

PAPER: Design of a multimedia object-oriented DBMS (Chen, Meliksetian, Chang, Liu)

REVIEWER: Αρκολάκης Δημήτριος

ΒΑΘΜΟΛΟΓΗΣΗ (κλίμακα: -3 έως 3)

- Αυθεντικότητα δημοσίευσης: 3
- Τεχνική αρτιότητα: 2
- Σημασία – Σημαντικότητα: 2
- Σαφήνεια παρουσίασης: 2
- Συνάφεια – Σχετικότητας: 3

ΔΙΑΡΚΕΙΑ:

- **Εντάξει X**
- Υπερβολική
- Ελλιπής

ΓΕΝΙΚΗ ΠΡΟΤΑΣΗ:

- Αποδοχή, άνευ αλλαγών
- **Αποδοχή, με μικρές αναθεωρήσεις X**
- Χρήζει δεύτερης αναθεώρησης
- Απορριπτέο

ΠΕΡΙΛΗΨΗ ΔΗΜΟΣΙΕΥΣΗΣ

Η παρούσα δημοσίευση δείχνει την ανάπτυξη πολυμεσικών λειτουργιών σε αντικειμενοστραφείς πολυμεσικές ΒΔ, την ενσωμάτωση και τον διαμοιρασμό πολυμέσων σε κατανομημένα περιβάλλοντα εργασίας, καθώς και τρόπους διαχείρισης βίντεο και εικόνων εντός συστήματος ΒΔ.

ΣΧΟΛΙΑΣΜΟΣ ΔΗΜΟΣΙΕΥΣΗΣ

Το paper αρχικά θέτει ως πιο σύγχρονη μέθοδο για ενσωμάτωση πολυμέσων σε ΒΔ τις **αντικειμενοστραφείς πολυμεσικές βάσεις**, έναντι των *σχεσιακών βάσεων με αντικειμενοστραφείς και πολυμεσικές προεκτάσεις* και των *αντικειμενοστραφών βάσεων με πολυμεσικές προεκτάσεις*. Βεβαίως, αναλύει τα υπέρ και κατά των δύο τελευταίων μεθόδων, αλλά **δεν επιχειρηματολογεί πειστικά** για τα πλεονεκτήματα (ή και τα μειονεκτήματα) της πρώτης μεθόδου σε σχέση με τις άλλες δύο, εντοπίζοντας κυρίως μόνο τις διαφορές τους.

Αναλύοντας το πολυμεσικό μοντέλο δεδομένων, οι συγγραφείς κάνουν **συνεχείς συγκρίσεις και παραλληλισμούς με το “κλασικό” σχεσιακό τύπο βάσεων δεδομένων**, προφανώς προς καλύτερη κατανόηση από τον αναγνώστη (πόσο μάλλον της εποχής εκείνης, 1995 σημ.) όπως και με το αντικειμενοστραφές μοντέλο, τονίζοντας ότι τα πολυμέσα στην προκειμένη, συμπεριφέρονται ως αντικείμενα. **Οι συσχετισμοί είναι αρκετά επιτυχημένοι** και κατανοητοί, τηρουμένων των αναλογιών και των δεδομένων της εποχής.

Αναλύεται διεξοδικά η αλληλεπίδραση των αντικειμένων σε κλασικές λειτουργίες (διαγραφή, καταχώρηση κλπ) ενώ **επεξηγούνται εν συντομία χρήσιμες έννοιες** όπως long και short objects, reference points, καθώς επίσης γίνεται και μία παρουσίαση των φορμών καταχώρησης των τότε γραφικών περιβάλλοντων εργασίας χρήστη, όπως και των hypermedia presentation effects. Ίσως οι συγκεκριμένες αναλύσεις **να ήταν καλύτερα να μετριάστουν**, μιας και τα μεν πρώτα φαίνεται να χρήσουν λίγο μεγαλύτερης ανάλυσης, ενώ τα δε άλλα κάπως μικρότερης και όχι τόσο διεξοδικής. Παρ'αυτα **το συνολικό μέγεθος της δημοσίευσης δεν φαίνεται να χρήζει αλλαγής**, μιας και αλληλομετριάζεται.

