234 – Ασφάλεια & Ιδιωτικότητα στο Διαδίκτυο του Μέλλοντος

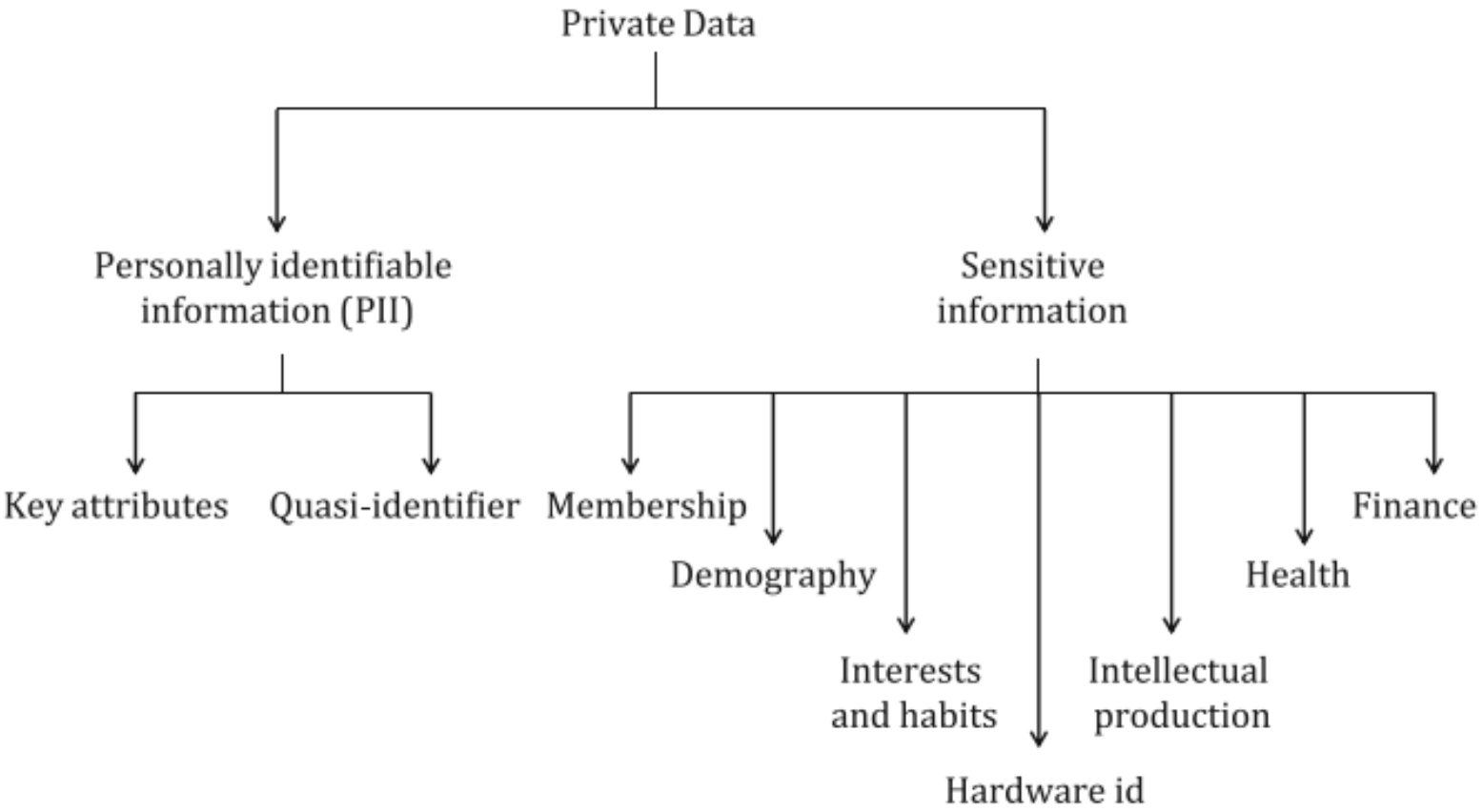
Ιδιωτικότητα στο Υπολογιστικό Νέφος

Κυριάκος Κρητικός
Αναπλ. Καθηγητής
Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων

Περίγραμμα

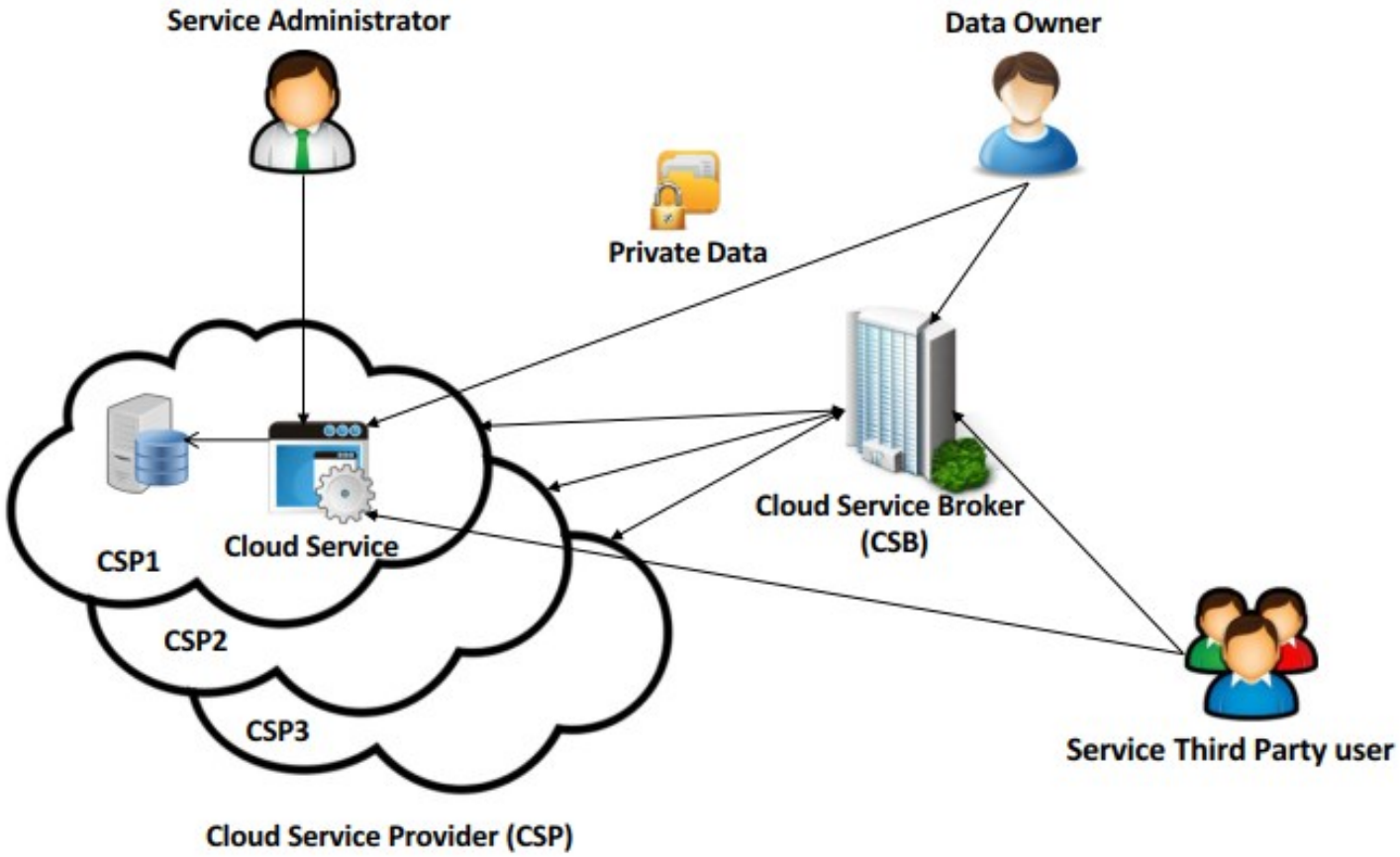
- Είδη Δεδομένων
- Γενική Αρχιτεκτονική
- Εννοιολογικό Μοντέλο
- Κατηγοριοποίηση & Ανάλυση Μέτρων

Ιδιωτικά Δεδομένα



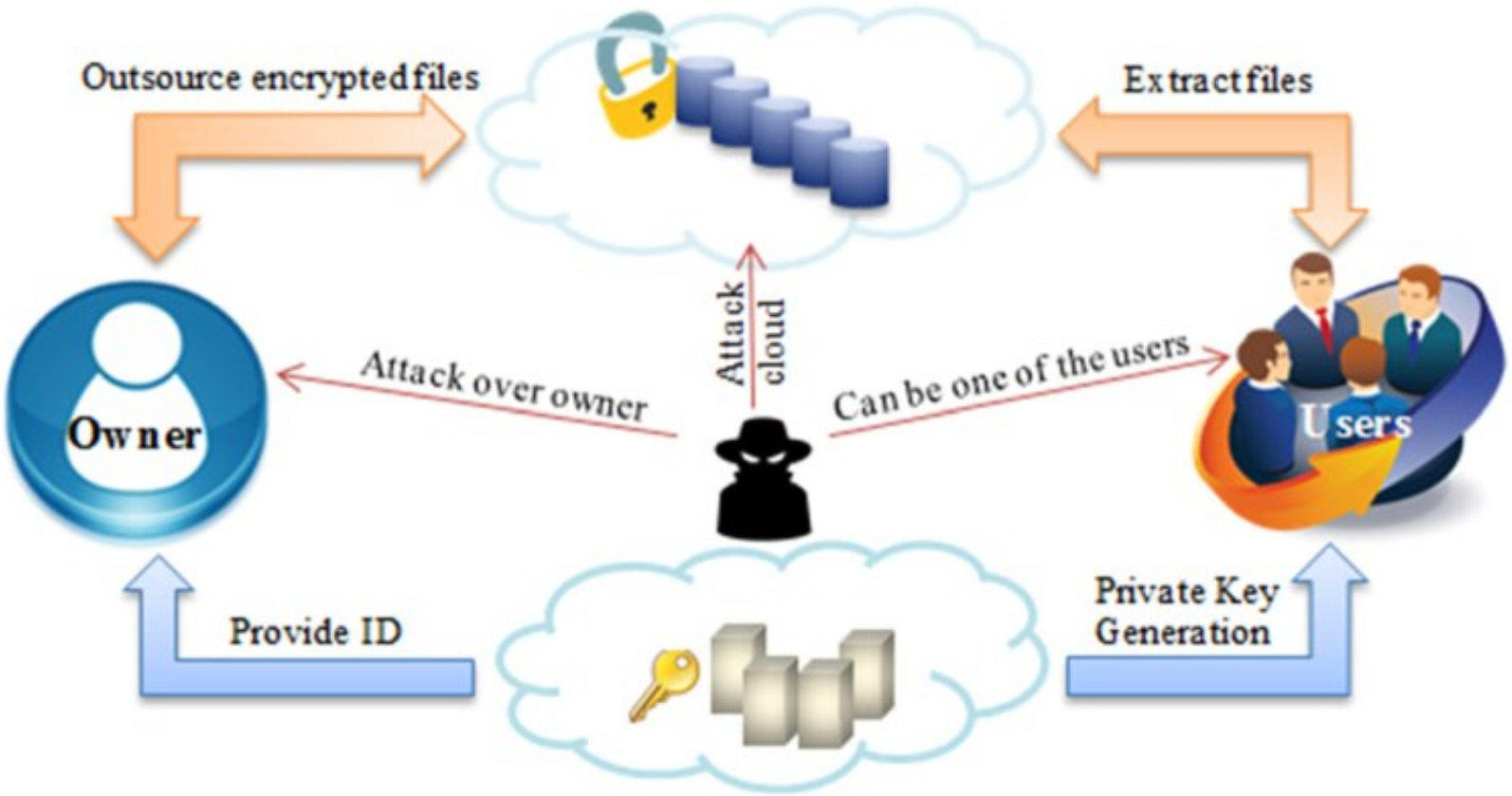
Πηγή: [1]

Γενική Αρχιτεκτονική



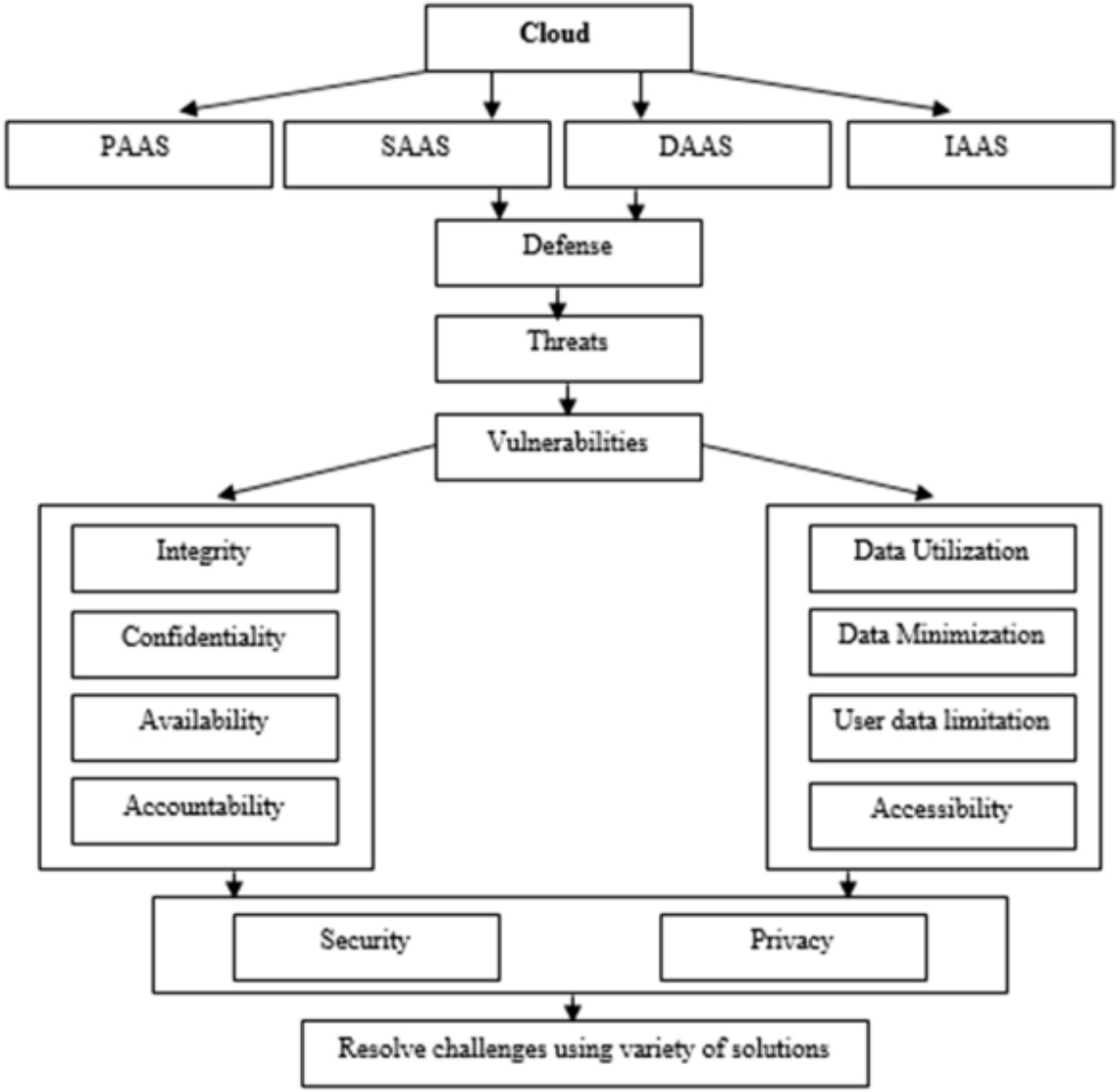
Πηγή: [1]

