

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Σαββάκη Σοφία

Μεταπτυχιακή Φοιτήτρια

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτ. Εργασίας: Καθηγητής Π. Τσακαλίδης

Παρασκευή, 4/11/2016, 10:00

Αίθουσα B108, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

" Καινοτόμες τεχνικές για την εκτίμηση πολυτροπικών χαμένων δεδομένων σε ασύρματα δίκτυα αισθητήρων "

ΠΕΡΙΛΗΨΗ

Κατά τη διάρκεια των τελευταίων δεκαετιών, τα ασύρματα δίκτυα αισθητήρων έχουν προσελκύσει μεγάλο ενδιαφέρον καθώς αποτελούν τη βασική τεχνολογία που επιτρέπει την υλοποίηση εξελιγμένων υπηρεσιών σε πολλαπλούς τομείς εφαρμογών, οι οποίοι περιλαμβάνουν την παρακολούθηση του περιβάλλοντος, ενός βιομηχανικού συστήματος και της ατομικής υγείας. Όμως, παρά την ευρεία εφαρμογή τους, τα ασύρματα δίκτυα αισθητήρων πάσχουν από δικτυακές και ενεργειακές ατέλειες, που αναπόφευκτα οδηγούν σε απώλεια μετρήσεων. Ως αποτέλεσμα, ο χαμηλός όγκος από διαθέσιμα δεδομένα επηρεάζει δραματικά τις επακόλουθες εργασίες επεξεργασίας και μάθησης, όπως η ανίχνευση ασυνήθιστων γεγονότων, η ομαδοποίηση και η ταξινόμηση.

Στην παρούσα διπλωματική εργασία αντιμετωπίζουμε το πρόβλημα αυτό, προτείνοντας τις μη συμβατικές μεθόδους της Συμπλήρωσης α) Πινάκων (ΣΠ) και β) Τανυστών (ΣΤ), οι οποίες είναι σε θέση να ανακτούν τις απύσυχες μετρήσεις με έναν αξιόπιστο τρόπο. Η βασική θεωρητική αρχή που υιοθετείται σε αυτές τις τεχνικές, είναι ότι ένα σύνθετο σήμα μπορεί να ανακτηθεί από ένα

μικρό αριθμό τυχαίων μετρήσεων, αξιοποιώντας τις υποκείμενες συσχετίσεις ανάμεσα στα δεδομένα. Ωστόσο, η υπόθεση αυτή δεν ικανοποιείται σε πραγματικά, και συχνά θορυβώδη, σύνολα δεδομένων, τα οποία τείνουν να είναι μεγάλης τάξεως. Αντιμετωπίζουμε αυτόν τον περιορισμό, εισάγοντας την ιδέα της κατάλληλης διαμόρφωσης των διαθέσιμων ροών δεδομένων σε δισδιάστατες και τρισδιάστατες δομές χαμηλής τάξης, η οποία επιτρέπει την χρήση των ΣΠ και ΣΤ στον τομέα των ασύρματων δικτύων αισθητήρων.

Για να δοκιμάσουμε την αποτελεσματικότητα της προσέγγισής μας, πειραματιζόμαστε σε δύο εξέχοντες τομείς και συγκεκριμένα στην έξυπνη διαχείριση του νερού μέσω των ασύρματων δικτύων αισθητήρων και στην αναγνώριση της ανθρώπινης δραστηριότητας. Συνθέτουμε τα αντίστοιχα συστήματα επεξεργασίας και ταξινόμησης, τα οποία ενσωματώνουν τις προτεινόμενες μεθόδους μας για δειγματοληψία, δόμηση και ανάκτηση των δεδομένων. Τα εν λόγω συστήματα αξιολογούνται κατά διάφορες πτυχές, που σχετίζονται με την ποιότητα της ανακατασκευής για ποικίλα ποσοστά ελλειπουσών τιμών, την ακρίβεια της μετέπειτα ανάλυσης (π.χ. ταξινόμηση) και την επίδραση της υπό-δειγματοληψίας στη διάρκεια ζωής του δικτύου. Η ανάλυσή μας τονίζει την αλληλεπίδραση μεταξύ διαφόρων σεναρίων ανάκαμψης, από την οπτική της δόμησης και της προέλευσης των δεδομένων, και πολλαπλών σύγχρονων ταξινομητών. Τα αποτελέσματα επιδεικνύουν ότι, οι μέθοδοι που έχουν αναπτυχθεί επιτυχάνουν υψηλή ακρίβεια ανακατασκευής, ακόμη και για περιπτώσεις εξαιρετικής υπο-δειγματοληψίας πολυτροπικών ροών δεδομένων, στις οποίες έχει χαθεί μέχρι και το 80% των μετρήσεων.

Savvaki Sofia

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Professor P. Tsakalides

Friday, 4/11/2016, 10:00

Room B108, Computer Science dept., University of Crete

“Novel techniques for the estimation of multi-modal missing data in wireless sensor networks”

ABSTRACT

Over the last decades Wireless Sensor Networks (WSNs) have attracted great attention, as they constitute a key enabling technology for implementing sophisticated services in numerous application domains, including area and environmental sensing, health care monitoring, and industrial control systems. Despite their wide applicability, WSNs suffer from network and energy imperfections, which inevitably often lead to missing measurements. The resulting low volume of

available data dramatically affects subsequent processing and learning tasks, such as detection of unusual events, clustering, and classification.

In this thesis, we address the problem of missing WSN data by employing two non-conventional techniques, which are capable of recovering measurements in a reliable fashion, namely: a) Matrix Completion (MC), and b) Tensor Completion (TC). The key theoretical principle adopted is that a complex signal can be recovered from a small number of random measurements, by exploiting the underlying redundancies of the sensing data. However, this assumption is not satisfied in real-life, and often, noisy datasets, which tend to be full rank. We tackle this limitation by introducing the concept of appropriately forming the available data streams into low-rank 2D and 3D structures, thereby enabling the utilization of MC and TC in the WSN domain.

To test the efficacy of our approach, we experiment on two prominent fields, namely WSN-based Smart Water Management (SWM) and Human Activity Recognition (HAR). We synthesize their respective processing and classification frameworks, which encapsulate our proposed modules for data sampling, structuring, and recovery. These frameworks are evaluated against numerous aspects, related to the quality of reconstruction on different volumes of missing data, the accuracy of subsequent analysis (e.g. classification), and the impact of sub-sampling on the network's lifetime. Our analysis highlights the interaction of different recovery scenarios in terms of data structuring and origin, with several state-of-the-art classifiers. The results demonstrate that high reconstruction accuracy can be achieved through the developed modules, even for the case of extremely under-sampled, multi-modal streams of data, lacking up to 80% of their measurements.