Στην περαιτέρω ανάλυση του πολυμεσικού μοντέλου, αναφέρονται οι απαιτήσεις βίντεο, εικόνας, ήχου και **αναλύονται οι τεράστιες απαιτήσεις** ταχύτητας, δικτύωσης, χώρου, δεδομένων κτλ των συνεχούς-ροής πολυμέσων (βίντεο, ήχος) για την υποστήριξη τους σε πολυμεσικές ΒΔ και τον διαμοιρασμό τους σε καταναμημένα περιβάλλοντα εργασίας. Παρατηρείται ότι αυτό **το θέμα είναι**, ουσιαστικά, **διαχρονικό** μιάς και όσο λύνονται αυτά τα προβλήματα ταχύτητας, χώρου με την πάροδο των εποχών τόσο γιγαντοποιούνται και οι απαιτήσεις των πολυμέσων με τα χρόνια!

Παραδειγματικά, στο paper αναφέρεται ο διαμοιρασμός ενός 240X200px video σε απομακρυσμένες διευθύνσεις, ενώ σήμερα ζούμε αντίστοιχα το live streaming 1024px hd video, διαμοιράζοντας το με παρόμοιες μεθόδους.

Στην συνέχεια του προηγούμενου, οι συγγραφείς δίνουν, ορθώς, **αρκετή βάση στο διαμοιρασμό πολυμέσων** παρουσιάζοντας “σύγχρονα” distributed collaboration services tools, αναλύοντας παράλληλα τους διαχωρισμούς και τις απαιτήσεις για βίντεο και εικόνα. Πάντως, **δεν αναλύεται ιδιαίτερα διεξοδικά το κομμάτι του ήχου**, πέραν της αναφοράς συγχρονισμού εικόνας-ήχου σε ένα βίντεο. Επίσης **οι παραγράφοι των collaboration tools είναι κάπως πτωχές**, ιδιαίτερα στα παραδείγματα.

Ένα σημείο στο οποίο θα ήθελα να σταθώ, πριν κλείσω, είναι τα **πολύ καλά παραδείγματα επεξηγηματικών εικόνων** που εμπεριέχονται στην δημοσίευση. Μέσα στο “χαος” του κειμένου και της **συχνής επανάληψης διάφορων όρων** (όπως multimedia, DB, video, audio κτλ) οι εικόνες και στα σχεδιαγράμματα είναι αρκετά κατατοπιστικά και “ξεθολώνουν” αρκετά το “τοπίο” της ανάγνωσης. **Συγκεκριμένα**, το σχ.13 για το εργαλείο διαμοιρασμού βίντεο, η εικ.15 για τα αρχεία εικόνας και οι εικ.4-5 για τις φόρμες γραφικών περιβάλλοντων βάσεων είναι αρκετά κατατοπιστικές για τις “ελλειψεις” παραγράφων στις οποίες αντιστοιχούν. Δυστυχώς, δεν μπορώ να πω το ίδιο για τις 2-3 και 6-7-8 αντίστοιχα.

Συμπερασματικά, πρόκειται για ένα paper αρκετά **σύγχρονο** για την εποχή του αλλά και **διαχρονικό**, με πρακτικές ακόμα και τωρινές. Μεθόδοι τεμαχισμού πολυμεσικών δεδομένων που αναφέρονται **βρήκαν εφαρμογή και αργότερα** που απλοποιήθηκαν οι τεράστιοι όγκοι δεδομένων ανταλλαγής.

ΠΡΟΤΑΣΕΙΣ ΒΕΛΤΙΩΣΕΩΝ:

- Στα ερωτήματα: “Τι γίνεται σε περίπτωση corrupted αρχείου ή διακοπής σύνδεσης κατά τη μεταφορά μέρους ενός τεμαχισμένου μεγάλου πολυμεσικού αρχείου όπως βίντεο;” δεν βρήκα απαντήσεις! Θέματα τα οποία απασχόλησαν αρκετά τους επιστήμονες τα επόμενα χρόνια.
- Οι τελευταίες παραγράφοι των collaboration tools έρχονταν μεγαλύτερης επέκτασης. Παρόλο που γενικά δίνει αρκετή βάση στα distributed collaboration services, οι τελευταίες δύο παράγραφοι του κεφαλαίου 4 είναι σχετικά πτωχές.
- Δεν εμβαθύνει καθόλου στο θέμα του ήχου και δεν αναφέρεται καν κάποιο audio collaboration tool. Χρειαζόταν κάποιας αναφοράς.
- Παρόλο που αναλύεται διεξοδικά η αλληλεπίδραση των αντικειμένων σε κλασικές

