποίηση του χρόνου αποθήκευσης προσωπικών δεδομένων : στην αποκεντρωμένη αρχιτεκτονική, στις περισσότερες περιπτώσεις, ο χρόνος αποθήκευσης προσωπικών δεδομένων από τον κεντρικό φορέα είναι μηδενικός, ενώ στην συγκεντρωτική αρχιτεκτονική είναι περιορισμένος αλλά όχι μηδενικός.
- Κριτήριο 4 Ελάχιστη προσβασιμότητα στα προσωπικά δεδομένα : στην αποκεντρωμένη αρχιτεκτονική, κατάλληλα τεχνικά μέτρα μπορούν να ληφθούν για την προστασία της συσκευής του χρήστη και της εφαρμογής που ο ίδιος εγκαθιστά προαιρετικά. Στην συγκεντρωτική αρχιτεκτονική, η προστασία έναντι της μη-εξουσιοδοτημένης χρήσης είναι ευθύνη του κεντρικού φορέα.

Ένας φορέας που επιλέγει μια συγκεντρωτική αρχιτεκτονική για μια εφαρμογή ιχνηλάτησης επαφών, αναλαμβάνει μεγαλύτερη άμεση ευθύνη για την προστασία των δεδομένων, την εξουσιοδοτημένη

επεξεργασία τους και τον έλεγχο πρόσβασης σε αυτά. Η συγκεντρωτική αποθήκευση και επεξεργασία δεδομένων σε έναν κεντρικό διακομιστή διευκολύνει πολύ την λήψη αυστηρών τεχνικών και διοικητικών μέτρων προστασίας, σε σύγκριση με την λήψη μέτρων προστασίας των εφαρμογών και των συσκευών των χρηστών. Όμως ο πιθανός αντίκτυπος κάποιας παραβίασης της ασφάλειας των δεδομένων είναι συγκριτικά πολύ μεγαλύτερος στην συγκεντρωτική αρχιτεκτονική, καθώς ο επιτιθέμενος είναι δυνατόν να αποκτήσει πρόσβαση σε δεδομένα μεγάλου αριθμού υποκειμένων, ενώ στην αποκεντρωμένη αρχιτεκτονική η παραβίαση της ασφάλειας θα πρέπει να πραγματοποιηθεί για κάθε διακριτό υποκείμενο.

Όμοια αρχιτεκτονική με το DP-3T έχει και το πρωτόκολλο Exposure Notification (GAEN)[8] που αναπτύχθηκε από τις εταιρείες Google και Apple όπως και τα περισσότερα από τα αποκεντρωμένα συστήματα που έχουν προταθεί. Η ενσωμάτωση των λειτουργιών του GAEN στο λειτουργικό σύστημα του κινητού τηλεφώνου, κάνει την ανάπτυξη εφαρμογών από τους κρατικούς οργανισμούς πιο εύκολη.

3.2.3 Pronto-C2 Fully Decentralized Automatic Contact Tracing System

Τον Φεβρουάριο του 2021 οι Avitabile, Bottay, Iovinoz και Visconti δημοσίευσαν [4] μια πρόταση για το Pronto-C2, ένα πλήρως αποκεντρωμένο σύστημα αυτόματης ιχνηλάτησης επαφών, σε μια προσπάθεια να μειώσουν τις ευπάθειες που είχαν αναγνωριστεί στα συστήματα αποκεντρωμένης αρχιτεκτονικής που είχαν προταθεί μέχρι τότε. Η σημαντικότερη αδυναμία που αναγνώρισαν, ήταν ότι σε περίπτωσης αναφοράς μόλυνσης από κάποιον χρήστη, η εφαρμογή του μεταφορτώνει τα δικά του εφήμερα ψευδώνυμα, καθώς και τα εφήμερα ψευδώνυμα των χρηστών που ήρθαν σε κοντινή επαφή μαζί τους κατά την διάρκεια μιας χρονικής περιόδου στον κεντρικό διακομιστή. Προσπάθησαν να περιγράψουν ένα σύστημα στο οποίο οι μολυσμένοι χρήστες θα μπορούν ανώνυμα να επικοινωνήσουν μέσω του κεντρικού διακομιστή με τους χρήστες που ήρθαν σε κοντινή επαφή μαζί τους.

Το πρωτόκολλο ανταλλαγής κλειδιών Diffie – Helman [12] επιτρέπει σε δύο μέρη να συμφωνήσουν σε ένα κοινό κλειδί κρυπτογράφησης K, συνεισφέροντας στον υπολογισμό του από έναν μυστικό αριθμό το καθένα και είναι δυνατόν να πραγματοποιηθεί με την ανταλλαγή δύο μόνο μηνυμάτων.

Μετά την ανταλλαγή των μηνυμάτων, τα δύο μέρη υπολογίζουν το ίδιο K χωρίς κανένα από τα δύο να γνωρίζει τον μυστικό αριθμό του άλλου και χωρίς κάποιος που θα υποκλέψει τα μηνύματα να είναι σε θέση να υπολογίσει το K. Στην περίπτωση κοντινής επαφής δύο χρηστών, οι εφαρμογές τους μπορούν με δύο μηνύματα μέσω Bluetooth LE και με την χρήση του αλγορίθμου Diffie – Helman να υπολογίσουν έναν κοινό αριθμό K.