Περιοχές Επίθεσης

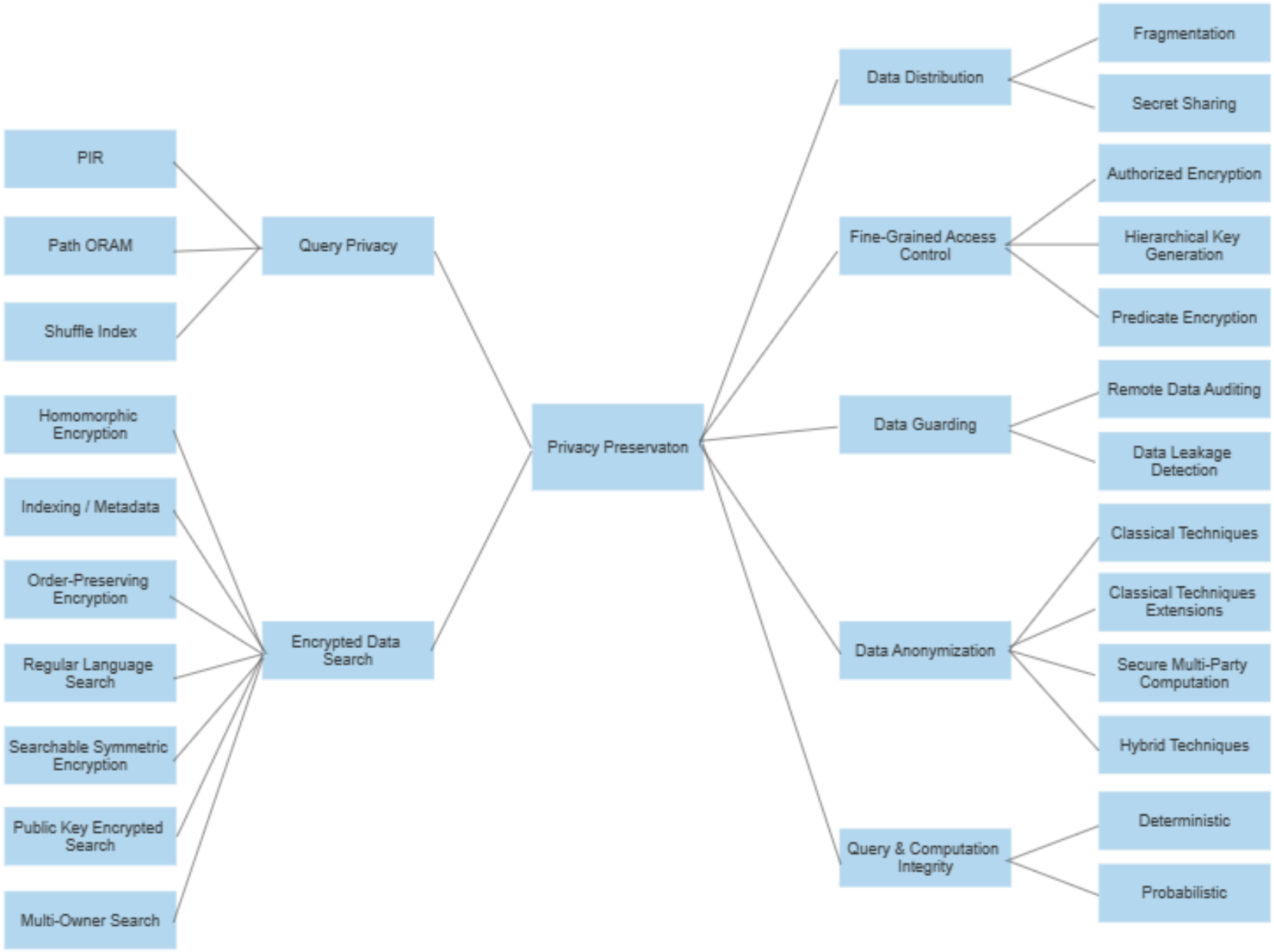


Πηγή: [2]

Εννοιολογικό Μοντέλο



Πηγή: [2]


PIR
Path ORAM
Shuffle Index
Query Privacy
Homomorphic Encryption
Indexing / Metadata
Order-Preserving Encryption
Regular Language Search
Searchable Symmetric Encryption
Public Key Encrypted Search
Multi-Owner Search
Encrypted Data Search
Privacy Preservaton
Data Distribution
Fragmentation
Secret Sharing
Fine-Grained Access Control
Authorized Encryption
Hierarchical Key Generation
Predicate Encryption
Data Guarding
Remote Data Auditing
Data Leakage Detection
Data Anonymization
Classical Techniques
Classical Techniques Extensions
Secure Multi-Party Computation
Hybrid Techniques
Query & Computation Integrity
Deterministic
Probabilistic

Τμηματοποίηση (Fragmentation)

- Επιτρέπει την τμηματοποίηση των δεδομένων και την αποθήκευσή τους σε διαφορετικές τοποθεσίες
  - Η τμηματοποίηση πρέπει να πραγματοποιηθεί με κατάλληλο τρόπο ώστε να προστατευθούν οι συσχετίσεις
    - Θα πρέπει να ληφθεί υπόψιν η περίπτωση του μη έμπιστου ή ημι-έμπιστου παρόχου
      - Λογικά θα πρέπει να χρησιμοποιούνται πολλαπλοί πάροχοι σε αυτή την περίπτωση
  - Δύο τεχνικές μπορούν να εφαρμοστούν – και οι δύο προτείνουν την κρυπτογράφηση των ευαίσθητων δεδομένων
    - “two can keep a secret”
      - Χρήση 2 παρόχων για την αποθήκευση των τμημάτων των δεδομένων
      - Οι τιμές των δεδομένων που δεν μπορούν να αποθηκευθούν αυτούσιες χωρίς να αποκαλύψουν δεδομένα θα πρέπει να κρυπτογραφηθούν
    - “multiple fragments”
      - Απλώς παραγωγή πολλαπλών κομματιών και αποθήκευσή τους σε πολλούς παρόχους ώστε κανείς να μην μπορεί να είναι σε θέση να συσχετίσει δεδομένα

Διαμοιρασμός Μυστικών (Secret Sharing)

- Αναφέρεται σε μεθόδους κατανομής ενός μυστικού σε μια ομάδα έτσι ώστε κανένα άτομο/μέλος από την ομάδα να έχει στην κατοχή του ένα έξυπνο μέρος του μυστικού
  - Αλλά όταν ένα επαρκές σύνολο από μέλη συνδυάζει το μερίδιό του, τότε το μυστικό μπορεί να ανακατασκευαστεί
- Σε ένα σχήμα διαμοιρασμού μυστικών υπάρχει ένας αντιπρόσωπος και n παίκτες
  - Ο αντιπρόσωπος στέλνει ένα μερίδιο του μυστικού στον κάθε παίκτη αλλά μόνο όταν συγκεκριμένες συνθήκες ικανοποιούνται τότε μόνο οι παίκτες θα μπορούν να αναδομήσουν το μυστικό
    - Ο διαμοιρασμός γίνεται με τέτοιο τρόπο έτσι ώστε μόνο ένας συγκεκριμένος αριθμός από παίκτες (άνω κατώφλι t) μπορεί να ανακατασκευάσει το μυστικό ενώ σε μικρότερες υπο-ομάδες παικτών αυτό δεν είναι εφικτό
    - Το σχήμα αυτό ονομάζεται (t, n)-κατωφλιού

Χρήσεις / Σημαντικότητα

- Τα σχήματα αυτά είναι ιδανικά για την αποθήκευση πληροφορίας που είναι υψηλά ευαίσθητη και σημαντική όπως κρυπτογραφικά κλειδιά, κωδικοί εκτόξευσης πυραύλων και πληροφορίες τραπεζικών λογαριασμών
  - Σημειώνουμε πως για τα 2 τελευταία είδη πληροφοριών δεν απαιτείται η χρήση κρυπτογράφησης

Ζητήματα Διαμοιρασμού Μυστικών

- Τα περισσότερα σχήματα διαμοιρασμού μυστικών δουλεύουν πάνω σε δεδομένα μικρού όγκου και δεν είναι αποδοτικά όταν πρέπει να ανανεωθεί το διαμοιραζόμενο μυστικό

Λύση [3]

- Με βάση την χρήση τεχνικών αποδοτικής διαδικασίας αποκωδικοποίησης της διάδοσης πεποιθήσεων (belief propagation – BP) σε LDPC & LT κωδικούς, προτείνεται η σχεδίαση BP-XOR κωδικών και η χρήση τους για την δημιουργία 3 κλάσεων σχημάτων διαμοιρασμού μυστικών με ονομασία BP-XOR, pseudo-BP-XOR & LDPC
- Τα σχήματα αυτά απαιτούν μόνο γραμμικό αριθμό από XOR λειτουργίες σε δυαδικά αλφαριθμητικά τόσο για την φάση διαμοιρασμού όσο και αναδόμησης του μυστικού
- Επίσης, πετυχαίνουν την βέλτιστη πολυπλοκότητα ανανέωσης
  - Πολυπλοκότητα ανανέωσης σημαίνει τον μέσο αριθμό από bits στα μερίδια των παικτών που πρέπει να ανανεωθεί όταν ένα συγκεκριμένο bit από το κύριο κλειδί αλλάξει
  - Παράδειγμα:
    - Έστω πως ο χρήστης έχει αποθηκεύσει 1 GB στο νέφος και επιθυμεί να αλλάξει 1 KB δεδομένων. Τότε, με βάση τα προτεινόμενα σχήματα, θα πρέπει να αλλάξουν το πολύ 1 KB δεδομένων σε κάθε μερίδιο παίκτη

Άλλα Ζητήματα Διαμοιρασμού Μυστικών

- Κακόβουλη συμπεριφορά των παικτών
  - Ένας παίκτης μπορεί να ψεύδεται όσον αφορά το δικό του μερίδιο για να έχει πρόσβαση στα μερίδια άλλων παικτών
- Μη αποδοτικότητα αποθήκευσης
  - Τα σχήματα διαμοιρασμού μυστικών απαιτούν τον ίδιο όγκο δεδομένων στο μέρος κάθε παίκτη σε σχέση με τα αρχικά δεδομένα (μυστικό)
    - Για 1 GB μυστικού και 10 παίκτες, θα απαιτηθούν 10GB χώρου

Αντιμετώπιση Κακόβουλης Συμπεριφοράς

- Χρήση ενός σχήματος επικυρώσιμου διαμοιρασμού δεδομένων (verifiable secret sharing scheme)
  - Επιτρέπει στους παίκτες να είναι σίγουροι με μια μικρή πιθανότητα λάθους πως κανένας άλλος παίκτης δεν ψεύδεται για τα περιεχόμενα του μεριδιού του
  - Δεν υπολογίζεται με συμβατικό τρόπο αλλά οι παίκτες πρέπει να κάνουν συνεργατικούς υπολογισμούς χωρίς να γνωρίζουν τι ακριβώς δεδομένα χρησιμοποιούν
- Ένα τέτοιο σύστημα προτάθηκε από τους Tal Rabin & Michael Ben-Or [4], το οποίο επιτρέπει την ανίχνευση ανειλικρίνειας στο μέρος του αντιπροσώπου ή στο ένα τρίτο των παικτών ακόμη και αν οι παίκτες αυτοί συντονίζονται από έναν προσαρμοστικό επιτιθέμενο που μπορεί να αλλάζει στρατηγικές σε πραγματικό χρόνο ανάλογα με ποια πληροφορία έχει αποκαλυφθεί

Αποδοτικότητα Αποθήκευσης

- Μια προσέγγιση αντιμετώπισης [5] που ονομάζεται σύντομος διαμοιρασμός δεδομένων (secret sharing made short) συνδυάζει τον αλγόριθμο IDA (διασποράς πληροφοριών) με κλασικά σχήματα διαμοιρασμού κλειδιών
  - Τα δεδομένα πρώτα κρυπτογραφούνται με ένα τυχαία κατασκευασμένο συμμετρικό κλειδί
  - Έπειτα, διαμοιράζονται σε N κομμάτια με βάση τον αλγόριθμο IDA
    - Ο αλγόριθμος αυτός πάλι βασίζεται σε κάποιο κατώφλι αλλά το μέγεθος των δεδομένων αυξάνεται σε μικρότερο βαθμό με βάση την διαίρεση: αριθμός τμημάτων / κατώφλι
      - Για παράδειγμα, αν το κατώφλι είναι 10 και το αριθμός των τμημάτων είναι 15, τότε το συνολικό μέγεθος των μεριδίων θα είναι 1,5 φορές μεγαλύτερο των αρχικών δεδομένων
  - Τέλος, το συμμετρικό κλειδί κρυπτογράφησης διαμοιράζεται με βάση κλασικό σχήμα διαμοιρασμού μυστικών
    - Οπότε κάθε παίκτης λαμβάνει ένα τμήμα των δεδομένων και ένα τμήμα του συμμετρικού κλειδιού
  - Καλή λύση και για την περίπτωση όπου ένα τμήμα μπορεί να έχει δεδομένα που μπορούν να συσχετισθούν με ευαίσθητα
  - Μια επέκταση [6] αφορά την διενέργεια ενός προπαρασκευαστικού βήματος για την εξασφάλιση πως ένας αριθμός μεριδίων μικρότερος από το κατώφλι είναι ανεπαρκής για την αποκρυπτογράφηση των δεδομένων

Αποδοτικότητα Αποθήκευσης

- Σε μια άλλη προσέγγιση [7], κάθε μερίδιο αντιστοιχεί περίπου στο μέγεθος του μυστικού δια k-1
  - Το προτεινόμενο σχήμα χρησιμοποιεί πολυώνυμη παρεμβολή (polynomial interpolation) για την τμηματοποίηση των δεδομένων

Ασφαλής Διαγραφή

- Τα δεδομένα θα πρέπει να διαγράφονται πραγματικά και όχι να μαρκάρονται προς διαγραφή όταν δεν χρειάζονται πια στο νέφος

Χρήση Κρυπτογραφίας

- Όταν τα δεδομένα κρυπτογραφούνται μέσω υπηρεσίας κρυπτογράφησης, η υπηρεσία αυτή διαχειρίζεται τα κλειδιά
  - Οπότε, αν δεν εμπιστευόμαστε τον πάροχο, τότε αυτός θα μπορεί εύκολα να εκμεταλλευτεί τα δεδομένα μας
- Η άλλη λύση είναι η κρυπτογράφηση των δεδομένων πριν σταλούν στο νέφος
  - Μπορούν να εφαρμοστούν διαφορετικές αρχιτεκτονικές που όμως βασίζονται σε 2 εναλλακτικές πολιτικές
    - Αποτίμηση εμπιστοσύνης
    - Κρυπτογράφηση κατηγορημάτων (predicate encryption)

Αποτίμηση Εμπιστοσύνης

- Βασίζεται σε 2 διαφορετικές τεχνικές
  - Εκλεπτυσμένη διαχείριση δικαιωμάτων (fine-grained access control)
  - Απομακρυσμένοι έλεγχοι δεδομένων (remote data auditing)