PAPER: A scalable continuous query system for internet databases (Chen, DeWitt, Tian, Wang)

REVIEWER: Αρκολάκης Δημήτριος

ΒΑΘΜΟΛΟΓΗΣΗ (κλίμακα: -3 έως 3)

- Αυθεντικότητα δημοσίευσης: 3
- Τεχνική αρτιότητα: 3
- Σημασία – Σημαντικότητα: 2
- Σαφήνεια παρουσίασης: 2
- Συνάφεια – Σχετικότητας: 2

ΔΙΑΡΚΕΙΑ:

- **Εντάξει X**
- Υπερβολική
- Ελλιπής

ΓΕΝΙΚΗ ΠΡΟΤΑΣΗ:

- Αποδοχή, άνευ αλλαγών
- **Αποδοχή, με μικρές αναθεωρήσεις X**
- Χρήζει δεύτερης αναθεώρησης
- Απορριπτέο

ΠΕΡΙΛΗΨΗ ΔΗΜΟΣΙΕΥΣΗΣ

Η παρούσα δημοσίευση παρουσιάζει το Niagara CQ, ένα σύστημα συνεχών ερωτημάτων για διαδικτυακές βάσεις δεδομένων, βασισμένο σε ομαδοποίηση των ερωτημάτων στη λογική της ομοιογένειας τους. Στο paper οι συντάκτες προτείνουν νέες μεθοδολογίες ομαδοποίησης, θεωρώντας τις πιο επεκτάσιμες, καθώς και λιγότερο τροποποιημένες προς τους κοινώς αποδεκτούς μηχανισμούς ερωτημάτων.

ΣΧΟΛΙΑΣΜΟΣ ΔΗΜΟΣΙΕΥΣΗΣ

Το paper αρχικά αναφέρει τον **τρόπο λειτουργίας** του μηχανισμού Niagara και το **πώς ανταπεξέρχεται στο μεγάλο πλήθος των ερωτημάτων** μέσω της ομαδοποίησης ανά ομοιογενείς δομές μοιράζοντας έτσι τον κοινό υπολογισμό και “αφαιρώντας” τεράστιο αριθμό “άχρηστων – επαναλαμβανόμενων” πληροφοριών στα ερωτήματα.

Κατά τις εισαγωγικές παραγράφους αναφέρονται **οι απαιτήσεις ενός συνεχούς ερωτήματος** και τα πλεονεκτήματα ομαδοποίησης συνεχών ερωτημάτων. **Δεν αναφέρονται** πιθανά μειονεκτήματα. Παράλληλα **αναφέρει τις καλύτερες μεθόδους ομαδοποίησης**, κάνοντας σύντομη σύγκριση με παλιότερες – διαφορετικές μεθόδους παλαιότερων. Η σύγκριση δεν συνοδεύεται από πειστικά παραδείγματα ανωτερότητας των προτεινόμενων μεθόδων. Επίσης **δεν παραλληλίζει τα περιβάλλοντα** για τα οποία χρησιμοποιούντουσαν οι παλιότερες μέθοδοι ομαδοποίησης, με αυτά που χρησιμοποιούνται οι νεότερες - προτεινόμενες, για να κατανοήσει ο αναγνώστης το μεγάλο όφελος των νεότερων.

Δίνεται βάση, επίσης, στη **δενδροειδή δομή σχηματισμού** του μηχανισμού (αναφέροντας μόνο τα πλεονεκτήματα και αποφεύγοντας αναφορές σε μειονεκτήματα ή συγκρίσεις), αναφέρονται οι **κατηγορίες συνεχών ερωτημάτων** (και σύγκριση τους) καθώς και ένα σύντομο ιστορικό του Niagara. Στην επεξήγηση των ιδιοτήτων του μηχανισμού, δεν αναφέρεται πούθενά αν υπάρχει **απώλεια στο θέμα της πληροφορίας** (lossy ή lossless) καθώς επίσης δεν αναφέρονται οι **απαιτήσεις μόνιμης μνήμης** και δεν αναλύονται και τόσο διεξοδικά οι απαιτήσεις προσωρινής.