Ο αριθμός K είναι ένα ομαδικό ψευδώνυμο (group pseudonym), κοινό και για τους δύο χρήστες που δεν ταυτοποιεί κανέναν από τους δύο. Αν θεωρήσουμε την κοντινή επαφή και την ανταλλαγή μηνυμάτων μεταξύ των φορητών συσκευών σαν μια συναλλαγή (transaction), το K είναι ψευδώνυμο της συγκεκριμένης συναλλαγής (επαφής) και όχι κάποιου από τους δύο χρήστες, λαμβάνοντας υπόψη ότι σε κάθε νέα συνάντηση των ίδιων χρηστών παράγεται ένα νέο K. Η εφαρμογή του χρήστη αποθηκεύει όλα τα K που παράγονται κατά την διάρκεια των κοντινών επαφών τους με άλλους χρήστες, μαζί με τα δεδομένα που είναι απαραίτητα για τον υπολογισμό του κινδύνου της συγκεκριμένης επαφής (ισχύς σήματος κ.λ.π). Στο συγκεκριμένο πρωτόκολλο δεν απαιτείται ανάκτηση κάποιου αναγνωριστικού ταυτοποίησης, συνεπώς η ταυτότητα των χρηστών δεν απαιτείται να είναι συνδέσιμη (linkable) με το ψευδώνυμο της συναλλαγής.

Σε περίπτωση αναφοράς μόλυνσης, ο χρήστης θα πρέπει ανώνυμα να μεταφορτώσει σε κάποιον κεντρικό διακομιστή τα ψευδώνυμα K των συναλλαγών (κοντινών επαφών) στις οποίες συμμετείχε. Για να το κάνει αυτό, ένας χρήστης, θα πρέπει να είναι εξουσιοδοτημένος από τις υγειονομικές αρχές, πράγμα που σημαίνει ότι θα πρέπει να έχει πιστοποιηθεί η ταυτότητά του και το αποτέλεσμα του διαγνωστικού ελέγχου. Δεν είναι προφανής κάποιος τρόπος για να ολοκληρωθεί μια τέτοια διαδικασία διατηρώντας συγχρόνως την ανωνυμία του χρήστη. Το Pronto-C2 υποδεικνύει την χρήση ενός τυχαίου αριθμού ενεργοποίησης:

- ο κεντρικός φορέας μοιράζει στα διαγνωστικά εργαστήρια έναν μεγάλο αριθμό από τυχαίους κωδικούς ενεργοποίησης
- μετά από θετική διάγνωση, το εργαστήριο δίνει στον χρήστη έναν τυχαίο αριθμό ενεργοποίησης[4]
- ο χρήστης χρησιμοποιεί τον κωδικό για την ενεργοποίηση της διαδικασίας μεταφόρτωσης των ψευδωνύμων K των επαφών του σε έναν κεντρικό διακομιστή

[4] Μπορούμε να παρατηρήσουμε εδώ, ότι η εξασφάλιση απόλυτης ανωνυμίας δεν είναι δυνατή, καθώς είναι αναγκαία η ταυτοποίηση του χρήστη από το διαγνωστικό εργαστήριο για την εξουσιοδότησή του να δηλώσει θετική διάγνωση. Ο χρήστης πρέπει να διαβεβαιώνεται ότι το εργαστήριο θα του αποδώσει έναν τυχαίο κωδικό χωρίς αυτός να είναι δυνατόν να συνδεθεί με την ταυτότητά του, και κατάλληλα διοικητικά μέτρα πρέπει να λαμβάνονται. Σε αυτή την περίπτωση και μόνο για την διαδικασία αναφοράς μόλυνσης, μπορούμε να θεωρήσουμε ότι η αρχή ψευδωνυμοποίησης είναι το εργαστήριο

Η εφαρμογή του χρήστη ελέγχει τακτικά την βάση δεδομένων που τηρείται στον κεντρικό διακομιστή και συγκρίνει τα ψευδώνυμα Κ που έχουν μεταφορτώσει οι μολυσμένοι χρήστες, και τα συγκρίνει με τα Κ των δικών του επαφών διαπιστώνοντας αν συμμετείχε σε κάποια επαφή από αυτούς. Θεωρείται ότι ο κεντρικός διακομιστής σε αυτή την περίπτωση είναι απλώς ένας διάμεσος που διευκολύνει την επικοινωνία μεταξύ των χρηστών, καθώς μόνο οι ίδιοι γνωρίζουν σε ποιες επαφές συμμετείχαν, και δεν μεταφορτώνονται ψευδώνυμα που μπορούν να συνδεθούν με συγκεκριμένους χρήστες.

Ο σχεδιασμός του Pronto-C2 περιλαμβάνει μια σειρά από τεχνολογικά μέτρα για την προστασία έναντι των τύπων επιθέσεων που έχουν καταγραφεί για τα υπόλοιπα αποκεντρωμένα συστήματα, αλλά και για την βελτίωση της απόδοσης του συστήματος. Μέχρι τώρα, δεν υπάρχει εμπειρία κάποιας υλοποίησης του συγκεκριμένου πρωτοκόλλου, που να οδηγήσει σε συμπεράσματα για την ασφάλεια και την απόδοσή του

4 Σενάρια χρήσης σε εφαρμογές Εκπαίδευσης

4.1 Εκπαίδευση μέσω τηλεδιάσκεψης

Η πρόσφατη πανδημία δημιούργησε σε πολλές χώρες την ανάγκη περιορισμού των κοινωνικών επαφών και συνεπώς την ευρύτερη εφαρμογή της τηλεκπαίδευσης σε όλες τις βαθμίδες της εκπαίδευσης, είτε πρόκειται για την υποχρεωτική εκπαίδευση, τις σπουδές στη τριτοβάθμια, την εκπαίδευση και την επαγγελματική κατάρτιση ενηλίκων. Σε πολλές περιπτώσεις, η σύγχρονη εξ' αποστάσεως εκπαίδευση με την χρήση τηλεδιάσκεψης ήταν η πλέον κατάλληλη επιλογή, καθώς είναι περισσότερο διαδραστική από την ασύγχρονη εξ' αποστάσεως εκπαίδευση, άρα και περισσότερο κατάλληλη στην υποκατάσταση της εκπαίδευσης με φυσική παρουσία. Η επιτυχημένη εφαρμογή, υποδεικνύει ότι είναι πιθανόν στο μέλλον να επιλεγούν όμοιες τακτικές σε περίπτωση κάθε είδους εκτάκτων καταστάσεων που καθιστούν την φυσική παρουσία ή τη μετακίνηση δύσκολη σε μεγάλες πληθυσμιακές ομάδες.