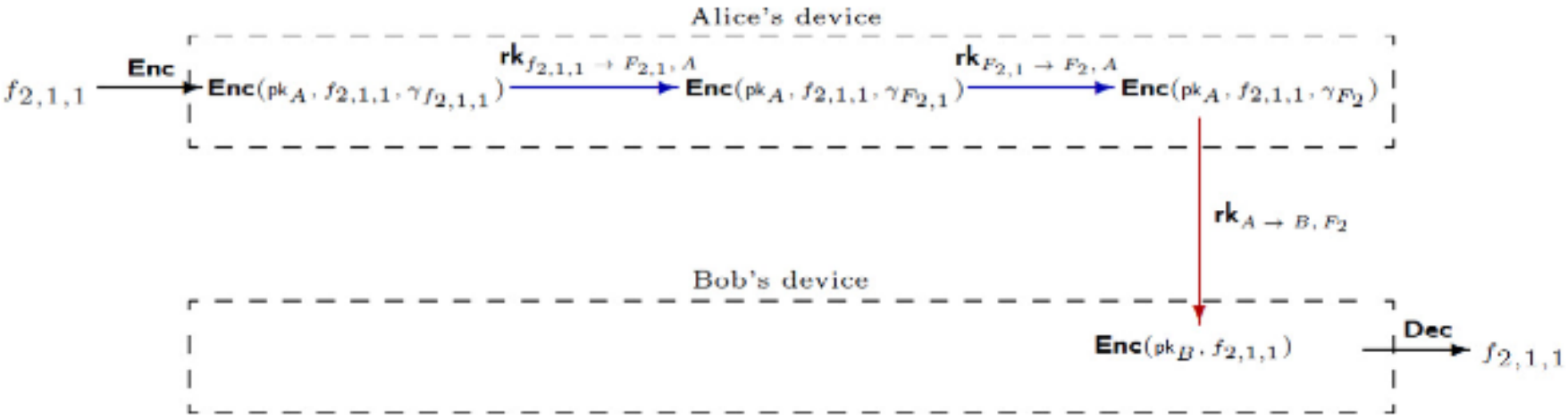
Εκλεπτυσμένη Διαχείριση Δικαιωμάτων

- Βασίζεται στο σχήμα ανα-κρυπτογράφησης πληρεξουσίου (proxy re-encryption)
  - Το σχήμα αυτό επιτρέπει την εξουσιοδότηση της ικανότητας αποκρυπτογράφησης σε ένα πληρεξούσιο (πχ. το νέφος) κατά τον διαμοιρασμό δεδομένων
    - Τα δεδομένα αποθηκεύονται στο πληρεξούσιο κρυπτογραφημένα με το δημόσιο κλειδί του κατόχου
    - Ο κάτοχος των δεδομένων υπολογίζει ένα κλειδί ανα-κρυπτογράφησης που παρέχεται στο πληρεξούσιο και σε άλλον έμπιστο χρήστη (πχ. υπάλληλος του κατόχου ή πελάτης)
    - Το κλειδί επιτρέπει την ανακρυπτογράφηση των κρυπτογραφημένων δεδομένων όταν τα ζητά ο άλλος έμπιστος χρήστης
    - Σημαντικό: Το πληρεξούσιο δεν μπορεί να μάθει την πληροφορία για απλό κείμενο ή οποιουδήποτε μυστικού κλειδιού
  - Το ζήτημα είναι πως όλα τα δεδομένα του κατόχου θα είναι διαθέσιμα στον έμπιστο χρήστη χωρίς κάποιον επιπλέον περιορισμό

Εκλεπτυσμένη Διαχείριση Δικαιωμάτων

- Το σχήμα επεκτάθηκε για την διαχείριση διαμοιραζόμενων δεδομένων μέσω μιας δενδρικής δομής
  - Ο κάτοχος των δεδομένων μπορεί να αποφασίζει ποιο μέρος του δένδρου θα διαμοιράζει σε ποιον χρήστη
  - Η βασική ιδέα αφορά την ανακρυπτογράφηση πληρεξουσίου υπό όρους
    - Η κρυπτογράφηση των δεδομένων βασίζεται σε μια συνθήκη C1 ενώ το κλειδί ανακρυπτογράφησης υπολογίζεται με βάση την συνθήκη C2
    - Αν C1==C2 τότε επιτρέπεται η ανακρυπτογράφηση
    - Αυτό ισχύει για έναν μόνο κόμβο στο δένδρο
      - Αν πρέπει να πηγαίνουμε ψηλότερα στην ιεραρχία διότι πχ. κάποιος ζητά ένα αρχείο ενώ έχει το κλειδί ανακρυπτογράφησης για τον φάκελο, τότε πρέπει να υπάρχει ένα κλειδί ανακρυπτογράφησης που τροποποιεί την συνθήκη του αρχείου (C2') ώστε να μετατραπεί σε συνθήκη για τον φάκελο (C2)
      - Οπότε, εφόσον C1==C2, ο χρήστης θα έχει πρόσβαση στο αρχείο διότι θα κρυπτογραφηθεί εν τέλει με το δικό του κλειδί ανακρυπτογράφησης

Εκλεπτυσμένη Διαχείριση Δικαιωμάτων



Σημείωση: Καλύπτεται η περίπτωση όπου ο χρήστης μπορεί να έχει πρόσβαση στον φάκελο F2 ενώ το αρχείο f2.1.1 βρίσκεται στον υποφάκελο F2,1

Πηγή: [8]

Εναλλακτική Προσέγγιση

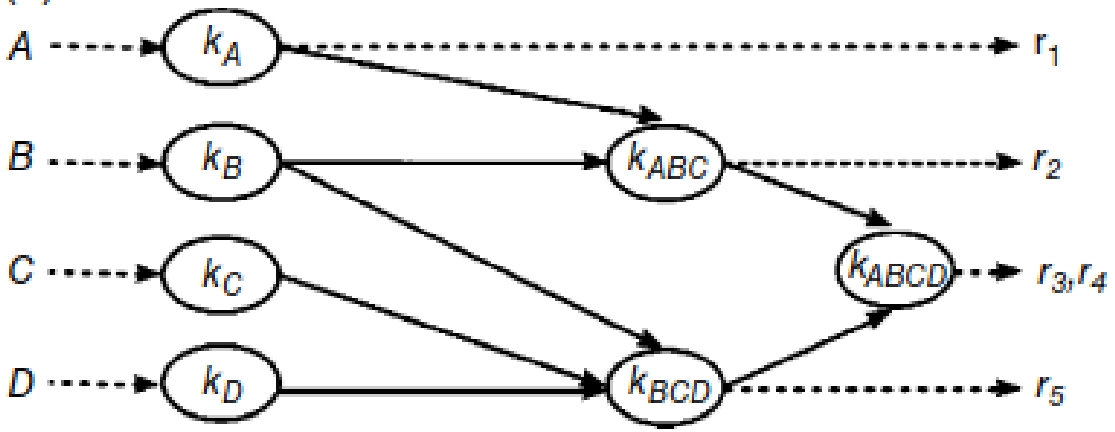
- Εφαρμογή μεθόδων παραγωγής κλειδιού (key derivation methods) όπου οι χρήστες παράγουν τα κλειδιά (που απαιτούνται για την πρόσβαση σε επιθυμητούς πόρους) από ένα απλό κλειδί που έχει ανατεθεί σε αυτούς και δημόσια tokens
  - Ο έλεγχος πρόσβασης εφαρμόζεται μέσω κατάλληλης οργάνωσης των κλειδιών σε μια ιεραρχία που ανακλά την εξουσιοδότηση όπου το κλειδί που αντιστοιχεί σε μια λίστα ελέγχου πρόσβασης επιτρέπει την πρόσβαση σε κλειδιά που έχουν συσχετιστεί με όλες τις λίστες ελέγχου πρόσβασης που είναι υπερσύνολό της

(a)

	r_1	r_2	r_3	r_4	r_5
A	1	1	1	1	0
B	0	1	1	1	1
C	0	1	1	1	1
D	0	0	1	1	1

Access matrix

(b)



Key derivation hierarchy

Πηγή: [9]

Απομακρυσμένοι Έλεγχοι Δεδομένων

- Εστιάζουν στο να εξασφαλίσουν είτε πιθανολογικά είτε ντετερμινιστικά πως τα δεδομένα (πχ. στο νέφος) είναι άθικτα χωρίς να επηρεάσουν την ιδιωτικότητά τους
- Θα πρέπει να εξασφαλίσουν 3 βασικές ιδιότητες:
  - Αποδοτικότητα: διενέργεια ελέγχων με την μικρότερη δυνατόν υπολογιστική πολυπλοκότητα
  - Δημόσια Επικυρωσιμότητα (Public Verifiability): δυνατότητα εξουσιοδότησης σε έναν έμπιστο τρίτο-μέρος ελέγχου (third-party auditor) για την μείωση του υπολογιστικού φόρτου στην πλευρά του πελάτη
  - Πιθανότητα Ανίχνευσης (Detection Probability): έλεγχος της πιθανότητας διαφθοράς των δεδομένων

Απομακρυσμένοι Έλεγχοι Δεδομένων

- Επειδή είναι ασύμφορη η εκφόρτωση όλων των δεδομένων προς επικύρωση, έχουν προταθεί διάφορες τεχνικές για τον πιο αποδοτικό απομακρυσμένο έλεγχο δεδομένων:
  - Επικυρώσιμη κατοχή δεδομένων (provable data possession – PDP)
  - Απόδειξη δυνατότητας ανάκτησης (proof of retrievability – POR)
    - Εφαρμογή πρωτοκόλλου πρόκλησης-απάντησης (challenge-response) για την ανίχνευση κακόβουλης συμπεριφοράς όπου σε κάθε πρόκληση ένα υποσύνολο από μπλοκ αρχείων δειγματοληπτείται και τα αποτελέσματα των υπολογισμών πάνω από τα μπλοκ επιστρέφονται στον πελάτη. Ο πελάτης έπειτα ελέγχει τα αποτελέσματα μέσω πρόσθετης πληροφορίας που έχει ενσωματωθεί στο αρχείο κατά την κωδικοποίησή του πριν αποσταλεί σε έναν απομακρυσμένο πάροχο νέφους
      - Το πρωτόκολλο αυτό είναι αποδοτικό ως προς το εύρος ζώνης και μπορεί να εγγυηθεί πιθανολογικά πως ένα αρχείο έχει παραμείνει άθικτο στον απομακρυσμένο χώρο αποθήκευσης

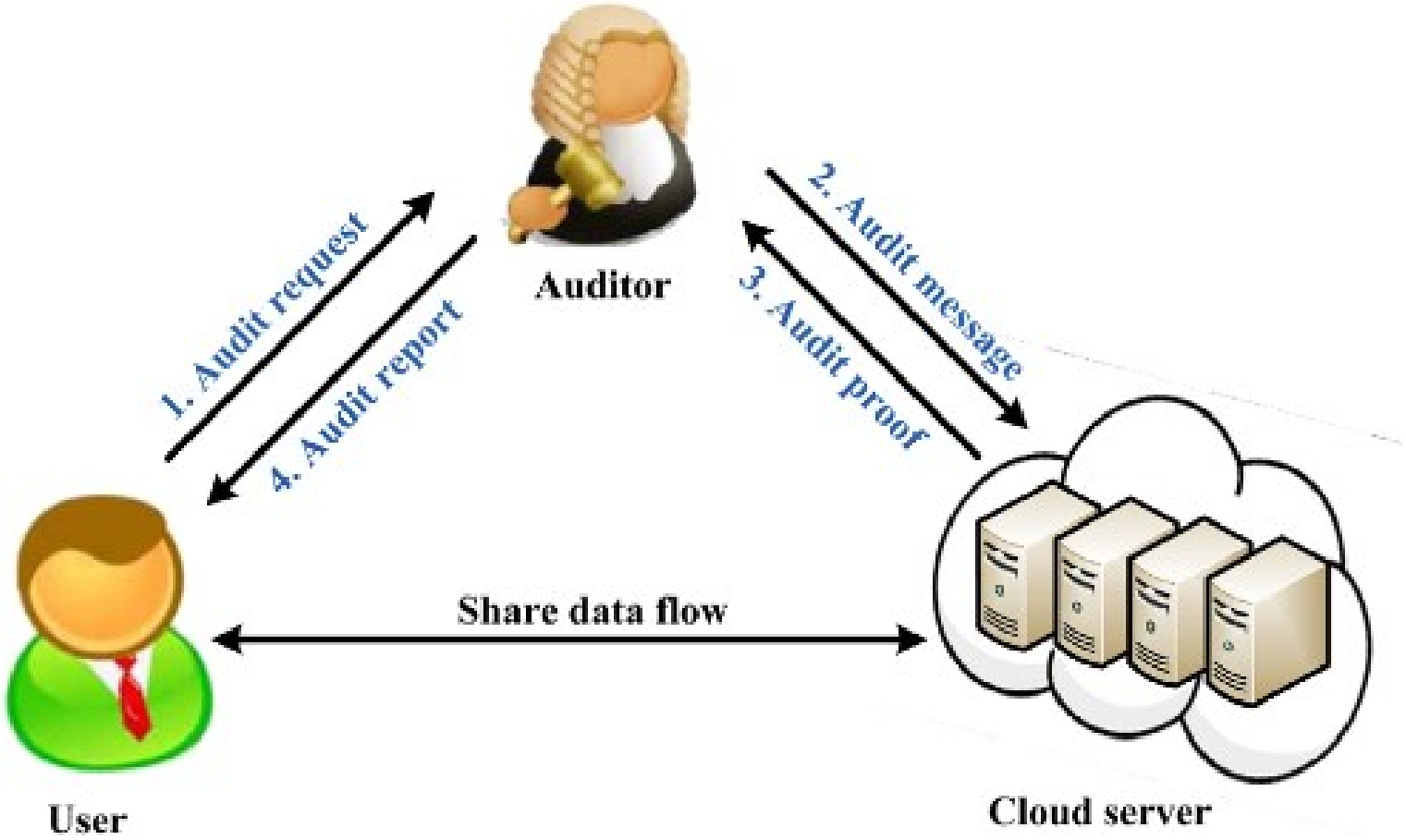
Απομακρυσμένοι Έλεγχοι Δεδομένων

- Βασίζονται σε κρυπτογραφία δημόσιου κλειδιού και άρα πρέπει να αντιμετωπίσουν το ζήτημα της λήξης σχετικών πιστοποιητικών (certificates)
  - Το θέμα είναι πώς ο κάτοχος των δεδομένων θα μπορεί να επικυρώσει την ακεραιότητα τους στο νέφος εφόσον οι παλαιοί ταυτοποιητές (authenticators) που έχουν αποθηκευθεί στο νέφος δεν είναι πια έγκυροι με βάση το νέο δημόσιο κλειδί
  - Μια απλοϊκή λύση είναι η εκφόρτωση όλων των μπλοκ του αρχείου και των ταυτοποιητών, ο επανα-υπολογισμός των ταυτοποιητών για κάθε μπλοκ και η ανανέωση/μεταφόρτωση τόσο των μπλοκ όσο και των ταυτοποιητών
    - Οδηγεί σε υψηλό κόστος επικοινωνίας καθώς και σε υψηλό κόστος υπολογισμού