Η δημοσίευση χαρακτηρίζεται από την **διεξοδική ανάλυση στο τεχνικό κομμάτι** από το κεφάλαιο 3 και μετά. Υπάρχουν **εκτενή παραδείγματα**, οι παράγραφοι είναι αρκετά επεξηγηματικές και οι **εικόνες πρακτικές** και με κατανοητή αναπαράσταση. Μία παρατήρηση, ίσως, υπάρχει **στο κομμάτι της γλώσσας εντολών** όπου η περιγραφή θεωρείται ελλιπής, μιας και προφανώς η γλώσσα αυτή δεν περιορίζεται σε “create” και “delete”. Υπάρχει και απουσία πιο πρακτικών παραδειγμάτων. Κατα τα άλλα **ο αναγνώστης μπορεί να καταλάβει πληρέστατα τις τεχνικές query split, grouping κλπ** στον συγκεκριμένο μηχανισμό. Παρ'αυτά η τεκμηρίωση στο “Memory Caching κρίνεται ως μη πειστική”.

Οι συγγραφείς **προσπαθούν να αποδείξουν τις δυνατότητες του Niagara με πειραματικό τρόπο**. Από την μία αποδεικνύει ότι ναι μεν υπάρχουν δυνατότητες στον παρόν μηχανισμό, **αλλά δεν τρέχει πειραματικά** με παρόμοιο τρόπο τα ίδια πειράματα **σε άλλους μηχανισμούς** και μεθόδους για να έχουμε ένα μέτρο σύγκρισης αποτελεσμάτων. Τέλος, πολύ αναφορά γίνεται σε **“μελλοντική δουλειά”** σε πολλά σημεία του paper, αφήνοντας κάποια κενά για το... μέλλον!

ΠΡΟΤΑΣΕΙΣ ΒΕΛΤΙΩΣΕΩΝ:

- Η γενική εικόνα του paper μοιάζει πολύ “διαφημιστική” και ελάχιστα “εγκυκλοπαιδική”. Στην προσπάθεια παρουσίασης της μεθοδολογίας τους, οι συγγραφείς, αποφεύγουν επαιδικτικά να αναφέρουν οποιαδήποτε μειονεκτήματα ή να προβούν σε οιοσδήποτε συγκρίσεις. Συγκρίσεις και αναφορά πιθανών μειονεκτημάτων επιβάλλονται.
- Παρόλο που αναφέρεται πολλάκις το θέμα “μνήμης” και κατά πόσο “γλυτώνουμε μνήμη” από την ομαδοποίηση υπολογισμών, εντούτοις ελάχιστες ουσιαστικές λεπτομέρειες σχετικά με το ζήτημα διαβάζουμε. Χρειάζοντουσαν “νούμερα” ή έστω αλγοριθμική επεξήγηση.
- Η λέξη “future work” αναφέρεται σε κάποια σημεία του εγγράφου αφήνοντας... μετέωρη αυτήν την “μελλοντική εργασία” αποφεύγοντας λεπτομέρειες για τα χαρακτηριστικά της και για την ενδεχόμενη συμβατότητα του συστήματος τους με μελλοντικές τεχνολογίες, επιχειρηματολογώντας.

PAPER: t-Closeness: Privacy Beyond k-Anonymity and l-Diversity
(Ninghui Li, Tiancheng Li, Suresh Venkatasubramanian)

REVIEWER: Αρκολάκης Δημήτριος

ΒΑΘΜΟΛΟΓΗΣΗ (κλίμακα: -3 έως 3)

- Αυθεντικότητα δημοσίευσης: 3
- Τεχνική αρτιότητα: 2
- Σημασία – Σημαντικότητα: 2
- Σαφήνεια παρουσίασης: 3
- Συνάφεια – Σχετικότητα: 3

ΔΙΑΡΚΕΙΑ:

- Εντάξει
- Υπερβολική
- **Ελλειπής X**

ΓΕΝΙΚΗ ΠΡΟΤΑΣΗ:

- **Αποδοχή, άνευ αλλαγών X**
- Αποδοχή, με μικρές αναθεωρήσεις
- Χρηρίζει δεύτερης αναθεώρησης
- Απορριπτέο

ΠΕΡΙΛΗΨΗ ΔΗΜΟΣΙΕΥΣΗΣ

Το παρόν Paper μας διευκρινίζει τους στόχους του από την παράγραφο της περίληψης: Να δείξει τους περιορισμούς του l-Diversity, να προτείνει το t-closeness επιλέγοντας το EMD ως μέθοδο μέτρησης και να συζητηθεί η λογική του μέσω πειραμάτων και παραδειγμάτων.

ΣΧΟΛΙΑΣΜΟΣ ΔΗΜΟΣΙΕΥΣΗΣ

Η παρούσα δημοσίευση καταπιάνεται με την ιδιωτικότητα στα δεδομένα και συγκεκριμένα με τις έννοιες t-Closeness, k-Anonymity & l-Diversity τις οποίες και μας τις ορίζει σύντομα αλλά και συγχρόνως κατανοητά από την εισαγωγή. Το εισαγωγικό κλίμα είναι κατατοπιστικότατο και στο “κοινωνικό” κομμάτι του θέματος, εξηγώντας με απλά λόγια πόσο σημαντική είναι η ιδιωτικότητα των δεδομένων, τι εστί ευαίσθητη πληροφορία καθώς και τι μπορούμε να κάνουμε για την προστασία αυτών.

Εν συνέχεια το paper επεξηγεί αναλυτικότερα τις έννοιες με τις οποίες καταπιάνεται, αφιερώνοντας περίπου ένα κεφάλαιο στην κάθε μια. Οι έννοιες γίνονται αρκετές κατανοητές μέσω της απλής επεξήγησης και του γενικά λιτού τρόπου που έχει το paper στην ανάπτυξη του. Τέλος, δίνονται οι τρόποι υπολογισμού του EMD καθώς και η συσχέτιση του ως μέτρο μέτρησης και γίνονται και τα πειράματα απόδειξης της λογικής του t-closeness.

Η δημοσίευση προτείνει το t-closeness ως λύση στα “κενά” και στα προβλήματα που υπάρχουν στο l-diversity. Μέσα στο paper υπάρχουν αρκετά επεξηγηματικά παραδείγματα και εικόνες τόσο για τα “αλλαγμένα” datasets, όσο για τους μαθηματικούς τύπους που χρησιμοποιούνται για τους υπολογισμούς καθώς και για τα πειράματα και τα γραφήματα αναπαράστασης.

Εν κατακλείδι, θα λέγαμε ότι είναι ένα ολοκληρωμένο σύγχρονο paper με λιτό τρόπο παρουσίασης, σύντομη διάρκεια και σχετική “λακωνικότητα”. Μπορούμε να πούμε ότι, έτσι όπως είναι στημένο, θα μπορούσε να χρησιμοποιηθεί και ως διδακτικό υλικό σε πανεπιστημιακή αίθουσα.

ΠΡΟΤΑΣΕΙΣ ΒΕΛΤΙΩΣΕΩΝ:

- Ο λιτός τρόπος παρουσίασης, ίσως να είναι βολικός προς άτομα χωρίς ιδιαίτερη εμπειρία στον συγκεκριμένο τομέα (όπως εγώ ο ίδιος που γράφω αυτήν την κριτική) αλλά ως “επιστημονική δημοσίευση” ίσως θα έπρεπε (με κάθε επιφύλαξη) να χρησιμοποιεί λίγο πιο εξειδικευμένη ορολογία.
- Το Paper χαρακτηρίζεται από σχετικά μικρή διάρκεια σε σχέση με τα “πολλά” που μας λέει. Βεβαίως, αυτό δεν είναι “κακό” μιας και στην τελική τον σκοπό της, ως δημοσίευση, τον καταφέρνει. Πάντως, θα μπορούσε να γίνει λίγο παραπάνω σχολαστική και μεγαλύτερης διάρκειας, σε γενικές γραμμές.
- Το πείραμα με το “similarity attack” είναι μεν επαρκές, αλλά ίσως να χρειαζόταν και κάτι