Η εμπειρία έδειξε ότι σε πολλές χώρες η διείσδυση των νέων τεχνολογιών και της χρήσης δημοσίων δικτύων ήταν επαρκής για την προσωρινή, αποκλειστικά εξ' αποστάσεως, εκπαίδευση σε πολύ μεγάλη έκταση χωρίς σημαντικά τεχνολογικά προβλήματα, πράγμα που δεν θα ήταν τόσο εύκολο σε προηγούμενες δεκαετίες. Σε πολλές περιπτώσεις οι χρήστες χρησιμοποίησαν ιδιόκτητο τεχνολογικό εξοπλισμό που ήδη διέθεταν, υπάρχουσες δικτυακές υποδομές και κυρίως υπάρχον αξιόπιστο λογισμικό τηλεδιάσκεψης.

Η επιλογή αυτή θα είναι σε πολλές περιπτώσεις αναγκαστική, καθώς σε έκτακτες καταστάσεις είναι πιθανόν να μην είναι διαθέσιμος ο απαραίτητος χρόνος ή ο εκπαιδευτικός οργανισμός να μην έχει την οικονομική δυνατότητα και τεχνογνωσία για την ανάπτυξη ιδιόκτητης εφαρμογής και δικτυακής υποδομής. Επιπλέον είναι πολύ πιθανόν να κριθεί ασύμφορη η συντήρηση μιας τέτοιας υποδομής για χρήση σε έκτακτες καταστάσεις, ενώ είναι δυνατή η χρήση εφαρμογών και υποδομών γενικής χρήσης , με σχετικά μικρής έκτασης παραμετροποίηση.

Η εμπειρία αυτή μας βοηθά να αναγνωρίσουμε ως εμπλεκόμενα μέρη

- έναν κεντρικό φορέα που μπορεί να είναι ένας οποιουδήποτε μεγέθους εκπαιδευτικός οργανισμός
- Χρήστες: εκπαιδευτές και εκπαιδευόμενους που αποτελούν ένα τμήμα, με την έννοια μιας ενιαίας, σταθερής σύνθεσης, ομάδας που παρακολουθεί μια σειρά εκπαιδευτικών περιόδων σε ορισμένο χρόνο και ορισμένης διάρκειας, με συγκεκριμένους εκπαιδευτικούς στόχους
- έναν πάροχο της εφαρμογής και της δικτυακής υποδομής

Υποθέτουμε ότι οι εκπαιδευτές και οι εκπαιδευόμενοι που αποτελούν τα υποκείμενα των δεδομένων,

έχουν ήδη μια σχέση εμπιστοσύνης με τον κεντρικό φορέα, στον οποίον έχουν καταχωρίσει προσωπικά δεδομένα και έχουν συναινέσει στην επεξεργασία τους για την εκπλήρωση των στόχων της εκπαιδευτικής διαδικασίας. Σημαντική παρέμβαση αποτελεί η εμπλοκή του παρόχου της τεχνολογίας τηλεδιάσκεψης, η παροχή των υπηρεσιών του οποίου μπορεί να απαιτεί την πρόσβαση σε δεδομένα των υποκειμένων.

4.1.1 Ο εκπαιδευτικός οργανισμός ως φορέας ταυτοποίησης

Η χρήση της υπηρεσίας τηλεδιάσκεψης είναι πιθανόν να προαπαιτεί την σύνδεση με κάποιο όνομα χρήστη και κωδικό πρόσβασης, τα οποία όμως δεν είναι απαραίτητο να αποτελούν αναγνωριστικά ταυτοποίησης του χρήστη. Είναι δυνατόν να αποδοθεί στον εκπαιδευτικό οργανισμό ο ρόλος του φορέα ταυτοποίησης. Ο χρήστης μετά από ταυτοποίηση και σύνδεση σε εφαρμογή του εκπαιδευτικού οργανισμού, μπορεί να εγγράφεται στην υπηρεσία τηλεκπαίδευσης και να αποκτά ένα νέο ξεχωριστό ζεύγος αναγνωριστικών, δηλαδή ένα νέο όνομα χρήστη και έναν νέο κωδικό πρόσβασης, με τα οποία ο εκπαιδευτικός οργανισμός θα δημιουργεί τον λογαριασμό του χρήστη στον πάροχο της υπηρεσίας τηλεδιάσκεψης.

Τα αναγνωριστικά τηλεδιάσκεψης δεν οδηγούν σε ταυτοποίηση του χρήστη από τον πάροχο, αφού αυτός δεν διαθέτει άλλα δεδομένα σχετικά με την ταυτότητα του χρήστη, συνεπώς η μόνη γνώση που έχει είναι το γεγονός ότι τα αναγνωριστικά ανήκουν σε ένα μέλος του εκπαιδευτικού οργανισμού, είτε εκπαιδευόμενο είτε εκπαιδευτή είτε διοικητικό προσωπικό, που έχει δικαίωμα μέσω του εκπαιδευτικού οργανισμού και της συμφωνίας που έχει αυτός συνάψει με τον πάροχο, να συμμετέχει σε τηλεδιασκέψεις. Σύμφωνα με το σχήμα που υποθέσαμε, τα αναγνωριστικά τηλεδιάσκεψης αποτελούν ψευδώνυμα των αντίστοιχων αναγνωριστικών ταυτοποίησης και σύνδεσης του χρήστη στην εφαρμογή του εκπαιδευτικού οργανισμού.