Απομακρυσμένοι Έλεγχοι Δεδομένων

- Ένα άλλο ζήτημα είναι αυτό της ιδιωτικότητας, τόσο ταυτότητας όσο και των δεδομένων
  - Η ταυτότητα του υπογράφων σε κάθε μπλοκ διαμοιραζόμενων δεδομένων πρέπει να μην αποκαλυφθεί στους δημόσιους επικυρωτές/ελεγκτές (public verifiers)
    - Μια λύση είναι η χρήση υπογραφών δακτυλιδιού και ομάδων (ring & group signatures)
  - Η διαδικασία ελέγχου δεν θα πρέπει να αποκαλύψει την γνώση των αρχείων που χρησιμοποιούνται στις προκλήσεις σε ελεγκτές τρίτων μερών (third party auditors)
    - Μια μη πρακτική λύση είναι η χρήση ομομορφικής κρυπτογράφησης

Αρχιτεκτονική Ελέγχου [10]



Προτεινόμενη Λύση [10]

- Ενοποιεί συστήματα αποδείξεων μηδενικής γνώσης, ομομορφικούς γραμμικούς ταυτοποιητές (homomorphic linear authenticators) και επανα-υπογραφές πληρεξούσιου (proxy re-signatures)
  - Προσφέρει δυνατότητα εξέλιξης των ταυτοποιητών χωρίς την εκφόρτωση του αρχείου/δεδομένων
- Βασίζεται σε 5 αλγορίθμους και ένα διαδραστικό σύστημα αποδείξεων

5 Αλγόριθμοι

- CrsGen(k): Λαμβάνει ως είσοδο μια παράμετρο ασφάλειας k και παράγει ένα κοινό αλφαριθμητικό αναφοράς crs
  - Το τελευταίο χρησιμοποιείται ως έμμεση είσοδος στους υπόλοιπους αλγορίθμους
- KeyGen(crs): Με βάση την είσοδο crs, ο αλγόριθμος παράγει ένα δημόσιο κλειδί pk και ένα ιδιωτικό κλειδί sk για τον κάτοχο των δεδομένων
  - Ο κάτοχος δημοσιεύει μόνο το δημόσιο κλειδί
  - Ο αλγόριθμος αυτός χρησιμοποιείται και για την ανανέωση κλειδιών
- AuthGen (sk, F):
  - Λαμβάνει ως είσοδο ένα ιδιωτικό κλειδί sk και ένα αρχείο F με N μπλοκ mi
  - Παράγει:
    - ένα σύνολο από ταυτοποιητές {Di} για το αρχείο
    - ένα σύνολο από δημόσιες παραμέτρους επικύρωσης φ προς χρήση για τον έλεγχο των δεδομένων του αρχείου κατά την φάση της απόδειξης
- KeyUpdate (sk, pk):
  - Λαμβάνει ως είσοδο (l – 1) παλαιά ζεύγη κλειδιών (skl–1, pkl–1)
  - Παράγει ένα νέο ζεύγος κλειδιών (skl, pkl)
- AuthUpdate (skl, pkl, ftl–1, φ)
  - Λαμβάνει ως είσοδο ένα νέο ζεύγος κλειδιών (pkl, skl), την αρχική ετικέτα του αρχείου ftl–1 και την παράμετρο φ
  - Παράγει μια νέα ετικέτα για το αρχείο ftl και το νέο κλειδί ανανέωσης βl που είναι έγκυρα υπό το πρίσμα του νέου ζεύγους κλειδιού

Χαρακτηριστικά Ασφάλειας Λύσης

- Πληρότητα (completeness)
  - Κατά την αλληλεπίδραση με έναν εξυπηρετητή νέφους που διατηρεί τα δεδομένα άθικτα, το διαδραστικό πρωτόκολλο θα οδηγεί σε θετική επικύρωση όταν τόσο ο εξυπηρετητής όσο και ο ελεγκτής τρίτου μέρους ακολουθούν το πρωτόκολλο ειλικρινά
- Ορθότητα (soundness)
  - Σημαίνει πως οποιοσδήποτε αποδεικνύων (prover) που μπορεί να πείσει έναν επικυρωτή (verifier) πως αποθηκεύει ένα αρχείο δεδομένων, τότε πράγματι το αποθηκεύει
- Ιδιωτικότητα δεδομένων (data privacy)
  - Ο ελεγκτής τρίτου μέρους δεν μπορεί να επάγει γνώση για το περιεχόμενο εκτός από το τυχαίο όνομα του αρχείου με βάση την πληροφορία που είναι δημοσίως διαθέσιμη

Κρυπτογράφηση Κατηγορημάτων

- Η κρυπτογράφηση με βάση τις ιδιότητες είχε αναφερθεί πως παρέχει εκλεπτυσμένο έλεγχο πρόσβασης πάνω από κρυπτογραφημένα δεδομένα αλλά αποκαλύπτει τις πολιτικές πρόσβασης
- Η κρυπτογράφηση κατηγορημάτων (predicate encryption) θεωρείται εξειδίκευση της κρυπτογράφησης με βάση τις ιδιότητες που δεν επιτρέπει την αποκάλυψη των πολιτικών πρόσβασης

Κρυπτογράφηση Κατηγορημάτων

- Σε ένα σχήμα κρυπτογράφησης κατηγορημάτων, τα μυστικά κλειδιά αντιστοιχούν σε κατηγορήματα και τα κρυπτογραφημένα δεδομένα αντιστοιχούν σε ιδιότητες
- Το μυστικό κλειδί SKf που αντιστοιχεί σε ένα κατηγόρημα f μπορεί να χρησιμοποιηθεί για την αποκρυπτογράφηση δεδομένων που σχετίζονται με μια ιδιότητα I μόνο και μόνο όταν f(I)=1
- Συνεπώς, ο διαμοιρασμός των δεδομένων μπορεί να γίνει με βάση τις ιδιότητες που τα αφορούν
  - Κάθε χρήστης θα λαμβάνει μόνο μυστικά κλειδιά για τις ιδιότητες που θα μπορεί να δει

Κρυπτογράφηση Κατηγορημάτων

- Βασικές ιδιότητες προς εξασφάλιση:
  - Απόκρυψη μηνύματος/δεδομένων:
    - Αν ένας επιτιθέμενος έχει πρόσβαση σε k μυστικά κλειδιά SKf1, …, SKfk, τότε δεν θα μπορεί να αποκρυπτογραφήσει μήνυμα/δεδομένα με κάποιο από τα κλειδιά του εφόσον τα δεδομένα αυτά αντιστοιχούν σε ιδιότητα I που δεν καλύπτεται από τα κατηγορήματά του, δηλαδή f1(I) == f2(I) == … == fk(I) == 0
  - Απόκρυψη ιδιοτήτων:
    - Τα κρυπτογραφημένα δεδομένα κρύβουν την πληροφορία για την σχετική ιδιότητα I εκτός από οτιδήποτε διαρρέει ρητά μέσω της χρήσης του ιδιωτικού κλειδιού που κάποιος κατέχει (δηλ. τις τιμές fi(I) και το ίδιο το μήνυμα ακόμη και αν αποτιμάται σε 1 από κάποιο κατηγόρημα)
    - Αυτή η ιδιότητα αντιστοιχεί στην μη αποκάλυψη της πολιτικής πρόσβασης

Κύρια Ζητήματα

- Αν και η κρυπτογράφηση κατηγορημάτων επιτρέπει τον εκλεπτυσμένο έλεγχο πρόσβασης σε κρυπτογραφημένα δεδομένα, δεν διευκολύνει τόσο πολύ την αναζήτηση παρά μόνο την πρόσβαση στα δεδομένα
- Απλοϊκή λύση
  - Ένας χρήστης θα πρέπει να εκφορτώσει τα κρυπτογραφημένα δεδομένα στο δικό του σύστημα, να τα αποκρυπτογραφήσει (όσα του επιτρέπονται) και έπειτα να προχωρήσει στην αναζήτηση
    - Μεγάλος κόστος επικοινωνίας και υπολογισμού
- Εξεζητημένη λύση
  - Ο χρήστης στέλνει την επερώτηση που θα σχετίζεται με μια ή περισσότερες ιδιότητες και τα κατηγορήματα που κατέχει στον πάροχο του νέφους/αποθήκευσης
    - Αυτός ψάχνει τα δεδομένα με βάση τις ιδιότητές τους και εφόσον τα κατηγορήματα του χρήστη «δουλεύουν» πάνω σε αυτές τις ιδιότητες, του επιστρέφει πίσω τα κρυπτογραφημένα δεδομένα
    - Ο πάροχος πρέπει να γνωρίζει τις συσχετίσεις των ιδιοτήτων με τα δεδομένα για να υποστηρίξει την αναζήτηση ενώ μαθαίνει και σε ποια δεδομένα & ιδιότητες έχει πρόσβαση ο χρήστης
      - Ο πάροχος μπορεί να είναι κακόβουλος ή ημι-έμπιστος

Αναζήτηση σε Κρυπτογραφικά Δεδομένα

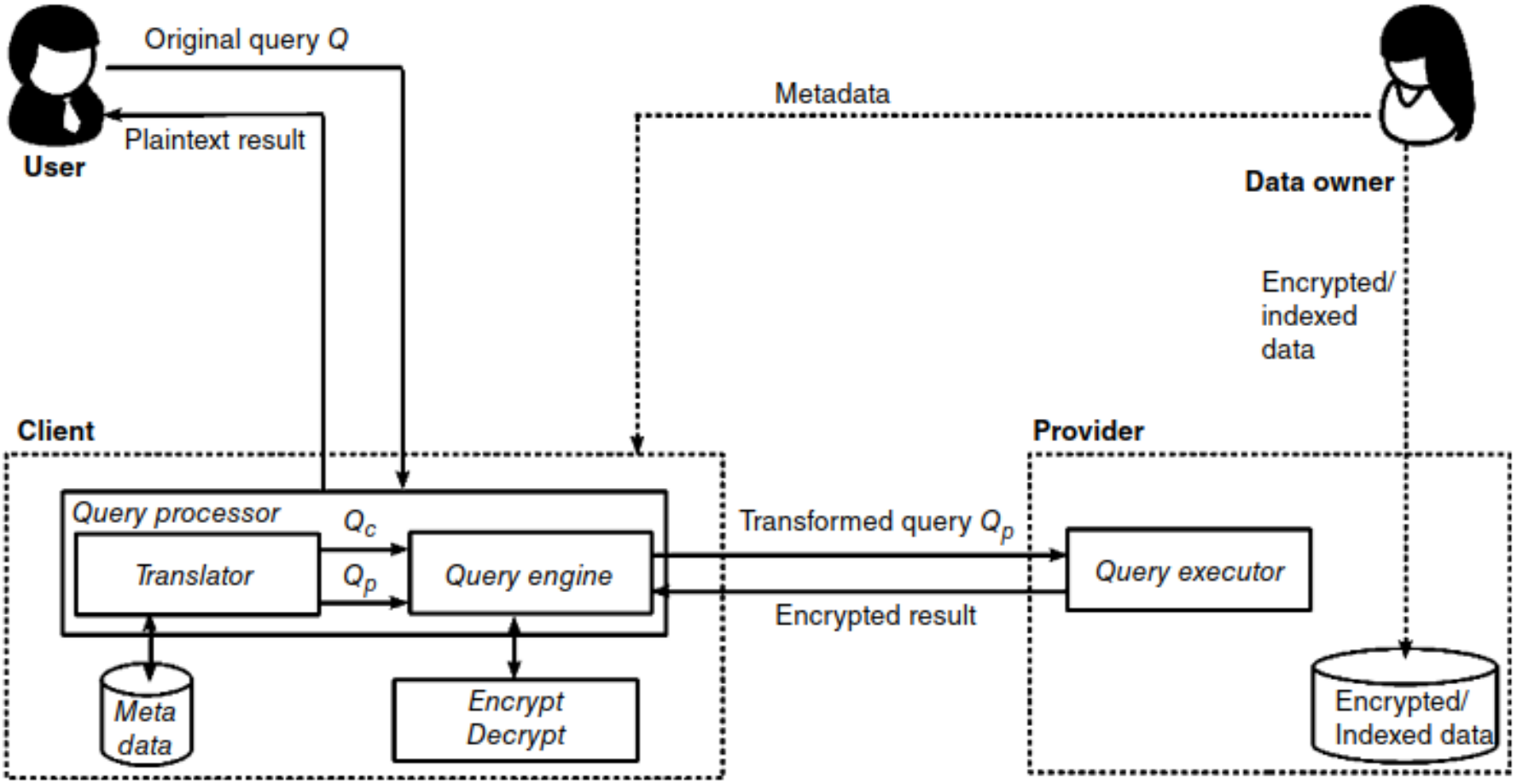
- Εν γένει, μπορούμε να κατατάξουμε τις λύσεις σε 2 κατηγορίες:
  - Απευθείας εφαρμογή επερωτήσεων στα κρυπτογραφημένα δεδομένα
    - Είναι δυνατή με την χρήση συγκεκριμένων τεχνικών κρυπτογράφησης (πχ. ομομορφική κρυπτογράφηση)
    - Κύρια μειονεκτήματα:
      - Κυρίως εφαρμόζονται σε αναζητήσεις για λέξεις κλειδιά και βασικές λειτουργίες
      - Περιορισμένα είδη πρόσβασης
        - Ο έλεγχος πρόσβασης ουσιαστικά θα πρέπει να εξουσιοδοτείται στον πάροχο αποθήκευσης
      - Υπολογιστική πολυπλοκότητα των λύσεων

Αναζήτηση σε Κρυπτογραφημένα Δεδομένα

- Εν γένει, μπορούμε να κατατάξουμε τις λύσεις σε 2 κατηγορίες:
  - Προσκόλληση στα κρυπτογραφημένα δεδομένα μεταδεδομένων (ευρετηρίων) που έπειτα χρησιμοποιούνται για εκλεπτυσμένη ανάκτηση πληροφορίας και εκτέλεση επερωτήσεων
    - Πχ. σε ένα σχεσιακό πίνακα που οι πλειάδες έχουν κρυπτογραφηθεί, διαφορετικά ευρετήρια (indexes) μπορούν να οριστούν σε διαφορετικές ιδιότητες στις οποίες συνθήκες θα πρέπει να αποτιμηθούν
    - Τα ευρετήρια θα πρέπει να είναι ακριβή και αποτελεσματικά ως προς την εκτέλεση επερωτήσεων αλλά συγχρόνως να μην αποκαλύπτουν πληροφορία
    - Η προστασία θα πρέπει να εξασφαλιστεί τόσο από στατικές (όσον αφορά τα κρυπτογραφημένα και ευρετηριασμένα δεδομένα) όσο και από δυναμικές (όσον αφορά τις επερωτήσεις που εκτελούνται στα δεδομένα αυτά) παρατηρήσεις

Αναζήτηση σε Κρυπτογραφημένα Δεδομένα

- Εν γένει, μπορούμε να κατατάξουμε τις λύσεις σε 2 κατηγορίες:
  - Προσκόλληση στα κρυπτογραφημένα δεδομένα μεταδεδομένων (ευρετηρίων) που έπειτα χρησιμοποιούνται για εκλεπτυσμένη ανάκτηση πληροφορίας και εκτέλεση επερωτήσεων (συνέχεια)
    - Διαφορετικά είδη ευρετηρίων έχουν μελετηθεί όπως απευθείας ευρετηριασμός (1-1 αντιστοίχιση καθαρού κειμένου και τιμών ευρετηρίου), βασισμένος σε κατακερματισμό ευρετηριασμού (N-1 αντιστοίχιση μεταξύ καθαρού κειμένου και τιμών ευρετηρίου) και επίπεδος (flat) ευρετηριασμός (1-N αντιστοίχιση)
    - Άλλα είδη ευρετηρίων σχετίζονται με δομές δεδομένων βασιζόμενες σε δένδρα καθώς και λύσεις διατήρησης σειράς ή ομομορφικής κρυπτογράφησης για την παροχή υποστήριξης σε επερωτήσεις εύρους και λειτουργίες συσσωμάτωσης (aggregation functions)
    - Διαφορετικές προσεγγίσεις στον ευρετηριασμό παρέχουν διαφορετικές εξασφαλίσεις προστασίας και διαφορετικά επίπεδα απόδοσης στην εκτέλεση επερωτήσεων
      - Πχ. οι N-1 & 1-N προσεγγίσεις παρέχουν καλύτερη προστασία εμπιστευτικότητας του ευρετηριασμού σε σχέση με την 1-1 αντιστοίχιση αλλά δημιουργούν επιπλέον πολυπλοκότητα στην διαδικασία των επερωτήσεων
      - Οι προσεγγίσεις ευρετηριασμού με κρυπτογράφηση διατήρησης σειράς (order-preserving) παρέχουν υποστήριξη για επερωτήσεις εύρους αλλά εκτίθενται σε κάποια έκθεση πληροφορίας


