Θέματα για πτυχιακές εργασίες

Θέμα διπλωματικής	Περιγραφή διπλωματικής	Παραδοτέα	Απαραίτητες γνώσεις
Επεκτάσεις και βελτιώσεις του Sodasense	Στο εργαστήριο συστημάτων λογισμικού και βάσεων δεδομένων έχει υλοποιηθεί μια πλατφόρμα συλλογής και ανάλυσης χωροχρονικών δεδομένων. Η πλατφόρμα περιλαμβάνει συστήματα που συλλέγουν τα χωροχρονικά δεδομένα (π.χ., από μετεωρολογικούς σταθμούς, από κινητά τηλέφωνα, από σταθμούς καλλιέργειας) και συστήματα που επεξεργάζονται και παρουσιάζουν της πληροφορία αυτή. Παράλληλα υπάρχει όλη η υποδομή που υποστηρίζει του χρήστες την ανωνυμία τους και την ασφάλεια των δεδομένων που αποθηκεύουν. Στα πλαίσια της εργασίας αυτή μπορούν να προταθούν και να υλοποιηθούν νέες λειτουργίες και υπηρεσίες ή να βελτιωθούν ήδη υπάρχουσες.	Σύστημα και έντυπη εργασία	Προγραμματισμός, συστήματα διαχείρισης βάσεων δεδομένων
Οπτικοποίηση γεωχρονικών σειρών δεδομένων	Ο στόχος της παρούσας πτυχιακής είναι η δημιουργία διαδραστικής εφαρμογής για την οπτικοποίηση γεωχρονικών σειρών δεδομένων. Η εφαρμογή θα χρησιμοποιεί υβριδικούς δείκτες που υπολογίζουν χωρικές αλλά και χρονικές συνόψεις. Οι παραπάνω συνόψεις θα χρησιμοποιηθούν για να παραχθούν οπτικοποιήσεις σε χάρτες που θα επιτρέπουν στους χρήστες να επεξεργαστούν τα δεδομένα και στα δύο πεδία τους.	Σύστημα και έντυπη εργασία	java, python
Δεικτοδότηση γεωχρονικών σειρών δεδομένων σε δίσκο (Dumpy)	Time series data is probably the most common type of modern information. Time series include trajectory data, sensor data (for climate, agriculture etc.), web information and even medical data that typically grow in vast sizes. To handle and process data series efficiently requires data indexing. Dumpy is a new time series index that claims to be much more efficient than the established iSAX time series index. The purpose of this thesis is to directly compare Dumpy and iSAX. The first step is to port the C++ implementation of Dumpy to Python. Then, we will compare the performance of this implementation with the existing Python implementation of iSAX. Finally, we will consider methods to improve the performance of iSAX based on the Dumpy implementation and methods to improve Dumpy.	Σύστημα και έντυπη εργασία	Python, Pandas, NumPy, C++.
Δεικτοδότηση γεωχρονικών σειρών δεδομένων σε δίσκο	Ο στόχος της παρούσας πτυχιακής είναι η πειραματική αξιολόγηση ενός υβριδικού δείκτη που χρησιμοποιείται για την αποτίμηση γεωχρονικών σειρών δεδομένων. Στη πτυχιακή περιλαμβάνεται η μετασχηματισμός του δείκτη ώστε να υποστηρίζει αποθήκευση των φύλλων του στο δίσκο. Εκεί τα δεδομένα μπορούν να αποθηκευτούν με τρεις διαφορετικούς τρόπους: ως απλά δυαδικά αρχεία, σε ένα σύστημα σχεσιακών βάσεων δεδομένων ή σε ένα σύστημα αποθήκευση κολόνων (column store). Ο σκοπός της εργασίας είναι η πειραματική μελέτη και αξιολόγηση των παραπάνω εναλλακτικών.	Σύστημα και έντυπη εργασία	java, javascript, PHP, python