Είναι πιθανώς επιθυμητό, τα ψευδώνυμα αναγνωριστικά τηλεδιάσκεψης να μην προδίδουν την ταυτότητα του χρήστη, για παράδειγμα το όνομα χρήστη προφανώς θα πρέπει να είναι ένα τυχαιοποιημένο όνομα και να μην είναι ούτε να παράγεται από το ονοματεπώνυμο ή άλλο αναγνωριστικό ταυτοποίησης του χρήστη. Σε αυτή την περίπτωση μπορούν αυτά να παράγονται και να αποδίδονται στον χρήστη από τον εκπαιδευτικό οργανισμό, και όχι να επιλέγονται και να καταχωρίζονται από τον ίδιο. Αυτό θα κάνει την φύλαξη και την διαχείριση των αναγνωριστικών τηλεδιάσκεψης από τον χρήστη λίγο πιο δύσκολη, καθώς δεν θα μπορεί να επιλέξει αναγνωριστικά που θα μπορεί να απομνημονεύσει εύκολα, ταυτόχρονα όμως θα αποτρέψει την αποκάλυψη της ταυτότητας του στον πάροχο από την επιλογή αυτών, ενισχύοντας την εμπιστοσύνη στην διαδικασία τηλεκπαίδευσης. Όπως συνήθως συμβαίνει σε τέτοιες εφαρμογές, σε περίπτωση απώλειας είναι

δυνατή η ανάκτηση των αναγνωριστικών τηλεδιάσκεψης από τον χρήστη ή παραγωγής νέων.

Σε αυτό το σχήμα, ο εκπαιδευτικός οργανισμός, αναλαμβάνει το μεγαλύτερο μέρος της διαχείρισης των χρηστών που μπορεί να περιλαμβάνει την δημιουργία λογαριασμών και παραγωγή των αναγνωριστικών ψευδωνύμων, την κατανομή σε τμήματα και την κατάρτιση προγράμματος τηλεκπαίδευσης, τον καθορισμό της πρόσβασης σε τηλεδιασκέψεις, αλλά και την επικοινωνία με τους χρήστες για τεχνική υποστήριξη και θέματα τήρησης των όρων χρήσης της υπηρεσίας τηλεδιάσκεψης. Συνεπώς υπάρχει η απαίτηση της διατήρησης από τον εκπαιδευτικό οργανισμό της δυνατότητας ανάκτησης των αναγνωριστικών από τα ψευδώνυμα.

Τα ψευδώνυμα αναγνωριστικά τηλεδιάσκεψης είναι δυνατόν να παράγονται με την χρήση κάποιας τεχνικής ψευδωνυμοποίησης από αυτές που έχουν ήδη αναφερθεί, είτε από τα ίδια τα αναγνωριστικά σύνδεσης στην εφαρμογή του εκπαιδευτικού οργανισμού, είτε από άλλα διαθέσιμα αναγνωριστικά δεδομένα του χρήστη. Η επιλογή συγκεκριμένης τεχνικής μπορεί να γίνει λαμβάνοντας υπόψιν την απαίτηση της δυνατότητας ανάκτησης (recovery) και της επεκτασιμότητας (scalability) ανάλογα με το μέγεθος του οργανισμού και το πλήθος των εκπαιδευόμενων, καθώς δεν μπορούμε να αναγνωρίσουμε απαιτήσεις χρησιμότητας (utility).

4.1.2 Ανά έγγραφο ντετερμινιστική πολιτική

Η πολιτική ψευδωνυμοποίησης των αναγνωριστικών μπορεί να είναι ντετερμινιστική, εφόσον για κάθε εκπαιδευόμενο εκδίδεται μόνο ένα ζεύγος ψευδώνυμων αναγνωριστικών για την γενικότερη χρήση της υπηρεσίας τηλεδιάσκεψης. Θεωρήσαμε ως «τμήμα» μια ενιαία ομάδα που παρακολουθεί μια σειρά εκπαιδευτικών περιόδων με συγκεκριμένους εκπαιδευτικούς στόχους. Σύμφωνα με αυτό τον ορισμό το τμήμα θα μπορούσε να σημαίνει ένα εκπαιδευτικό έτος, μια τάξη, ένα σεμινάριο, μια ομάδα εκπαιδευτικών αντικειμένων, ένα μάθημα, ένα εργαστήριο ή οποιοσδήποτε άλλος διαχωρισμός επιβάλλει η οργάνωση μιας εκπαιδευτικής διαδικασίας.

Τα ονόματα των τμημάτων και των αντικειμένων των τηλεδιασκέψεων θα ήταν δυνατόν να κωδικοποιηθούν από τον εκπαιδευτικό οργανισμό με μια διαδικασία ψευδωνυμοποίησης όμοια με αυτήν που θα επιλεγεί για την ψευδωνυμοποίηση των αναγνωριστικών των χρηστών. Έτσι ο πάροχος θα έχει πρόσβαση σε δεδομένα όπως ο προγραμματισμός, η χρονική διάρκεια, και ο αριθμός των συμμετεχόντων αλλά δεν θα γνωρίζει την ταυτότητα των συμμετεχόντων ούτε το θέμα της τηλεδιάσκεψης. Ακόμα και σε αυτή την περίπτωση, αν θεωρήσουμε την ύπαρξη εσωτερικού

αντιπάλου στον πάροχο και με συγκεκριμένες προϋποθέσεις, ίσως να είναι δυνατόν να εξάγει συμπεράσματα που θα οδηγήσουν σε μερική ταυτοποίηση κάποιων χρηστών που θα μπορούσαν να οδηγήσουν σε κάποια πολιτική διακρίσεων. Καθώς με την ντετερμινιστική πολιτική ψευδωνυμοποίησης κάθε χρήστης διατηρεί το ίδιο ψευδώνυμο σε μια εκπαιδευτική διαδικασία, μπορούν να εξαχθούν συμπεράσματα όπως για παράδειγμα το ότι μια ομάδα χρηστών ανήκουν σε ένα τμήμα ή ένας συγκεκριμένος χρήστης ανήκει σε κάποια τμήματα. Τέτοιες πληροφορίες σε συνδυασμό με την γνώση της συχνότητας και της διάρκειας των τηλεδιασκέψεων αλλά και με δημόσιες πληροφορίες σχετικά με τα προγράμματα σπουδών που διαθέτει ο οργανισμός μπορούν θεωρητικά να οδηγήσουν σε κάποια μορφή διαχωρισμού των χρηστών από κάποιον επιτιθέμενο.