Original query Q
Plaintext result
User
Metadata
Data owner
Encrypted/
indexed
data
Client
Provider
Query processor
Q_c
Q_p
Translator
Query engine
Transformed query Q_p
Query executor
Encrypted result
Meta
data
Encrypt
Decrypt
Encrypted/
Indexed data

Πηγή: [9]

Κρυπτογράφηση Διατήρησης Σειράς

- Η κρυπτογράφηση διατήρησης σειράς (order-preserving encryption) είναι ένα ντετερμινιστικό σχήμα του οποίου η συνάρτηση κρυπτογράφησης διατηρεί την αριθμητική ταξινόμηση του κυρίως κειμένου
- Το κύριο πλεονέκτημα είναι πως επιτρέπονται αποδοτικές επερωτήσεις εύρους σε κρυπτογραφημένα δεδομένα μέσω ευρετηριασμού των κρυπτογραφημένων δεδομένων
- Επιπλέον, ο ευρετηριασμός και η επεξεργασία επερωτήσεων μπορεί να πραγματοποιηθεί το ίδιο καλά όσο και για τα μη κρυπτογραφημένα δεδομένα
  - Η επερώτηση απλώς περιλαμβάνει τις κρυπτογραφήσεις των a & b οπότε ο εξυπηρετητής μπορεί να εντοπίσει τα επιθυμητά δεδομένα σε λογαριθμικό χρόνο μέσω πρότυπων δενδροειδών δομών δεδομένων

Ιδιωτικότητα Επερωτήσεων (Query Privacy)

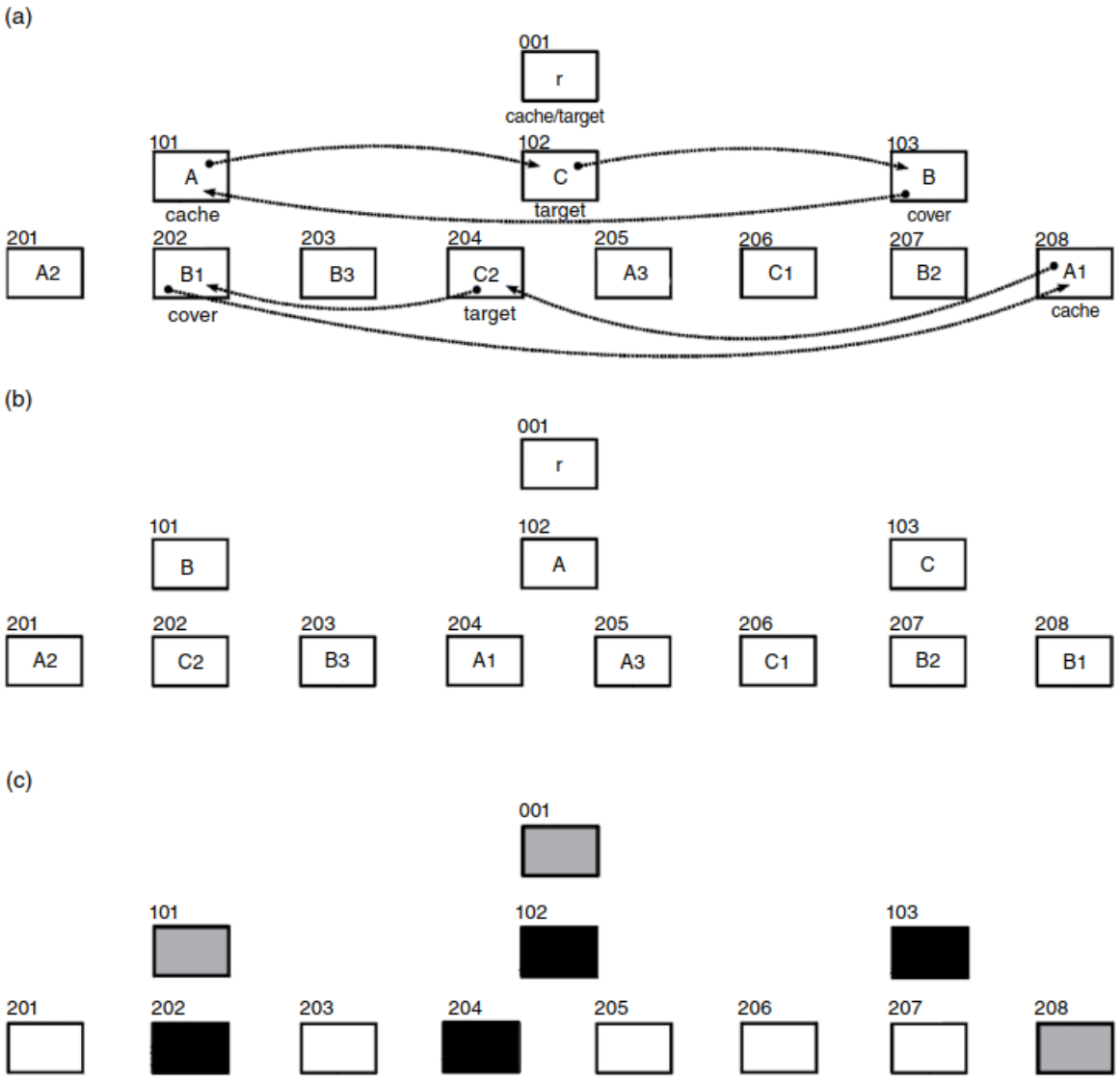
- Ιδιωτικότητα επερωτήσεων (Query privacy)
  - Μερικές φορές η πρόσβαση στα δεδομένα θα πρέπει να είναι εμπιστευτική
    - Ειδικότερα η εμπιστευτικότητα θα πρέπει να εξασφαλιστεί ακόμη και στα μάτια του παρόχου ως προς την πρόσβαση σε συγκεκριμένα δεδομένα (εμπιστευτικότητα πρόσβασης) ή ως προς το γεγονός πως 2 προσβάσεις προσβλέπουν στα ίδια δεδομένα (εμπιστευτικότητα μοτίβων – pattern confidentiality)
    - Οι παραδοσιακές προσεγγίσεις για την εξασφάλιση αυτών των 2 ιδιοτήτων βασίζονται σε τεχνικές ιδιωτικής ανάκτησης πληροφορίας (private information retrieval – PIR) που υποθέτουν πως μια βάση δεδομένων μοντελοποιημένη ως ένα αλφαριθμητικό N-bit παρέχει πρωτόκολλα για την ανάκτηση του i-οστού bit στο αλφαριθμητικό χωρίς να εκθέσει στον πάροχο το bit που προσπελαύνεται
      - Εκτός από τους περιορισμούς μιας τέτοιας μοντελοποίησης και του γεγονότος της μη θεώρησης της εμπιστευτικότητας δεδομένων, υπάρχει και θέμα υψηλούς υπολογιστικής πολυπλοκότητας και κόστους επικοινωνίας

Ιδιωτικότητα Επερωτήσεων (Query Privacy)

- Ιδιωτικότητα επερωτήσεων (Query privacy)
  - Πρόσφατες προσπάθειες εστιάζουν στην εφαρμοσιμότητα και πρακτικότητα της PIR
    - Προστατεύουν την εμπιστευτικότητα δεδομένων με την κρυπτογράφηση και την εμπιστευτικότητα πρόσβασης και μοτίβων με την δυναμική τροποποίηση σε κάθε πρόσβαση της φυσικής τοποθεσίας των δεδομένων ώστε να καταστραφεί η στατική αντιστοίχιση μεταξύ των δεδομένων και των φυσικών μπλοκ όπου αυτά έχουν αποθηκευθεί
    - Επίσης, εφαρμόζουν κρυφή μνήμη για την διατήρηση ορισμένων δεδομένων στην πλευρά του πελάτη

Ιδιωτικότητα Επερωτήσεων (Query Privacy)

- Ιδιωτικότητα επερωτήσεων (Query privacy)
  - Πρόσφατες προσπάθειες εστιάζουν στην εφαρμοσιμότητα και πρακτικότητα της PIR (συνέχεια)
    - Εκτός από την χρήση της κρυφής μνήμης και την δυναμική ανάθεση, η προσέγγιση Path ORAM προϋποθέτει μια δεντρική δομή όπου οι κόμβοι μπορούν να περιέχουν εκτός από κανονικά και πλαστά μπλοκ για να διασφαλίσουν ότι οι κόμβοι έχουν πάντοτε το ίδιο μέγεθος
    - Το ευρετήριο ανακατέματος (Shuffle index) προϋποθέτει πως για κάθε πρόσβαση επιπρόσθετες ψευδείς επερωτήσεις, με όνομα επερωτήσεις κάλυψης, εκτελούνται μαζί με την πραγματική αναζήτηση
      - Οι επερωτήσεις κάλυψης συγχέουν τον πάροχο όσον αφορά το μπλοκ που στοχεύεται
      - Σε κάθε πρόσβαση, το περιεχόμενο των μπλοκ που προσπελαύνονται και στην κρυφή μνήμη ανακατεύεται και επαναγράφεται
      - Αυτό δυναμικά τροποποιεί την δέσμευση των κόμβων οπότε ο πάροχος μπορεί μόνο να παρατηρήσει ότι ορισμένα μπλοκ έχουν διαβαστεί και γραφθεί
      - Αν υποθέσουμε μια ιεραρχική οργάνωση των δεδομένων με βάση την τιμή τους (δηλ. ένα B-tree με κρυπτογραφημένο περιεχόμενο κόμβων και χωρίς δείκτη μεταξύ των φύλλων), το Shuffle index μπορεί να υποστηρίξει επερωτήσεις εύρους


(a)
001
r
cache/target
101
A
cache
102
C
target
103
B
cover
201
A2
202
B1
cover
203
B3
204
C2
target
205
A3
206
C1
207
B2
208
A1
cache
(b)
001
r
101
B
102
A
103
C
201
A2
202
C2
203
B3
204
A1
205
A3
206
C1
207
B2
208
B1
(c)
001
101
102
103
201
202
203
204
205
206
207
208

Πηγή: [9]

Κύρια Ζητήματα

- Πως μπορεί να γίνει η ανάκληση των μυστικών κλειδιών & κατηγορημάτων που έχει δώσει ο κάτοχος σε χρήστες
  - Προφανώς και με αποδοτικό τρόπο χωρίς την ανάγκη ανακρυπτογράφησης των δεδομένων

Λύση [11]

- Για την ανάκληση:
  - Ο κάτοχος των δεδομένων δημιουργεί τόσο ένα μυστικό κλειδί όσο και ένα token χρονικού περιορισμού και τα μοιράζει στους κατάλληλους χρήστες
  - Παράλληλα στέλνει ένα μέρος του token χρονικού περιορισμού στον εξυπηρετητή αποθήκευσης
  - Οπότε, ο εξυπηρετητής αποθήκευσης θα ελέγχει αν η τρέχουσα χρονική στιγμή (όπως στέλνεται από τον χρήστη με βάση το δικό του token χρονισμού) ταιριάζει με το μέρος του token χρονισμού που έχει στην κατοχή του
  - Ο κάτοχος των δεδομένων, εφόσον επιθυμεί την ανάκληση, ζητά από τον εξυπηρετητή αποθήκευσης να διαγράψει/αφαιρέσει το μέρος του token χρονισμού που έχει αποθηκεύσει
    - Οπότε ο κάτοχος του token αυτού δεν θα μπορεί πλέον να λάβει τα σχετικά δεδομένα
    - Λογικά ο κάτοχος θα μπορεί να δημιουργεί νέο token χρονισμού και να το κατανέμει ανάλογα

Λύση [11]

- Προσφέρει ακόμη την δυνατότητα μη κρυπτογραφήσιμης εξουσιοδοτημένης αναζήτησης
  - ο κάτοχος των δεδομένων μπορεί να αποφασίζει σε ποιους χρήστες να επιτρέπει την αποκρυπτογράφηση των δεδομένων και σε ποιους όχι
    - Οι δεύτεροι, ακόμη και αν έχουν την ικανότητα να λάβουν πίσω τις απαντήσεις, δεν θα μπορούν να τις αποκρυπτογραφήσουν
  - Αυτό γίνεται μέσω διαφοροποίησης στο περιεχόμενο του μυστικού κλειδιού που διανέμεται
    - Το πλήρες κλειδί στέλνεται στους χρήστες με την δυνατότητα κρυπτογράφησης
    - Ενώ μέρος του κλειδιού (ουσιαστικά το πλήρες κλειδί με ορισμένα κενά μέρη) στέλνεται σε άλλους χρήστες

Αναζητήσιμη Συμμετρική Κρυπτογράφηση

- Αναζητήσιμη Συμμετρική Κρυπτογράφηση (Searchable Symmetric Encryption)
  - Νέο είδος κρυπτογραφικού σχήματος που υποστηρίζει την αναζήτηση σε κρυπτογραφημένα δεδομένα
    - Ο κάτοχος πρώτα χρησιμοποιεί έναν ειδικό κρυπτογραφικό αλγόριθμο που παράγει μια κρυπτογραφημένη έκδοση της βάσης (των δεδομένων και των μεταδεδομένων) και την αποθηκεύει σε έναν απομακρυσμένο εξυπηρετητή
    - Έπειτα, ο κάτοχος μπορεί να αλληλεπιδρά με τον εξυπηρετητή και να πραγματοποιεί επερωτήσεις στην βάση

Αναζητήσιμη Συμμετρική Κρυπτογράφηση

- Βασικά ζητήματα:
  - Μόνο ο κάτοχος θα κάνει επερωτήσεις;
  - Οι επερωτήσεις θα βασίζονται σε μια ή πολλαπλές λέξεις-κλειδιά;
    - Μια λέξη-κλειδί είναι περιοριστική και μπορεί να οδηγήσει σε χιλιάδες αποτελέσματα
  - Μπορούν να συνδυάζουν τις λέξεις-κλειδιά με λογικούς τελεστές;
  - Παρέχεται κάποιο είδος κλιμακωσιμότητας;
  - Πως μπορεί να προστατευθεί η ιδιωτικότητα των επερωτήσεων;

