

## Αποδοτική Επεξεργασία Ερωτήσεων σε Δίκτυα Αισθητήρων

*Αντώνιος Δεληγιαννάκης (University of Maryland)*  
*Ιωάννης Κωτίδης (AT&T Labs-Research)*

## Σχεδιάγραμμα Ομιλίας

- ◆ Εισαγωγικά
  - Αισθητήρες (sensors)
  - Εφαρμογές, χαρακτηριστικά
  - Μέθοδοι περιορισμού μεταδιδόμενων δεδομένων
- ◆ Συμπύεση Ιστορικών Μετρήσεων
- ◆ Προσεγγιστικές Συνεχείς Συναθροιστικές Ερωτήσεις
- ◆ Στιγμιότυπα Δικτύου
- ◆ Συμπεράσματα

## Αισθητήρες: Τι είναι?

- ♦ Αντικείμενα που μετράνε/παρατηρούν το περιβάλλον τους
  - Μετεωρολογικά
    - Θερμοκρασία, Πίεση, Υγρασία...
  - Πεδίο Μάχης
    - Θόρυβος, Κίνηση/Θέση Αντικειμένων, Χημικά
  - Μετρήσεις Υγείας
    - Παλμοί καρδιάς (HR), κορεσμός οξυγόνου (SpO2), δεδομένα ηλεκτροκαρδιογραφήματος (EKG)
- ♦ Μεγάλη ποικιλία μοντέλων, ανάλογα με εφαρμογή

26/08/2005

Αντόνιος Δεληγιαννάκης

3

## Τμήματα ενός Αισθητήρα



	Berkeley Mica2	Stargate (Intel PXA255 cpu)
CPU	7.38 MHz	400 MHz
Μνήμη	4KB SRAM, 512 KB EEPROM	μέχρι 256 MB FLASH
Ασύρματος	300 μέτρα	Εξαρτάται από το μοντέλο
Μπαταρία	2 AA	Li-Ion

26/08/2005

Αντόνιος Δεληγιαννάκης

4

## Προτεινόμενες Εφαρμογές

---

- ♦ Μετεωρολογικές μετρήσεις
- ♦ Παρακολούθηση οικοσυστήματος
- ♦ Ανίχνευση αντικειμένων
- ♦ Επίβλεψη μηχανημάτων
- ♦ Επίβλεψη κατάστασης κτισμάτων (σεισμοί)
- ♦ Ιατρικά δεδομένα
  - Στρατιώτες, ασθενείς επειγόντων
- ♦ Περιοχές καταστροφής
- ♦ Ποιότητα νερού

26/08/2005

Αντόνιος Δεληγιαννάκης

5

## Προκλήσεις/Περιορισμοί

---

- ♦ Περιορισμοί ενέργειας
  - 3-5% ετήσια αύξηση χωρητικότητας
  - Συχνά μη επιβλεπούμενη τοποθέτηση
- ♦ Περιορισμοί χωρητικότητας δικτύου (bandwidth)
  - Ασύρματη επικοινωνία
  - Ασύρματος ανοικτός για μικρό διάστημα
- ♦ Μετάδοση = κύρια πηγή κατανάλωσης ενέργειας
  - Μετάδοση 1 byte απαιτεί περισσότερη ενέργεια από χιλιάδες εντολές CPU
  - Περιορισμένες δυνατότητες ασυρμάτων, μετάδοση μέσω πολλών κόμβων

26/08/2005

Αντόνιος Δεληγιαννάκης

6

## Στόχος Έρευνάς μας

- ♦ Κοινός στόχος: Μείωση δεδομένων στο δίκτυο
  - Συμπίεση/περιορισμό μεταδιδόμενων δεδομένων
  - Μείωση κόμβων που χρησιμοποιούνται σε ερωτήσεις
    - Αντιπρόσωποι υπολογίζουν εκτιμήσεις για γειτονικούς κόμβους
- ♦ Ερωτήσεις με διαφορετικά χαρακτηριστικά
  - Μετάδοση ιστορικών μετρήσεων
    - Περιοδική μετάδοση / απόκριση σε ερωτήσεις
    - Μέγεθος δεδομένων μπορεί να είναι μεγάλο
      - ♦ “Μετάδωσε όλες τις μετρήσεις της τελευταίας ώρας”
  - Συνεχείς συναθροιστικές ερωτήσεις
    - Συνεχής παρακολούθηση στατιστικών τιμών από μετρήσεις
    - Συχνές μεταδόσεις, μικρός όγκος δεδομένων / μετάδοση
      - ♦ “Ανέφερε τον αριθμό παρατηρούμενων αντικειμένων κάθε 1 sec”