Στην περίπτωση αναγνώρισης τέτοιων κινδύνων και αυστηρών απαιτήσεων προστασίας δεδομένων, μπορεί να ληφθεί υπόψη μια ανά έγγραφο ντετερμινιστική πολιτική (document deterministic policy).

Αυτή η κατηγορία πολιτικών αν και δεν περιλαμβάνεται στην κατηγοριοποίηση του ENISA όπως την είδαμε στην παράγραφο 2.3, χρησιμοποιείται στην βιομηχανία για να καλύψει παρόμοιες ανάγκες. Σε μια τέτοια πολιτική, σε μια βάση δεδομένων, αντιστοιχίζεται σε ένα αναγνωριστικό ψευδώνυμο, το ίδιο σε όλες τις εμφανίσεις του στην ίδια βάση, ενώ σε κάθε άλλη βάση δεδομένων, στο ίδιο αναγνωριστικό αντιστοιχίζεται ένα διαφορετικό για κάθε βάση δεδομένων ψευδώνυμο.

Στην περίπτωση της εκπαίδευσης μέσω τηλεδιάσκεψης, μπορούμε να θεωρήσουμε μια διαφορετική βάση δεδομένων με τους συμμετέχοντες σε κάθε τμήμα. Ο εκπαιδευτικός οργανισμός θα παράγει ένα διαφορετικό ζεύγος ψευδωνύμων των αναγνωριστικών σύνδεσης χρήστη, διαφορετικό για κάθε τμήμα στο οποίο συμμετέχει ένας χρήστης, με αυτό τον τρόπο κάποιος που έχει πρόσβαση στα ψευδώνυμα δεν θα μπορεί να αναγνωρίσει τα τμήματα στα οποία συμμετέχει ένας χρήστης, αφού σε καθένα από αυτά συνδέεται με διαφορετικά αναγνωριστικά. Περιορίζονται έτσι σημαντικά οι πληροφορίες που μπορεί ένας επιτιθέμενος να πάρει από την επεξεργασία των ψευδωνυμοποιημένων δεδομένων σε συνδυασμό με πληροφορίες σχετικές με τις συνδέσεις και τις σχετικές με τα προγράμματα σπουδών δημόσιες πληροφορίες, με την προϋπόθεση βέβαια ότι το πλήθος των τμημάτων και το πλήθος των χρηστών δεν είναι πολύ μικρά.

Η επιλογή μιας τέτοιας πολιτικής ψευδωνυμοποίησης προσθέτει ένα διαχειριστικό κόστος στον οργανισμό, ο οποίος θα πρέπει να παράγει και να διαχειρίζεται περισσότερα του ενός ζεύγη αναγνωριστικών για κάθε χρήστη. Επίσης προσθέτει διαχειριστικό κόστος στον χρήστη, ο οποίος θα πρέπει να φυλάσσει τοπικά και να διαχειρίζεται περισσότερα του ενός αναγνωριστικά τηλεδιάσκεψης, αυτό όμως είναι ένα πρακτικό πρόβλημα που μπορεί να απλοποιηθεί με την κατάλληλη παραμετροποίηση κάποιου λογισμικού σύνδεσης που θα χρησιμοποιεί ο χρήστης, ώστε να διαχειρίζεται με ευκολία τοπικά το πρόγραμμα και τις λεπτομέρειες σύνδεσης των σπουδών του χωρίς αυτές οι πληροφορίες να γνωστοποιούνται στον πάροχο.

4.3 Παρουσιολόγιο σύγχρονης εξ’ αποστάσεως εκπαίδευσης

Σε κάποιες εκπαιδευτικές διαδικασίες, ανάλογα με την βαθμίδα, την νομοθεσία και τον σκοπό της εκπαίδευσης είναι δυνατόν να ζητείται η υποχρεωτική παρακολούθηση από τους εκπαιδευόμενους. Οι περισσότερες εμπορικές εφαρμογές που χρησιμοποιούνται γι’ αυτόν τον σκοπό παρέχουν την δυνατότητα της ενεργοποίησης της εξαγωγής, μετά το τέλος μιας συνεδρίας, αναφοράς που περιλαμβάνει τα αναγνωριστικά των συμμετεχόντων και τον χρόνο σύνδεσης του καθένα στην συγκεκριμένη συνεδρία.

Η δυνατότητα αυτή μπορεί να είναι χρήσιμη, όμως σε κάποιες εκπαιδευτικές διαδικασίες όπως είναι η υποχρεωτική εκπαίδευση ανηλίκων μαθητών, μπορεί αν μην είναι αξιόπιστη, για διάφορους λόγους: η σύνδεση δεν σημαίνει απαραίτητα και παρουσία του εκπαιδευόμενου καθώς μπορεί κάποιος άλλος να διαχειρίζεται να αναγνωριστικά πρόσβασης, η μεγάλη έκταση της εφαρμογής της εξ΄ αποστάσεως εκπαίδευσης δεν επιτρέπει να αποκλείσουμε δυσκολία ή αδυναμία σύνδεσης λόγω τεχνικών προβλημάτων και σε πολλές περιπτώσεις δεν είναι δυνατόν να απαιτηθεί η διάθεση και η ορθή χρήση του εξοπλισμού και της σύνδεσης δεδομένων από τον εκπαιδευόμενο.