Κρυπτογραφία Δημόσιου Κλειδιού με Αναζήτηση Λέξεων-Κλειδιών

- Αν και η Αναζητήσιμη Συμμετρική Κρυπτογράφηση είναι αποδοτική, έχει περιορισμένη εκφραστικότητα ως προς την αναζήτηση
- Για αυτό τον λόγο, έχει προταθεί η Κρυπτογραφία Δημόσιου Κλειδιού με Αναζήτηση Λέξεων-Κλειδιών (Public Key Encryption with Keyword Search – PKES)
  - Η PKES προσφέρει πολύ καλύτερη εκφραστικότητα
    - Διαχωρίζει τα αναζητήσιμα περιεχόμενα από τις λέξεις-κλειδιά προς αναζήτηση
    - Υποστηρίζει όχι μόνο συζευκτικές (conjunctive) αλλά και επερωτήσεις υποσυνόλων/εύρους (subset/range)
  - Επιτρέπει στον οποιοδήποτε να κρυπτογραφεί αναζητήσιμα δεδομένα
  - Αλλά
    - μόνο μια καθορισμένη ομάδα χρηστών μπορεί να δημιουργεί trapdoor αναζήτησης
    - δεν κατοχυρώνει την ιδιωτικότητα των επερωτήσεων
      - Ο εξυπηρετητής θα γνωρίζει ποια είναι τα δεδομένα που ενδιαφέρουν τον χρήστη

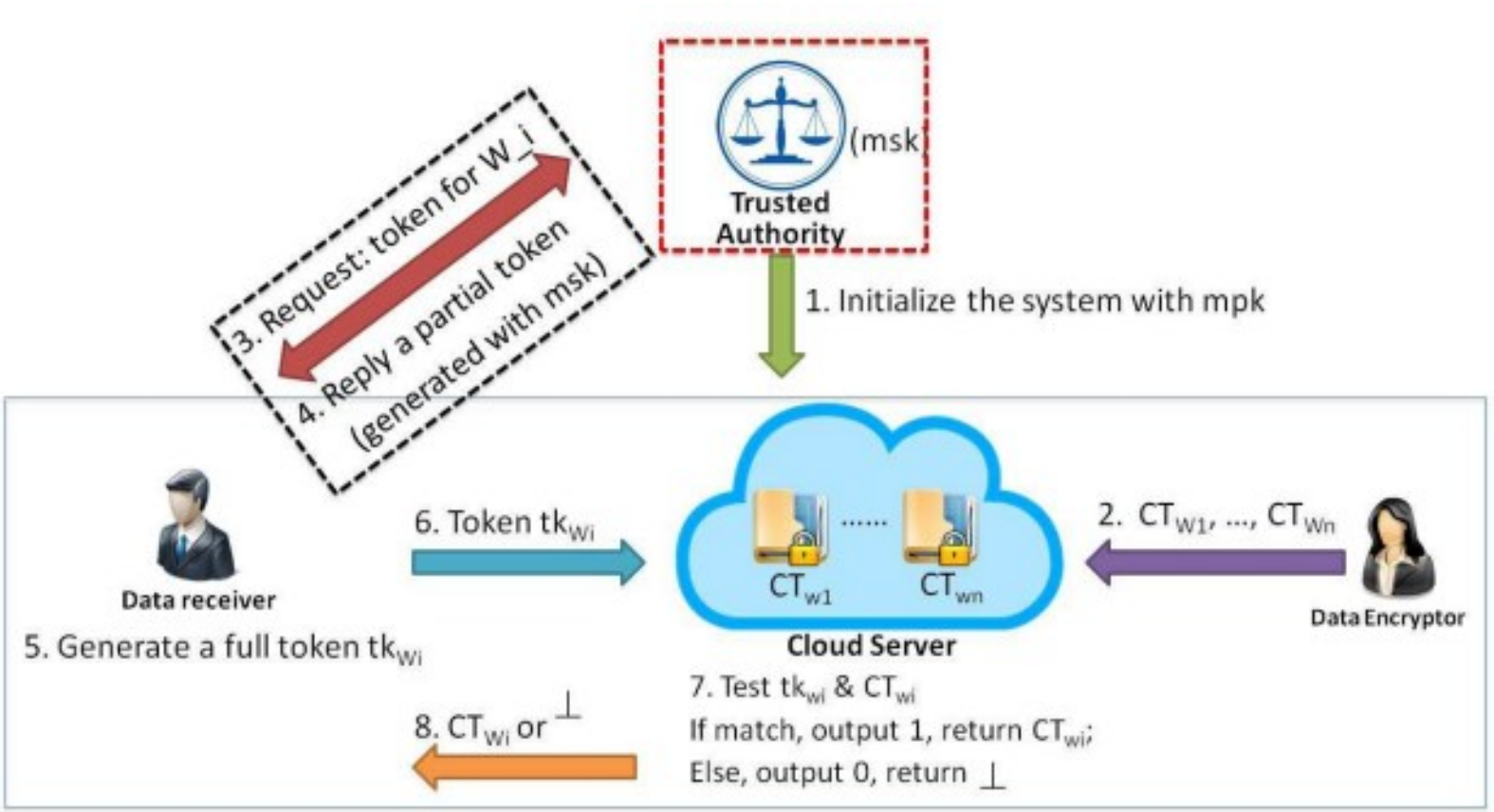
Αναζήτηση με Τυπική Γλώσσα [12]

- Η Αναζήτηση με Τυπική Γλώσσα (Regular Language Search) έχει την βάση της στην PKES αλλά σχετίζεται με τα αυτόματα γλωσσών
  - Ορίζει μια νέα έννοια που ονομάζεται: αναζητήσιμη λειτουργική κρυπτογράφηση με βάση τα πεπερασμένα αυτόματα (searchable fine automata-based functional encryption)

Πλεονεκτήματα

- Οδηγεί σε αποδοτικές επερωτήσεις στα κρυπτογραφημένα δεδομένα
- Μπορεί να υποστηρίξει οποιαδήποτε αναζήτηση με βάση αλφάβητο ή τυπική γλώσσα οπότε είναι πιο φιλική ως προς τον άνθρωπο ως προς την σχεδίαση των κλειδιών αναζήτησης
- Δεν απαιτείται από τον κάτοχο να διαλέξει ειδικές λέξεις-κλειδιά πριν την δόμηση των δομών ευρετηριασμού (πχ. ελάχιστα αναφερόμενη λέξη-κλειδί) ενώ εκμεταλλεύεται την δομή DFA ώστε αν ενσωματώσει ευέλικτη εκφραστικότητα αναζητήσεων (που καλύπτει διάφορους λογικούς τελεστές και όχι μόνο έναν)
- Ο εξυπηρετητής δεν μπορεί εύκολα να εντοπίσει:
  - την συσχέτιση μεταξύ των λέξεων κλειδιών και ενός token
  - την συσχέτιση μεταξύ των λέξεων κλειδιών και των κρυπτογραφημένων δεδομένων
  - Επομένως, υποστηρίζεται η ιδιωτικότητα των επερωτήσεων και δεδομένων

Αρχιτεκτονική



Πηγή: [12]

Χρησιμοποιούμενοι Αλγόριθμοι

- Παραγωγής κλειδιών:
  - Λαμβάνει μια παράμετρο ασφάλειας και το αλφάβητο (λέξεις-κλειδιά) και παράγει ένα ζεύγος δημόσιου & ιδιωτικού κλειδιού
- Κρυπτογράφησης:
  - Κρυπτογραφεί ένα αλφαριθμητικό μήκους l σε κρυπτογραφημένο κείμενο W
- Παραγωγής token:
  - Με βάση το ιδιωτικό κλειδί και μια περιγραφή DFA παράγεται ένα token σε συνεργασία μεταξύ της αρχής και του λήπτη
- Ελέγχου token: ελέγχεται αν ένα κρυπτογραφημένο κείμενο ταιριάζει με το token αναζήτησης

Βασικές Αλληλεπιδράσεις Οντοτήτων

- Η πλήρως εμπιστευμένη αρχή δημιουργεί τις δημόσιες παραμέτρους του συστήματος και αρχικοποιεί το σύστημα καλώντας τον αλγοριθμο παραγωγής κλειδιών και διατηρώντας κρυφό το μυστικό κλειδί
- Ο κάτοχος/κρυπτογραφητής δεδομένων χρησιμοποιεί τον αλγόριθμο κρυπτογράφησης για να κρυπτογραφήσει τα δεδομένα του και τα μεταφορτώνει σε έναν εξυπηρετητή νέφους
- Ο λήπτης δεδομένων εγκαθιδρύει ένα ασφαλές κανάλι επικοινωνίας με την αρχή και αποστέλλει μια DFA πολιτική μαζί με την επερώτηση
  - Η αρχή παράγει με βάση τον τρίτο αλγόριθμο ένα μερικό token
  - Ο λήπτης κατασκευάζει ένα πλήρες token από το μερικό και το στέλνει στον εξυπηρετητή μέσω δημόσιου καναλιού επικοινωνίας
- Ο εξυπηρετητής νέφους λαμβάνει ένα token αναζήτησης που σχετίζεται με την πολιτική αναζήτησης σε DFA (deterministic fine automata) και έπειτα αναζητά τα δεδομένα που ταιριάζουν με το token αυτό

Μέτρα Ιδιωτικότητας Βασιζόμενα στην Πιθανότητα

- Σκοπός είναι να εντοπιστεί αν τα δεδομένα που έχει παραδώσει ένας διανομέας δεδομένων (data distributor) σε ένα σύνολο από φερόμενους εμπιστευμένους πράκτορες έχουν διαρρεύσει σε κάποια μη εξουσιοδοτημένη τοποθεσία από κάποιον από αυτούς καθώς και ποιος είναι υπεύθυνος για την διαρροή
- Υποθέτουμε πως δεν πρέπει να γίνει κάποια διατάραξη των δεδομένων (πχ. δεδομένα ασθενών) με χρήση τεχνικών όπως προσθήκη κάποιου θορύβου στα δεδομένα ή την αντικατάσταση τιμών με εύρη

Μέτρα Ιδιωτικότητας Βασιζόμενα στην Πιθανότητα

- Ένα γνωστό μέτρο είναι το υδατογράφημα (watermarking)
  - Αφορά την εισαγωγή μοναδικών κωδικών στα αντίγραφα των δεδομένων που διανέμονται
    - Οπότε, αν διαρρεύσει κάποιο αντίγραφο, τότε εντοπίζεται άμεσα ο υπαίτιος
  - Όμως:
    - Οδηγεί σε τροποποίηση των δεδομένων
    - Ένας κακόβουλος χρήστης μπορεί να το αφαιρέσει/καταστρέψει

Μέτρα Ιδιωτικότητας Βασιζόμενα στην Πιθανότητα

- Στο άρθρο [13] μελετώνται διακριτικές τεχνικές για τον εντοπισμό διαρροών για ένα σύνολο αντικειμένων ή εγγραφών
  - Η προτεινόμενη λύση βασίζεται στον υπολογισμό της πιθανότητας διαρροής ενός αντιγράφου από έναν πράκτορα
    - Αν η πιθανότητα είναι μεγάλη, τότε υπάρχει αρκετά αποδεικτικά στοιχεία πως ο πράκτορας αυτός προκάλεσε την διαρροή

Μέτρα Ιδιωτικότητας Βασιζόμενα στην Πιθανότητα

- Η πιθανότητα αυτή ονομάζεται ως «ενοχή» ενός πράκτορα
- Η λύση επίσης παρουσιάζει αλγορίθμους κατανομής αντικειμένων στους πράκτορες που αυξάνουν την πιθανότητα εντοπισμού του αίτιου της διαρροής
- Τέλος, προτείνεται η επιλογή της προσθήκης ψεύτικων αντικειμένων στο διανεμημένο σύνολο
  - Τα ψεύτικα αντικείμενα δεν αντιστοιχούν σε πραγματικές οντότητες αλλά μοιάζουν ρεαλιστικά
    - Οπότε δρουν ως υδατογράφημα για το σύνολο των δεδομένων χωρίς να τροποποιούν τα μέλη του
    - Αν ένας πράκτορας έχει λάβει ένα ή παραπάνω από αυτά τα αντικείμενα και αυτά έχουν διαρρεύσει, τότε ο διανομέας μπορεί να είναι περισσότερο πεπεισμένος πως ο πράκτορας είναι ένοχος

Ανωνυμοποίηση Δεδομένων

- Θεωρείται μια αρκετά σημαντική προσέγγιση για την εξόρυξη και δημοσίευση ιδιωτικών/ευαίσθητων δεδομένων
- Έχουν προταθεί πολλές κλασικές τεχνικές για την υλοποίησή της όπως k-ανωνυμία, t-εγγύτητα, l-ποικιλομορφία, e-διαφορική ιδιωτικότητα
  - Όμως είναι επιρρεπής σε γνωστές επιθέσεις ενώ δεν μπορούν να επιτύχουν την αναγκαία ισορροπία μεταξύ της ιδιωτικότητας και χρηστικότητας
    - Ειδικότερα, δεν εφαρμόζονται εύκολα σε κατανεμημένα περιβάλλοντα

Ανωνυμοποίηση Δεδομένων

- Οι σχετικές ερευνητικές εργασίες μπορούν να χωριστούν σε [14]:
  - Τεχνικές που αντιμετωπίζουν κοινές επιθέσεις
  - Τεχνικές ασφαλούς υπολογισμού πολλαπλών μερών (secure multi-party computation)
  - Υβριδικές τεχνικές

Τεχνικές Αντιμετώπισης Κοινών Επιθέσεων

- Η (α-κ) ανωνυμία μπορεί να επιτύχει τις ιδιότητες της κ-ανωνυμίας και α-αποσύνδεσης (deassociation) για την διατήρηση της ιδιωτικότητα δημοσιευόμενων δεδομένων
  - Η σχετική συχνότητα της τιμής που συμβαίνει πιο συχνά σε κάθε κλάση ισοδυναμίας πρέπει να είναι μικρότερη ή ίση από ένα οριζόμενο από τον χρήστη κατώφλι
  - Η επίτευξή της είναι NP-hard

Τεχνικές Αντιμετώπισης Κοινών Επιθέσεων

- Η (ρ,α) & (ρ+,α) ευαίσθητη k-ανωνυμία αφορά την επίτευξη διακριτών ευαίσθητων τιμών σε κάθε κλάση ισοδυναμίας με ένα συνολικό βάρος τουλάχιστον α
  - Αλλά δεν είναι αρκετή για την αντιμετώπιση επιθέσεων ομοιότητας (similarity)
- Η (L, α) ποικιλομορφία βασίζεται στην l- ποικιλομορφία και την κατηγοριοποίηση των ευαίσθητων τιμών ιδιοτήτων σε διαφορετικά επίπεδα εμπιστευτικότητας
- Στην λειτουργική (τ, l) ποικιλομορφία όλες οι ευαίσθητες τιμές γενικεύονται μαζί με τις ιδιότητες οιονεί αναγνωριστικών μέχρι να ικανοποιηθεί η (τ, l) ποικιλομορφία
- Όλοι οι αλγόριθμοι βασίζονται σε παραμέτρους οριζόμενες από τον χρήστη για να τεθούν κατάλληλα κατώφλια στο μέγεθος των κλάσεων ισοδυναμίας και τον αριθμό των διακριτών τιμών των ευαίσθητων ιδιοτήτων
  - Οπότε είναι δύσκολη η διαμόρφωση των αλγορίθμων με βάση και το σύνολο δεδομένων προς προστασία και τα χαρακτηριστικά του

Τεχνικές Αντιμετώπισης Κοινών Επιθέσεων

- Το μοντέλο k-ανωνυμίας και β-πιθανότητας εστιάζει την προστασία από την αποκάλυψη ταυτότητας και ιδιοτήτων ενώ αντιμετωπίζει την επίθεση γνώσης υποβάθρου
  - Βασίζεται σε 2 αλγόριθμους ταξινόμησης για την δημιουργία του μοντέλου των 2 αυτών μερών/πλευρών
- Για την περίπτωση δεδομένων τροχιάς (trajectory), μια προσέγγιση είναι η γενίκευση των ευαίσθητων δεδομένων και η κατάπνιξη των δεδομένων τροχιάς ώστε να αντιμετωπισθούν οι επιθέσεις συσχέτισης και ομοιότητας