Θέμα διπλωματικής	Περιγραφή διπλωματικής	Παραδοτέα	Απαραίτητες γνώσεις
Αποδοτική διαχείριση συσχετίσεων κυβερνο-απειλών	Στις μέρες μας, πληροφορίες για κυβερνο-απειλές συγκεντρώνονται συνεχώς από μια μεγάλη ποικιλία πηγών, αναλύονται (αυτόματα ή από ειδικούς) και αποθηκεύονται σε κατάλληλα συστήματα όπως για παράδειγμα το MiSP. Η εργασία αυτή θα επικεντρώσει το ενδιαφέρον της στην αναζήτηση και αποθήκευση των συσχετισμών μεταξύ των διάφορων κυβερνο-απειλών. Οι συσχετισμοί αυτοί σχηματίζουν ένα γράφο με κόμβους τις απειλές και ακμές τις συσχετίσεις τους. Τα υπάρχοντα συστήματα (τόσο το MiSP όσο και άλλα) αποθηκεύουν τον εν λόγω γράφο χρησιμοποιώντας παραδοσιακές σχεσιακές μεθόδους. Οι παραπάνω τεχνικές εμφανίζουν όμως σημαντικά μειονεκτήματα οπότε με την παρούσα εργασία θέλουμε να εξετάσουμε πιο σύγχρονες μεθόδους και συστήματα που εξειδικεύονται στην αποθήκευση δεδομένων γράφων (όπως για παράδειγμα τα Neo4j, ArangoDB, Virtuoso, Dgraph).	Σύστημα και έντυπη εργασία	MySQL, Neo4j, ArangoDB
Πειραματική επισκόπηση συστημάτων αποθήκευσης και διαχείρισης κυβερνο-απειλών	Στις μέρες μας ο όγκος των δεδομένων που αφορούν κυβερνο-απειλές και κυβερνο-επιθέσεις αυξάνεται με τεράστιους ρυθμούς. Για την διαχείριση των δεδομένων αυτών έχουν προταθεί πολλά συστήματα και πλατφόρμες διαχείρισης και επεξεργασίας. Ο σκοπός της παρούσας εργασίας είναι η μελέτη, παρουσίαση και αξιολόγηση τέτοιων συστημάτων. Η αξιολόγησή τους θα βασίζεται σε πραγματικά δεδομένα κυβερνο-απειλών και σε πραγματικές συνθήκες χρήσης και ανάλυσης.	Σύστημα και έντυπη εργασία	Προγραμματισμός, συστήματα διαχείρισης βάσεων δεδομένων
Αντικατάσταση συστήματος διαχείρισης δεδομένων στο MISP	Το Misp είναι μια πλατφόρμα ανοικτού κώδικα που αναλαμβάνει την διαχείριση, την αποθήκευση και τον διαμοιρασμό κυβερνο-απειλών. Στο πυρήνα του Misp υπάρχει το σύστημα διαχείρισης βάσεων δεδομένων MySQL. Δυστυχώς οι η MySQL σε αρκετές περιπτώσεις δεν επαρκεί για να χειριστεί τον όγκο των πληροφοριών που αποθηκεύονται. Ο στόχος της παρούσας πτυχιακής εργασίας είναι η αντικατάσταση την MySQL με ένα πιο αποδοτικό σύστημα διαχείρισης βάσεων δεδομένων και η πειραματική αξιολόγηση του αποτελέσματος σε συνθήκες πραγματικής λειτουργίας.	Σύστημα και έντυπη εργασία	PHP, συστήματα διαχείρισης βάσεων δεδομένων
Έλεγχος κόστους τιμών παροχής ηλεκτρικού ρεύματος σε εύρος χρονικού διαστήματος	Ο σκοπός της πτυχιακής εργασίας είναι να συγκεντρώνει από την σελίδα της ρυθμιστικής αρχής (ΡΑΕΕ) τις τιμές παροχής ηλεκτρικού ρεύματος που θέτει κάθε πάροχος. Θα αποθηκεύει τις τιμές αυτές σε μια τοπική βάση δεδομένων και θα υπολογίζει το κόστος που θα πλήρωνε ο καταναλωτής σε κάθε πάροχο αλλά για οποιοδήποτε διάστημα επιλέξει ο χρήστης.	Σύστημα και έντυπη εργασία	web crawling, διαχείριση δεδομένων