Η καταγραφή των απουσιών σε πολλές περιπτώσεις στην δια ζώσης εκπαίδευση γίνεται από τον εκπαιδευτή, ο οποίος είναι σε θέση να διαπιστώνει άμεσα την παρουσία και την συμμετοχή των εκπαιδευόμενων. Το πλεονέκτημα αυτό διατηρείται και στην περίπτωση της τηλεδιάσκεψης καθώς η επικοινωνία είναι αμφίδρομη και η συμμετοχή των εκπαιδευόμενων στην εκπαιδευτική διαδικασία δυνατή. Έτσι είναι δυνατόν ο εκπαιδευτικός να αναλάβει την αξιόπιστη καταγραφή των απουσιών, χωρίς να τίθεται θέμα ψευδωνυμοποίησης και προστασίας των αναγνωριστικών των εκπαιδευόμενων.

Είτε επιλεγεί η ενεργοποίηση της δυνατότητας εξαγωγής αναφοράς σύνδεσης των συμμετεχόντων είτε όχι, ο πάροχος θα στις περισσότερες περιπτώσεις διατηρεί την καταγραφή και την δυνατότητα πρόσβασης στις πληροφορίες σύνδεσης, τόσο για λόγους εξασφάλισης της ποιότητας υπηρεσιών και τεχνικής υποστήριξης των χρηστών. Όπως έχουμε αναφέρει, αυτό μπορεί να οδηγήσει έναν επιτιθέμενο σε συμπεράσματα τόσο για την συμμετοχή του χρήστη στην εκπαιδευτική διαδικασία όσο και σχετικά με την ταυτότητα του χρήστη, αν για παράδειγμα γνωρίζει τον ακριβή χρόνο σύνδεσης ή την διεύθυνση IP κάποιου συγκεκριμένου χρήστη.

Το γεγονός αυτό κάνει περισσότερο χρήσιμη την επιλογή κάποιας ασφαλούς μεθόδου ψευδωνυμοποίησης των αναγνωριστικών, όσο και την ψευδωνυμοποίηση της διεύθυνσης IP του χρήστη που θα εξετάσουμε ακολούθως.

4.2 Ψευδωνυμοποίηση των IP διευθύνσεων

Οι IP (Internet Protocol) διευθύνσεις από τις οποίες συνδέονται οι χρήστες αποτελούν ένα ακόμα αναγνωριστικό για το οποίο θα μπορούσε να υπάρχει απαίτηση ψευδωνυμοποίησης. Η IP διεύθυνση προσδιορίζει το δίκτυο αλλά και τον συγκεκριμένο κόμβο (host) από τον οποίο συνδέεται ένας χρήστης και έχει χαρακτηριστεί ως προσωπικό δεδομένο [13]. Ταυτόχρονα, η IP διεύθυνση του χρήστη είναι απαραίτητη για την παροχή υπηρεσιών τηλεδιάσκεψης από τον πάροχο, την συντήρηση της υπηρεσίας και την ανίχνευση ανωμαλιών στην λειτουργία της με χρήση δεδομένων τηλεμετρίας. Στην περίπτωση που επιλεγεί κάποια διαδικασία ψευδωνυμοποίησης της IP διεύθυνσης, μπορούμε να υποθέσουμε ότι απαιτείται δυνατότητα ανάκτησης του αναγνωριστικού από τον πάροχο, για την διερεύνηση περιστατικών ασφάλειας ή παραβίασης των όρων χρήσης της υπηρεσίας, ή για την εξάσκηση του δικαιώματος πρόσβασης του υποκειμένου των δεδομένων στα προσωπικά δεδομένα του που διατηρεί ο πάροχος. Η ψευδωνυμοποίηση φαίνεται να είναι κατάλληλη πρακτική για την προστασία των δεδομένων διατηρώντας την απαραίτητη χρηστικότητα αλλά και την δυνατότητα ανάκτησης.

Μια IP διεύθυνση του πρωτοκόλλου IPv4 που είναι προς το παρόν το πρωτόκολλο που χρησιμοποιείται περισσότερο συγκροτείται από 32 δυαδικά ψηφία (bit) και χωρίζεται σε δύο μέρη, το πρόθεμα δικτύου (network prefix) που είναι τα πλέον σημαντικά bit και το αναγνωριστικό κόμβου (host identifier), ενώ το μήκος του κάθε μέρους δεν είναι σταθερό και εξαρτάται από τα χαρακτηριστικά του δικτύου.

Θέματα ψευδωνυμοποίησης των IP διευθύνσεων των χρηστών από παρόχους τηλεπικοινωνιακών υπηρεσιών εξετάζονται στο [2]:

- Η χρήση hash function δεν προσφέρει ικανοποιητικό επίπεδο προστασίας, λόγω του μικρού μεγέθους του πεδίου τιμών των αναγνωριστικών που κάνει εφικτές και ταχύτατες τις επιθέσεις εξαντλητικών αναζητήσεων, συμπέρασμα που υποδεικνύει την χρήση κάποιας τεχνικής που κάνει χρήση κάποιου μυστικού ψευδωνυμοποίησης όπως είναι ένα κρυπτογραφικό κλειδί.
- Παρουσιάζεται η δυνατότητα ψευδωνυμοποίησης μόνο του αναγνωριστικού κόμβου αφήνοντας ανέπαφο το πρόθεμα δικτύου. Η διατήρηση μιας τέτοιας χρησιμότητας είναι συχνά απαραίτητη, καθώς δίνει την δυνατότητα στον πάροχο τηλεπικοινωνιακών υπηρεσιών

για ανίχνευση ανωμαλιών στην λειτουργία του δικτύου του, ενώ παράλληλα προστατεύει το αναγνωριστικό του συγκεκριμένου κόμβου σύνδεσης του χρήστη

- Η επιλογή της πολιτικής ψευδωνυμοποίησης (ντετερμινιστική, ανά έγγραφο τυχαιοποιημένη, πλήρως τυχαιοποιημένη) εξαρτάται ως συνήθως από το επιθυμητό επίπεδο προστασίας και την απαραίτητη χρησιμότητα και είναι δυνατόν να επιλεγεί η πολιτική με το ανώτερο επίπεδο προστασίας από αυτές που διατηρούν την ελάχιστη χρησιμότητα που είναι απαραίτητη έτσι ώστε ο πάροχος να είναι σε θέση να προσφέρει την επιθυμητή ποιότητα υπηρεσιών.