26/08/2005

Αντώνιος Δεληγιαννάκης

7

## Σχεδιάγραμμα Ομιλίας

- ♦ Εισαγωγικά
- ♦ Συμπίεση Ιστορικών Μετρήσεων
- ♦ Προσεγγιστικές Συνεχείς Συναθροιστικές Ερωτήσεις
- ♦ Στιγμιότυπα Δικτύου
- ♦ Συμπεράσματα

26/08/2005

Αντώνιος Δεληγιαννάκης

8

## Συμπίεση Ιστορικών Μετρήσεων

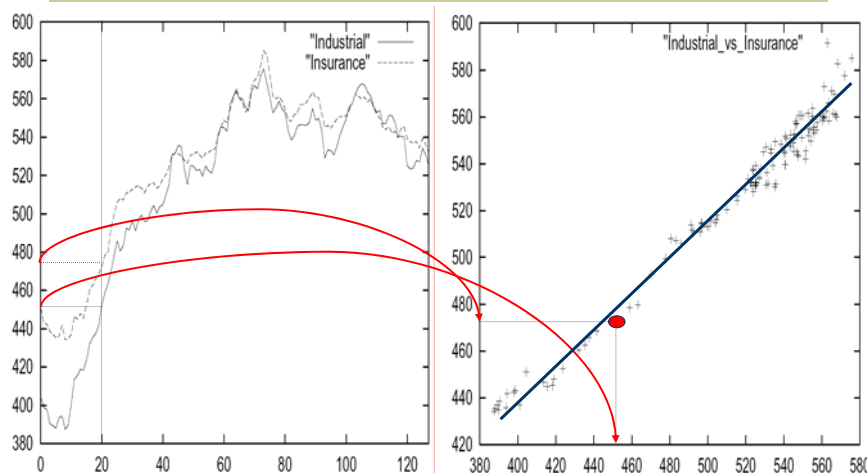
- ♦ Αισθητήρες συχνά συλλέγουν μετρήσεις για διαφορετικές ποσότητες
- ♦ Πιθανές συσχετίσεις σε συλλεγόμενες τιμές
  - Στην ίδια ποσότητα
    - Περιοδικότητα, παρόμοιες τάσεις
  - Μεταξύ διαφορετικών ποσοτήτων
    - Θερμοκρασία και τάση [Deshpande04], πίεση και υγρασία
  - Μεταξύ κόμβων σε μία περιοχή
    - Παρόμοιες θερμοκρασίες ή επίπεδα θορύβου
- ♦ Πώς να εκμεταλλευτούμε αυτές τις συσχετίσεις?