Ασφαλής Υπολογισμός Πολλαπλών Μερών

- Αποτελεί κρυπτογραφική τεχνική για την επίτευξη ιδιωτικότητας δεδομένων κατά την παρουσία τόσο ημι-εμπιστευτικών αλλά και κακόβουλων κόμβων
  - Μόνο τα δεδομένα που απαιτούνται για έναν κατανεμημένο υπολογισμό (πχ. στατιστικά για μια ιδιότητα) ανταλλάσσονται μεταξύ των μερών με έναν ασφαλή τρόπο
  - Επομένως, δεν ανταλλάσσονται όλα τα δυνατά δεδομένα
  - Ενώ και η ανταλλαγή των δεδομένων γίνεται για την ασφαλή εκτέλεση μόνο μιας λειτουργίας

Υβριδικές Τεχνικές

- Συνδυάζουν τεχνικές ανωνυμίας δεδομένων με άλλες για την επίτευξη της ιδιωτικότητας
- Μια πρώτη τεχνική ανωνυμοποιεί κατανεμημένα δεδομένα χρησιμοποιώντας ασφαλή υπολογισμό πολλαπλών μερών
  - Κάθε μέρος στέλνει τα κρυπτογραφημένα δεδομένα στο επόμενο και έπειτα τα τελικά ενοποιημένα δεδομένα ανωνυμοποιούνται για την προστασία από επιθέσεις συσχέτισης
- Η ανωνυμοποίηση με βάση τις μεταθέσεις (permutation) λαμβάνει υπόψιν το υποκείμενο των δεδομένων, τον εισβολέα και την διαφάνεια της ανωνυμοποίησης
  - Εστιάζει στην ποικιλομορφία των δεδομένων στην ίδια κλάση ισοδυναμίας και άρα αντιμετωπίζει την επίθεση ομοιότητας
  - Αλλά εφαρμόζεται μόνο σε αριθμητικά δεδομένα

Υβριδικές Τεχνικές

- Για την συνεργατική δημοσίευση δεδομένων προτάθηκε το m-μοντέλο ιδιωτικότητας, το οποίο προστατεύει από μέχρι και m κακόβουλα μέρη, τα οποία μπορεί να συνεργάζονται συντεχνιακά ώστε να χρησιμοποιήσουν τις δικές τους εγγραφές για να επάγουν τις εγγραφές δεδομένων των άλλων νόμιμων μελών
- Μια υβριδική προσέγγιση που συνδυάζει την ανωνυμία και την κρυπτογραφία προτάθηκε για την διατήρηση της ιδιωτικότητας σε περιβάλλοντα νέφους
  - Τα δεδομένα διαιρούνται σε τρία μέρη: αναγνωριστικά, οιονεί αναγνωριστικά και ιατρικά δεδομένα
  - Τα αναγνωριστικά κρυπτογραφούνται, τα οιονεί αναγνωριστικά ανωνυμοποιούνται ενώ τα ιατρικά δεδομένα μένουν ως έχουν

Ζητήματα

- Αν και πολλά υποσχόμενες διότι εφαρμόζονται σε κατανεμημένα περιβάλλοντα, οι υβριδικές τεχνικές είναι είτε υπολογιστικά πολύπλοκες είτε μπορούν να αντιμετωπίσουν μόνο τις επιθέσεις συσχέτισης
- Επίσης, εστιάζουν στην ασφαλή ενοποίηση των δεδομένων και δεν μπορούν να αντιμετωπίσουν τις επιθέσεις ομοιότητας ή πιθανολογικής επαγωγής (probabilistic inference)

Λύση [14]

- Υλοποίηση αλγορίθμου ομαδοποίησης KNN(G, S), ο οποίος προσφέρει τις εξής ικανότητες:
  - Αντιμετώπιση επιθέσεων ομοιότητας μέσω μοντέλου ιδιωτικότητας που σχηματίζει ανωνυμοποιημένες κλάσεις ισοδυναμίας, η κάθε μια από τις οποίες έχει την μέγιστη δυνατή ποικιλομορφία ως προς τις τιμές ευαίσθητων ιδιοτήτων
  - Αντιμετώπιση επιθέσεων πιθανολογικής επαγωγής επιτυγχάνοντας ενιαία κατανομή των τιμών ευαίσθητων ιδιοτήτων τόσο στα αρχικά δεδομένα όσο και στις ανωνυμοποιημένες κλάσεις ισοδυναμίας

Μοντέλα Πολλαπλών Κατόχων

- Οι πιο πολλές προσεγγίσεις για την ασφαλή αναζήτηση σε κρυπτογραφημένα δεδομένα θεωρούν πως ένας πάροχος νέφους εξυπηρετεί μόνο έναν κάτοχο δεδομένων
  - Αυτό είναι μη ρεαλιστικό
  - Υπάρχουν περιπτώσεις όπου επιβάλλεται η ύπαρξη πολλαπλών κατόχων δεδομένων όπως στον διαμοιρασμό ιατρικών δεδομένων (πχ. για την εκμετάλλευσή τους από δημόσιες δομές υγείας)

Μοντέλα Πολλαπλών Κατόχων

- Προκλήσεις:
  - Διαθεσιμότητα κατόχων
    - Όταν υπάρχει ένας κάτοχος δεδομένων θα πρέπει να είναι διαθέσιμος για την παραγωγή κρυπτογραφημένων λέξεων-κλειδιών (τα ονομαζόμενα trapdoors)
    - Αλλά αν υπάρχει ένας μεγάλος αριθμός από κατόχους δεδομένων, δεν είναι εφικτό να ζητηθεί όλοι αυτοί να είναι διαθέσιμοι πάντοτε διότι αυτό θα επιδρούσε αρνητικά στην ευελιξία και χρηστικότητα του συστήματος αναζήτησης
  - Διαφορετικότητα κλειδιών
    - Κάθε κάτοχος δεδομένων θα προτιμήσει να χρησιμοποιήσει τα δικά του ασφαλή κλειδιά για την κρυπτογράφηση των δεδομένων του
      - Επομένως, είναι πρόκληση η εκτέλεση ασφαλούς, βολικής και αποδοτικής αναζήτησης πάνω από τα κρυπτογραφημένα δεδομένα με διαφορετικά κρυφά κλειδιά

Μοντέλα Πολλαπλών Κατόχων

- Προκλήσεις:
  - Εγγραφή Χρηστών και Ανάκληση
    - Εφόσον υπάρχουν πολλαπλοί κάτοχοι δεδομένων, θα πρέπει να εξασφαλιστεί η χρήση αποδοτικών μηχανισμών εγγραφής χρηστών και ανάκλησης ώστε το τελικό σύστημα αναζήτησης να έχει κατάλληλα επίπεδα ασφάλειας και κλιμακωσιμότητας

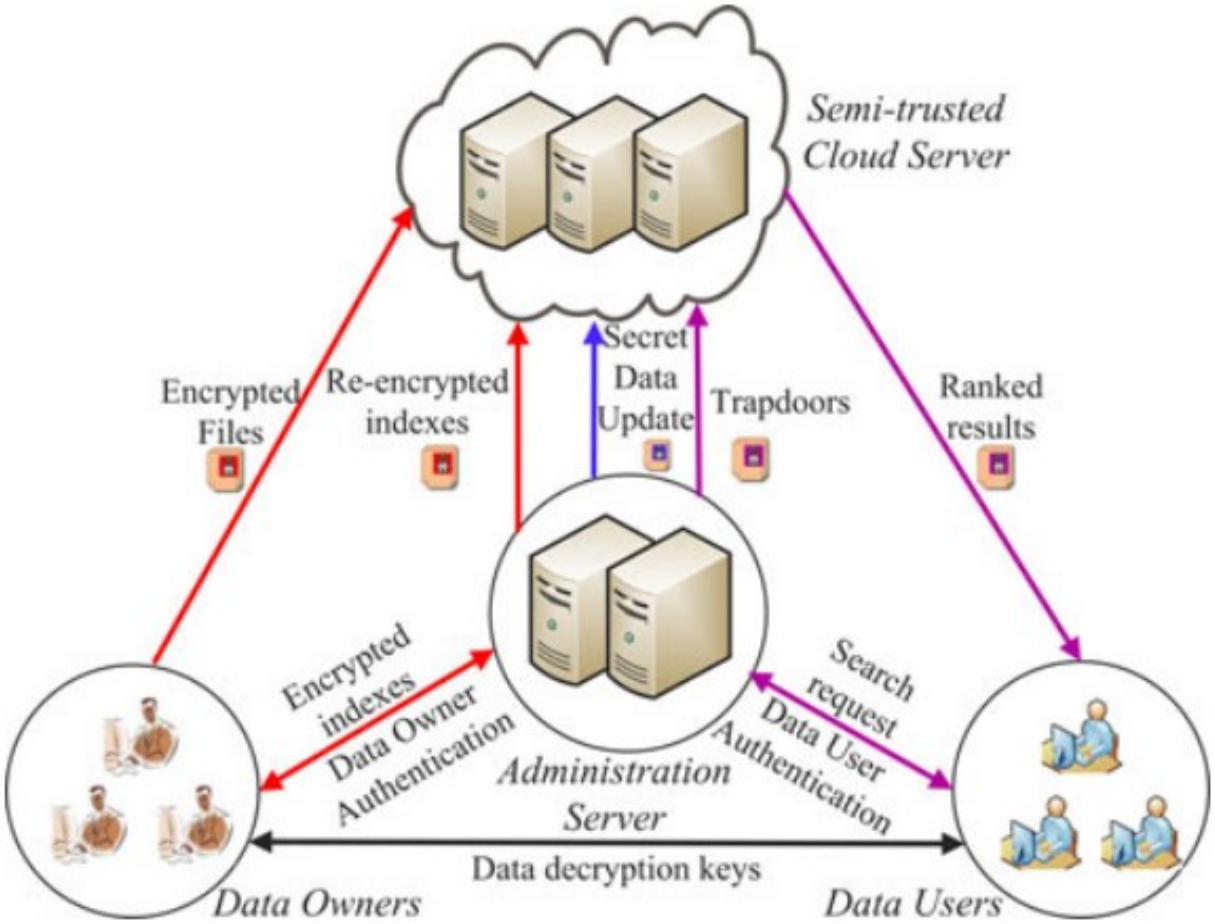
Λύση [15]

- Προτείνεται το PRMSM, ένα πρωτόκολλο ταξινομημένης αναζήτησης πολλαπλών λέξεων κλειδιών (ranked multi-keyword search protocol) που διατηρεί την ιδιωτικότητα σε ένα μοντέλο νέφους πολλαπλών κατόχων (δεδομένων)
  - Το πρωτόκολλο αυτό προσφέρει ασφαλή αναζήτηση χωρίς οι εξυπηρετητές νέφους να γνωρίζουν τις πραγματικές τιμές των κλειδιών αναζήτησης και των trapdoors
  - Επίσης, επιτρέπει την χρήση διαφορετικών κλειδιών και λέξεων κλειδιών από τους κατόχους
  - Ενώ οι χρήστες μπορούν να θέτουν επερωτήσεις χωρίς να γνωρίζουν τα μυστικά κλειδιά των κατόχων
- Για την ταξινόμηση των αποτελεσμάτων αναζήτησης και την διατήρηση της ιδιωτικότητας των σκορ σχετικότητας (relevance scores), προτείνεται μια νέα οικογένεια συναρτήσεων που υποστηρίζει την προσθετική ταξινόμηση και την διατήρηση της ιδιωτικότητας (με βάση τις προτιμήσεις των κατόχων) των σκορ σχετικότητας
  - Αυτό επιτρέπει στον εξυπηρετητή νέφους να αποστείλει τα πιο σχετικά αποτελέσματα αναζήτησης στους χρήστες χωρίς να μπορεί να κατανοήσει τα κωδικοποιημένα σκορ σχετικότητας

Λύση [15]

- Επιπλέον χαρακτηριστικά:
  - Προτείνεται ένα νεωτεριστικό πρωτόκολλο δυναμικής δημιουργίας κρυφών κλειδιών και ένα νέο πρωτόκολλο ταυτοποίησης χρηστών για την αποτροπή επιθέσεων παρακολούθησης και πλαστοπροσωπίας νόμιμων χρηστών
    - Αυτό έχει ως αποτέλεσμα όταν οι επιτιθέμενοι υποκλέπτουν ένα κρυφό κλειδί και εκτελούν παράνομες επερωτήσεις να ανιχνεύονται εύκολα
  - Αποδοτική ανάκληση χρηστών

Αρχιτεκτονική Λύσης



Πηγή: [15]

Ανάλυση Αρχιτεκτονικής

- Οι κάτοχοι δεδομένων έχουν μια συλλογή από αρχεία F. Για να επιτρέψουν την αποδοτική αναζήτηση των αρχείων αυτών (όταν κρυπτογραφηθούν), οι κάτοχοι πρέπει πρώτα να κατασκευάσουν ένα ασφαλή, αναζητήσιμο ευρετήριο I με βάση το σύνολο λέξεων κλειδιών W που έχει εξαχθεί από το F. Έπειτα, αποστέλλουν το I στον εξυπηρετητή διοίκησης. Τέλος, κρυπτογραφούν τα αρχεία τους, δημιουργώντας το σύνολο C, το οποίο αποστέλλεται στον εξυπηρετητή νέφους

Ανάλυση Αρχιτεκτονικής

- Μόλις ληφθεί το ευρετήριο I από τον εξυπηρετητή διοίκησης, αυτός ανα-κρυπτογραφεί το I εκ μέρους του κατόχου δεδομένων και αποθηκεύει το ευρετήριο αυτό (I’) στον εξυπηρετητή νέφους
- Όταν ένας χρήστης επιθυμεί να αναζητήσει t λέξεις κλειδιά πάνω από τα κρυπτογραφημένα αρχεία στον εξυπηρετητή νέφους, πρώτα υπολογίζει τα αντίστοιχα trapdoors και τα αποστέλλει στον εξυπηρετητή διοίκησης
- Ο εξυπηρετητής διοίκησης θα ταυτοποιήσει τον χρήστη και έπειτα θα ανακρυπτογραφήσει τα trapdoor για να τα αποστείλει στον εξυπηρετητή νέφους

Ανάλυση Αρχιτεκτονικής

- Κατά την λήψη ενός trapdoor T, ο εξυπηρετητής νέφους θα προχωρήσει σε σχετική αναζήτηση στα ευρετήρια I’ του κάθε κατόχου και έπειτα θα αποστείλει το σύνολο των ταιριασμένων κρυπτογραφημένων αρχείων
  - Για την βελτίωση της ακρίβειας ανάκτησης και την μείωση του κόστους επικοινωνίας, ο χρήστης μπορεί να περάσει μια παράμετρο k στον εξυπηρετητή για την λήψη μόνο των πρώτων/κορυφαίων k αποτελεσμάτων
- Μόλις λάβει τα αρχεία ο χρήστης, τα αποκρυπτογραφεί