Καθώς γίνεται χρήση συγκεκριμένης IP διεύθυνσης για την σύνδεση σε κάποιον εξυπηρετητή του παρόχου υπηρεσίας τηλεδιάσκεψης, είναι προφανές ότι ο ρόλος της οντότητας ψευδωνυμοποίησης θα πρέπει να αποδοθεί σε κάποια οντότητα που αποτελεί μέρος του παρόχου, με την παράλληλη λήψη διοικητικών μέτρων που θα κάνουν δυνατή την ανάκτηση μόνο στις περιπτώσεις που αυτό είναι απαραίτητο.

Μια διαφορετική προσέγγιση ψευδωνυμοποίησης θα ήταν να ανατεθεί στον εκπαιδευτικό οργανισμό ένας ρόλος διάμεσου με την παρεμβολή ενός δικού του proxy εξυπηρετητή. Υπάρχουν διάφοροι τρόποι να υλοποιηθεί κάτι τέτοιο, όπως για παράδειγμα η σύνδεση των χρηστών χωρίς λογισμικό – πελάτη, αλλά η σύνδεση σε κάποια διαδικτυακή εφαρμογή του εκπαιδευτικού οργανισμού. Σε αυτό το παράδειγμα όλες οι απ' ευθείας συνδέσεις προς τον πάροχο θα γίνονται από τον κόμβο του του εκπαιδευτικού οργανισμού, που θα χαρακτηρίζεται από μία μοναδική IP του εκπαιδευτικού οργανισμού, από την οποία ο πάροχος θα βλέπει να προέρχονται. Σε μια τέτοια προσέγγιση θα υπήρχε το πλεονέκτημα της δυνατότητας διαχείρισης των τμημάτων και της τηλεκπαίδευσης από τον εκπαιδευτικό οργανισμό, αλλά και σημαντικά μειονεκτήματα.

Τα μειονεκτήματα σχετίζονται κυρίως με το κόστος, τον χρόνο και την τεχνογνωσία που θα πρέπει να διατεθούν από τον εκπαιδευτικό οργανισμό για να εγκαταστήσει, να συντηρεί και να διαχειρίζεται μια τεχνολογική υποδομή, η οποία θα χρησιμοποιείται στις περιπτώσεις μετάπτωσης σε διαδικασίες τηλεκπαίδευσης, ενώ ανάλογα με τον τρόπο εφαρμογής μπορεί να απαιτεί τροποποίηση του λογισμικού του παρόχου της υπηρεσίας τηλεδιάσκεψης. Κάτι τέτοιο μπορεί να είναι πολύ δύσκολο ή και αδύνατο να γίνει κάτω από την πίεση μιας κατάστασης έκτακτης ανάγκης όπως είναι μια υγειονομική κρίση, θα μπορούσε όμως να είναι πιο ρεαλιστική εάν ο εκπαιδευτικός οργανισμός έχει προετοιμαστεί εκ των προτέρων.

Καθώς μεγάλος αριθμός εκπαιδευόμενων και εκπαιδευτών έχουν ήδη κάνει χρήση τηλεδιάσκεψης για εκπαιδευτικούς σκοπούς ως υποκατάσταση της δια ζώσης εκπαίδευσης και όχι συμπληρωματικά με αυτήν, υπάρχει πιθανότητα κάποιοι από αυτούς να αναγνωρίσουν κάποια πλεονεκτήματα που τους εξυπηρετούν και να προτιμήσουν την τηλεκπαίδευση στο μέλλον. Επίσης από την στιγμή που έγινε

αντιληπτό ότι μια προσωρινή αλλά πλήρης μετάπτωση σε εξ αποστάσεως εκπαίδευση είναι εφικτή, είναι πιθανότερο να επιλεγεί στο μέλλον σε κάποια κατάσταση έκτακτης ανάγκης, ενώ παλαιότερα σε αντίστοιχες καταστάσεις θα είχαμε διακοπή της εκπαιδευτικής διαδικασίας.

Δεν θα πρέπει να θεωρούμε απίθανο κάποιοι εκπαιδευτικοί οργανισμοί να επιλέξουν ασφαλέστερες και μεγαλύτερου κόστους τακτικές για να προσφέρουν υψηλό επίπεδο διασφάλισης των δεδομένων των χρηστών.

5 Συμπεράσματα – Συζήτηση

Προσπαθήσαμε να περιγράψουμε ρεαλιστικά σενάρια βασισμένοι σε πληροφορίες ευρείας δημοσιοποίησης, συνεπώς οι περισσότερες εφαρμογές που περιγράψαμε έχουν εφαρμοστεί κατά την διάρκεια της πανδημίας COVID-19, με διαφορετικές υλοποιήσεις από πολλές χώρες στην Ευρώπη και στον υπόλοιπο κόσμο. Τα κράτη, επέλεξαν να υλοποιήσουν ή και όχι παρόμοιες εφαρμογές ανάλογα με τις τοπικές συνθήκες, την επιδημιολογική παρακολούθηση, τα διαθέσιμα μέσα καθώς και το ισχύον νομοθετικό πλαίσιο.

Οι εφαρμογές αυτές υλοποιήθηκαν ταχύτατα, καλύπτοντας τις έκτακτες ανάγκες που επέβαλλε μια έκτακτη κατάσταση. Η εμφάνιση των αναγκών συνοδεύτηκε από την συνειδητοποίηση της ωριμότητας της τεχνολογίας, της επάρκειας των υποδομών, της καταλληλότητας έτοιμων εφαρμογών και της δυνατότητας των επιστημόνων και των τεχνικών να καλύψουν τα υπάρχοντα κενά για την ταχύτατη εφαρμογή. Η τηλεκπαίδευση σε κάποιες περιπτώσεις υποκατέστησε την δια ζώσης εκπαίδευση, κάνοντας δυνατή την προσωρινή αποφυγή συνωστισμού σε αίθουσες εκπαίδευσης. Τα αυτόματα συστήματα ιχνηλάτησης επαφών διευκόλυναν και επιτάχυναν την υπάρχουσα διαδικασία. Πιθανότατα όλες οι υλοποιήσεις συμμορφώνονται με την υπάρχουσα νομοθεσία, καλύπτοντας και τις απαιτήσεις του ΓΚΠΔ στην Ευρώπη.

Μεγάλος αριθμός πολιτών συμμετείχαν στις έκτακτες διαδικασίες, ενώ εκφράστηκαν και προβληματισμοί σχετικά με την προστασία των προσωπικών δεδομένων. Δεν μπορούμε να αποκλείσουμε κάποιοι να συμμετείχαν, παρά τους προβληματισμούς τους, λόγω του έκτακτου της κατάστασης. Επίσης δεν μπορούμε να αποκλείσουμε την συμβολή της εμπειρίας που αποκτήθηκε στην χρήση όμοιων εφαρμογών ή την ανάπτυξη νέων, σε μελλοντικές όμοιες ή και διαφορετικές συνθήκες.

Η επιλογή κατάλληλων τεχνικών ψευδωνυμοποίησης μπορεί να ενισχύσει την προστασία των προσωπικών δεδομένων των πολιτών. Σε συνδυασμό με την διαφάνεια και την δημοσιοποίηση των τεχνικών λεπτομερειών, μπορεί να προσφέρει στους πολίτες την διαβεβαίωση ότι η συλλογή, αποθήκευση επεξεργασία των προσωπικών δεδομένων γίνεται αποκλειστικά για τους δηλωθέντες σκοπούς, από εξουσιοδοτημένο προσωπικό και με τα αυστηρότερα μέτρα ασφάλειας, αυξάνοντας την συμμετοχή τους και συνεπώς την αποτελεσματικότητα των εφαρμογών. Η πρόσφατη εμπειρία μπορεί να βοηθήσει στην επιλογή ασφαλέστερων τεχνικών ψευδωνυμοποίησης

6 Βιβλιογραφία

[1] ENISA, “Recommendations on shaping technology according to GDPR provisions,” 2018. doi: 10.2824/518496.

[2] ENISA, “Pseudonymisation techniques and best practices,” 2019. doi: 10.2824/247711.

[3] ENISA, “DATA PSEUDONYMISATION: ADVANCED TECHNIQUES & USE CASES,” 2021, doi: 10.2824/860099.

[4] G. Avitabile, V. Botta, V. Iovino, and I. Visconti, “Towards Defeating Mass Surveillance and SARS-CoV-2: The Pronto-C2 Fully Decentralized Automatic Contact Tracing System,” *Cryptol. ePrint Arch.*, no. Report 2020/493, pp. 1–29, 2020, Accessed: May 16, 2021. [Online]. Available: https://eprint.iacr.org/2020/493.

[5] European Commission, “How tracing and warning apps can help during the pandemic | European Commission,” 2021. https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/travel-during-coronavirus-pandemic/how-tracing-and-warning-apps-can-help-during-pandemic_en (accessed Oct. 28, 2021).

[6] PEPP-PT/PEPP, “Data Protection and Information Security Architecture Illustrated on German Implementation,” no. April, pp. 1–26, 2020, [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation/blob/master/10-data-protection/PEPP-PT-data-protection-information-security-architecture-Germany.pdf.

[7] The Dp-3t Project, “Security and privacy analysis of the document ‘ PEPP-PT : Data Protection and Information Security Architecture ,’” no. April, pp. 1–5, 2020, [Online]. Available: https://github.com/DP-3T/documents/blob/master/Security analysis/PEPP-PT_ Data Protection Architechture - Security and privacy analysis.pdf.

[8] P. C. Troncoso *et al.*, “Decentralized Privacy-Preserving Proximity Tracing University of Torino / ISI Foundation ,” no. April, p. 33, 2020, [Online]. Available: https://github.com/DP-3T/documents/blob/master/DP3T White Paper.pdf.

[9] “Contact Tracing Joint Statement.” https://www.esat.kuleuven.be/cosic/sites/contact-tracing-joint-statement/ (accessed Jan. 11, 2022).

[10] S. Vaudenay, “Analysis of DP3T Between Scylla and Charybdis,” 2020. Accessed: Dec. 29, 2021. [Online]. Available: https://eprint.iacr.org/2020/399.

[11] S. Vaudenay, “Centralized or Decentralized? The Contact Tracing Dilemma,” *Cryptol. ePrint Arch.*, no. Report 2020/531, 2020, Accessed: Jan. 10, 2022. [Online]. Available: https://eprint.iacr.org/2020/531.

[12] W. Diffie, W. Diffie, and M. E. Hellman, “New Directions in Cryptography,” *IEEE Trans. Inf. Theory*, vol. 22, no. 6, 1976, doi: 10.1109/TIT.1976.1055638.

[13] Article 29 Data Protection Working Party, “Article 29 Working Party Opinion 4/2007 in the concept of personal data,” vol. 136, no. Lx, pp. 1–26, 2007, [Online]. Available: ec.europa.eu.