Στόχοι Ασφάλειας Λύσης

- Σημασιολογική ασφάλεια των λέξεων-κλειδιών
- Μυστικότητα/Ιδιωτικότητα λέξεων-κλειδιών
  - Η λύση εξασφαλίζει πως η πιθανότητα για την επαγωγή της πραγματικής τιμής μιας λέξεως-κλειδιού είναι μεγαλύτερη από την πιθανότητα τυχαίας εικασίας της τιμής αυτής
- Μυστικότητα του σκορ συσχέτισης
  - Ο εξυπηρετητής δεν μπορεί να επάγει την πραγματική τιμή των κωδικοποιημένων σκορ σχετικότητας
- Η διαρροή του κλειδιού ενός κατόχου δεν οδηγεί σε διαρροή των δεδομένων των άλλων κατόχων
- Ο εξυπηρετητής δεν μπορεί να συσχετίσει τις λέξεις-κλειδιά μεταξύ τους

Ανάκληση Χρηστών

- Βασίζεται στο γεγονός πως ο εξυπηρετητής διοίκησης δεν στέλνει μόνο το ανα-κρυπτογραφημένο trapdoor αλλά και ένα κρυπτογραφημένο δεδομένο Sa εκ μέρους του χρήστη προς τον εξυπηρετητή νέφους
  - Επομένως, αρκεί η αλλαγή του Sa ώστε να αποτύχει η εκτέλεση επερωτήσεων στον εξυπηρετητή νέφους από ανακλημένο χρήστη
- Επίσης, εφόσον ο χρήστης θα ανακληθεί, δεν θα μπορεί να ζητήσει την ανα-κρυπτογράφηση και αποστολή του trapdoor από τον εξυπηρετητή διοίκησης

Ακεραιότητα Επερωτήσεων & Υπολογισμών

- Εφόσον δεν υπάρχει πλήρης εμπιστοσύνη σε έναν πάροχο, ο κάτοχος δεδομένων ή ο χρήστης θα πρέπει να έχει την δυνατότητα να αποτιμήσει αν ένα αποτέλεσμα που επιστρέφεται από μια επερώτηση / υπολογισμό είναι σωστό, ολοκληρωμένο και φρέσκο
  - Η ορθότητα σημαίνει πως το αποτέλεσμα έχει υπολογιστεί πάνω από τα αρχικά δεδομένα και ο υπολογισμός πραγματοποιήθηκε σωστά
  - Η πληρότητα (completeness) σημαίνει πως κανένα δεδομένο δεν λείπει από το αποτέλεσμα
  - Η φρεσκάδα (freshness) σημαίνει πως ο υπολογισμός/επερώτηση έχει γίνει πάνω από την πιο πρόσφατη έκδοση των δεδομένων

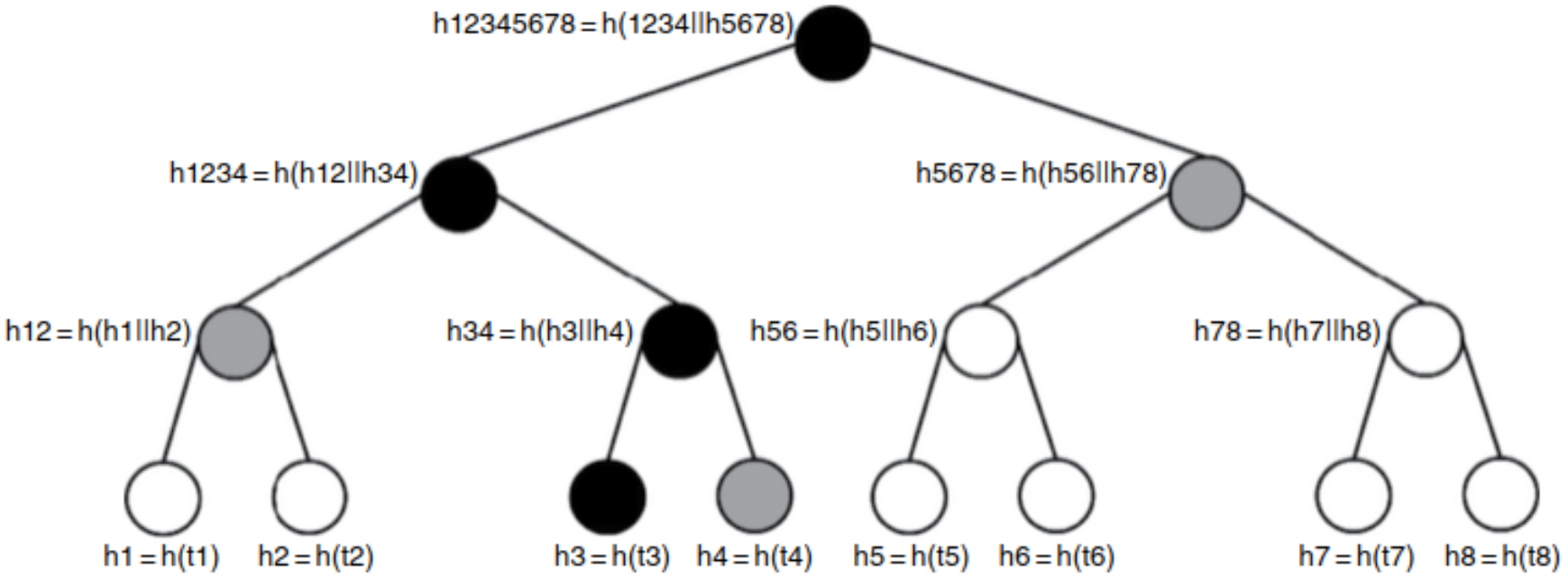
Ακεραιότητα Επερωτήσεων & Υπολογισμών

- Μηχανισμοί / Προσεγγίσεις
  - Οι περισσότερες προσεγγίσεις εστιάζουν στην παροχή εγγυήσεων για την πληρότητα και ορθότητα ενώ προτείνονται και συμπληρωματικές προτάσεις με την χρήση χρονοσφραγίδων και περιοδικής ανανέωσης (periodical refreshing) για την παροχή εγγυήσεων φρεσκάδας
  - Οι τρέχουσες λύσεις μπορούν να ταξινομηθούν σε δύο κατηγορίες: ντετερμινιστικές και πιθανοκρατικές

Ντετερμινιστικές Προσεγγίσεις

- Παρέχονται από αυθεντικοποιημένες δομές δεδομένων, οι οποίες, όπως και τα σχήματα υπογραφών για στατικά δεδομένα, επιτρέπουν την ανίχνευση παραβιάσεων ακεραιότητας με ένα επίπεδο σιγουριάς
  - Τα σχήματα αλυσίδας υπογραφών επιτρέπουν την επικύρωση της σειράς μεταξύ των πλειάδων και έπειτα μπορούν να χρησιμοποιηθούν για την επικύρωση της ακεραιότητας των επερωτήσεων εύρους, όπου η συνθήκη επιλογής βασίζεται στις ιδιότητες στις οποίες το σχήμα υπογραφής έχει εφαρμοστεί
  - Τα δένδρα Merkle και οι εκδόσεις τους οργανώνουν τα δεδομένα σε μια δενδρική δομή πάνω από μια δεδομένη ιδιότητα (πχ. κλειδί αναζήτησης). Το αποτέλεσμα επερώτησης με συνθήκες επιλογής ως προς τις ιδιότητες περιλαμβάνει, εκτός από τις πλειάδες αποτελέσματος, ένα αντικείμενο επικύρωσης που επιτρέπει την αποτίμηση της ακεραιότητας της επερώτησης
- Αυτές οι δομές δεδομένων αυθεντικοποίησης παρέχουν ντετερμινιστικές εγγυήσεις ακεραιότητας μόνο για επερωτήσεις που περιέχουν ιδιότητες με βάση τις οποίες έχει οργανωθεί η δομή δεδομένων

Δένδρο Merkle



Πηγή: [9]

Πιθανοκρατικές Προσεγγίσεις

- Οι ακόλουθες προσεγγίσεις έχουν εφαρμοστεί μεμονωμένα ή σε συνδυασμό για την παροχή πιθανοκρατικών εγγυήσεων ακεραιότητας
  - Εισαγωγή ψευδών πλειάδων στα δεδομένα που αν δεν περιλαμβάνονται στο αποτέλεσμα επερώτησης σηματοδοτούν παραβίαση ακεραιότητας
  - Αναπαραγωγή μέρους των δεδομένων με αντίγραφα που δεν αναγνωρίζονται ως αναπαραγωγές έτσι ώστε η εμφάνιση διπλότυπων δεδομένων όπου το αντίγραφο λείπει σηματοδοτεί παραβίαση ακεραιότητας
  - Προϋπολογισμός tokens που συσχετίζονται με επιλεγμένα αποτελέσματα επερωτήσεων που επιτρέπει την επικύρωση της ακεραιότητας των επερωτήσεων αυτών
- Αν και η απουσία μιας αναμενόμενης ψευδής πλειάδας ή αντιγράφουν σηματοδοτεί ένα πρόβλημα ακεραιότητας, η παρουσία τέτοιων σημαδιών δεν μας επιτρέπει να αποφανθούμε πως η ακεραιότητα δεν έχει παραβιαστεί διότι μπορεί απλώς οι πάροχοι να ήταν τυχεροί ώστε να μην υπάρχει έλλειψη των ελέγχων που εισάγονται από τους ιδιοκτήτες δεδομένων

Πιθανοκρατικές Προσεγγίσεις

- Η πιθανότητα ανίχνευσης παραβίασης ακεραιότητας τυπικά εξαρτάται από το πλήθος των ελέγχων που εφαρμόζονται όπου περισσότερος ο έλεγχος, περισσότερες είναι οι εγγυήσεις αλλά και υψηλότερο είναι και το επιπρόσθετο τίμημα απόδοσης της επικύρωσης
- Η χρήση πολλαπλών παρόχων στην αποθήκευση ή τον υπολογισμό περιπλέκει τα πράγματα και απαιτεί την επινόηση επιπρόσθετων ελέγχων
  - Μια πιθανή λύση για την αποτίμηση της ακεραιότητας των joins που υπολογίζονται από ένα μη έμπιστο πάροχο πάνω από δεδομένα που αποθηκεύονται σε 2 επιστευμένους πάροχους αποθήκευσης υποθέτει την συνεργασία των τελευταίων για την εισαγωγή πληροφορίας ελέγχου που αποτελείται από ψευδής και διπλότυπες πλειάδες ώστε η απουσία τέτοιων σημαδιών να οδηγήσει στην σηματοδότηση μη πληρότητας

Συνεργατική Εκτέλεση Επερωτήσεων

- Τα δεδομένα που αποθηκεύονται και διαχειρίζονται από διαφορετικούς πάροχους μπορεί να διαμοιραστούν και προσπελαστούν επιλεκτικά με ένα συνεργατικό τρόπο
  - Αυτό μπορεί να γίνει στα πλαίσια και διαφορετικών ιδιοκτητών δεδομένων
  - Σε αυτή την περίπτωση, η ανταλλαγή δεδομένων και οι συνεργατικοί υπολογισμοί πρέπει να ελεγχθούν ώστε να εξασφαλίσουν πως η πληροφορία δεν προσπελαύνεται, εκδίδεται ή εκτίθεται με αντικανονικό τρόπο
    - Πχ τα δεδομένα που έχουν αποθηκευθεί σε έναν πάροχο μπορεί επιλεκτικά να εκδοθούν μόνο σε συγκεκριμένους παρόχους και μέσα σε συγκεκριμένα πλαίσια (contexts)
  - Κάποιες λύσεις έχουν αντιμετωπίσει το πρόβλημα του ιδιωτικού και ασφαλούς υπολογισμού από πολλαπλά μέρη
    - Παρέχουν την ικανότητα για διαφορετικά μέρη να εκτελέσουν έναν συνεργατικό υπολογισμό μαθαίνοντας μόνο τα αποτελέσματα των επερωτήσεων και τίποτε για τις εισόδους αυτών

Συνεργατική Εκτέλεση Επερωτήσεων

- Στην ίδια γραμμή, υπάρχουν λύσεις για τον υπολογισμό κυρίαρχων joins πάνω από δεδομένα για την ανάκτηση αποτελεσμάτων από μια join λειτουργία πάνω σε διαφορετικούς πίνακες με την εξασφάλιση εμπιστευτικότητας της πληροφορίας που δεν ανήκει στο αποτέλεσμα
- Πρόσφατες προσεγγίσεις αντιμετωπίζουν ένα πιο γενικό σενάριο όπου διαφορετικά μέρη (κάτοχοι δεδομένων ή πάροχοι νέφους) πρέπει να συνεργαστούν και να διαμοιράσουν πληροφορία για την κατανεμημένη εκτέλεση μιας επερώτησης με επιλεγμένη έκθεση των δεδομένων
- Έχει επίσης μελετηθεί το πρόβλημα του προσδιορισμού ενός αποδοτικού και ασφαλούς πλάνου εκτέλεσης για τον υπολογισμό επερωτήσεων στο οποίο διαφορετικοί πάροχοι συνεργάζονται εκδίδοντας στους άλλους την εξουσιοδοτημένη πληροφορία που απαιτείται για τον υπολογισμό της απάντησης για την επερώτηση

Βιβλιογραφία

1. Ghorbel, A., Ghorbel, M. & Jmaiel, M. Privacy in cloud computing environments: a survey and research challenges. J Supercomput 73, 2763–2800 (2017). https://doi.org/10.1007/s11227-016-1953-y
2. Singh, N., Singh, A.K. Data Privacy Protection Mechanisms in Cloud. Data Sci. Eng. 3, 24–39 (2018). https://doi.org/10.1007/s41019-017-0046-0
3. Wang Y (2015) Privacy-preserving data storage in cloud using array BP-XOR codes. IEEE Trans Cloud Comput 3(4):425–436
4. Tal Rabin, Michael Ben-Or: Verifiable Secret Sharing and Multiparty Protocols with Honest Majority (Extended Abstract). STOC 1989: 73-85
5. Krawczyk, Hugo (1993). Secret Sharing Made Short. CRYPTO ’93
6. Resch, Jason; Plank, James (February 15, 2011). AONT-RS: Blending Security and Performance in Dispersed Storage Systems. Usenix FAST’11
7. Parakh, Abhishek; Kak, Subhash (January 2011). "Space efficient secret sharing for implicit data security". Information Sciences. 181 (2): 335–341. doi:10.1016/j.ins.2010.09.013

Βιβλιογραφία

8. Canard S, Devigne J (2016) Highly privacy-protecting data sharing in a tree structure. Future Gener Comput Syst 62:119–127
9. San Murugesan, Irena Bojanova. Encyclopedia of Cloud Computing. John Wiley & Sons, Ltd. 2016
10. Li Y et al (2016) Privacy preserving cloud data auditing with efficient key update. Future Gener Comput Syst 78:789–798
11. Chun-I Fan S-YH (2013) Controllable privacy preserving search based on symmetric predicate encryption in cloud storage. Future Gener Comput Syst 29:1716–1724
12. Liang K, Huang X, Guo F, Liu JK (2016) Privacy-preserving and regular language search over encrypted cloud data. IEEE Trans Inf Forensics Secur 11(10):2365–2376
13. Garcia-Molina H et al (2011) Data leakage detection. IEEE Trans Knowl Data Eng 23(1):51–63
14. Jesu Vedha Nayahi J, Kavitha V (2016) Privacy and utility preserving data clustering for data anonymization and distribution on Hadoop. Future Gener Comput Syst. https://doi.org/10.1016/j.future.2016.10.022
15. Zhang W et al (2016) Privacy preserving ranked multi-keyword search for multiple data owners in cloud computing. IEEE Trans Comput 65(5):1566–1578