26/08/2005

Αντόνιος Δεληγιαννάκης

9

## Παράδειγμα Συσχετισμένων Σημάτων



- ♦ Παλινδρόμηση σε XY επίπεδο = κλιμάκωση και μεταφορά

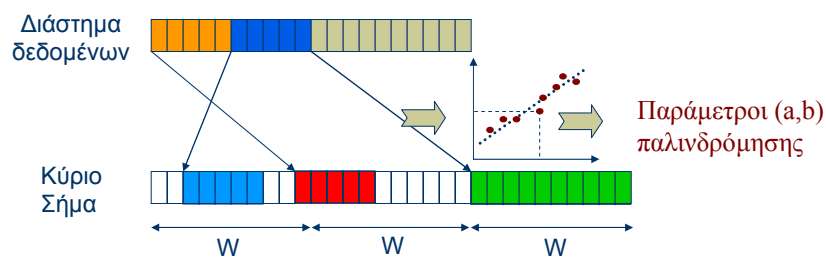
26/08/2005

Αντόνιος Δεληγιαννάκης

10

## Κύρια Ιδέα

- ♦ Δημιούργησε μικρό λεξικό τμημάτων των δεδομένων που εμφανίζονται συχνά (κύριο σήμα)
- ♦ Διαχώρισε δεδομένα σε διαστήματα
- ♦ Κωδικοποίησε κάθε διάστημα μέσω του λεξικού
  - Χρήση γραμμικής παλινδρόμησης :  $\vec{Y} = a \vec{X} + b$

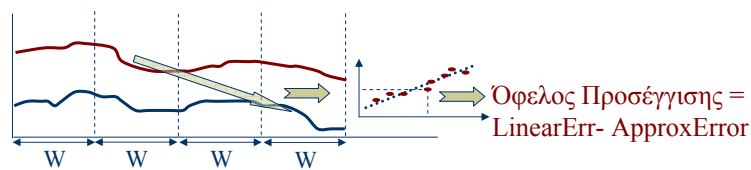


26/08/2005

Αντόνιος Δεληγιαννάκης

11

## Λεπτομέρειες Αλγορίθμου



- ♦ Υπολόγισε όφελος προσέγγισης διαστήματος  $j$  χρησιμοποιώντας το  $i$ 
  - Μείωση λάθους αντί για απλή γραμμική παλινδρόμηση για το διάστημα  $j$  ?
- ♦ Επέλεξε διάστημα με μεγαλύτερο συνολικό όφελος
  - Προσάρμοσε οφέλη άλλων διαστημάτων και επανέλαβε
- ♦ Προσέγγιση σημάτων
  - Αρχικά 1 διάστημα ανά σήμα
  - Συνεχώς σπάσε το διάστημα με το μεγαλύτερο λάθος
- ♦ Για διαφορετικά είδη σφαλμάτων, άλλαξε συνάρτηση παλινδρόμησης
- ♦ Αλγόριθμος δίνει περισσότερο χώρο σε “δύσκολα” σήματα
- ♦ Για μικρότερους κόμβους, οργάνωση σε ομάδες
  - Ηγέτης της ομάδας παρέχει λεξικό, κατανέμει χώρο (bandwidth)

26/08/2005

Αντόνιος Δεληγιαννάκης

12

## Αποτελέσματα – SSE & SSRE Σφάλματα

- ♦ Σύγκριση με Wavelets, DCT, Ιστογράμματα
- ♦ Κλάσμα λαθών δεύτερης καλύτερης προσέγγισης προς τον αλγόριθμό μας
- ♦ Σπάνιες ενημερώσεις λεξικού μετά από αρχικές μεταδόσεις

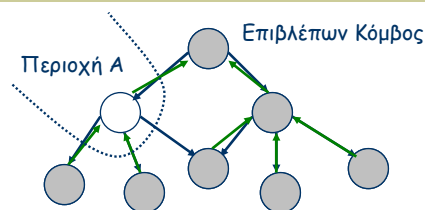
	SSE			SSRE
	Weather	Phone	Stock	Phone
5%	1.89	1.63	1.38	9.78
10%	2.04	2.01	1.70	5.97
15%	2.46	2.12	2.00	4.31
20%	3.02	2.61	2.44	5.37
25%	3.74	2.93	2.50	5.21
30%	4.44	3.24	3.67	6.22

26/08/2005

Αντόνιος Δεληγιαννάκης

13

## Συνάθροιση Δεδομένων



- ♦ Συνεχή συναθροιστική ερώτηση από επιβλέπων κόμβο
  - Πχ: “Ανέφερε μέση θερμοκρασία στην περιοχή B κάθε 1 sec”
- ♦ Διάδοση ερώτησης στο δίκτυο
  - Αναγνώριση σχετικών κόμβων
- ♦ Σχηματίζεται αντίστροφο “Δέντρο Συνάθροισης”
- ♦ Αποφασίζονται χρόνοι μετάδοσης/λήψης [TAG]
- ♦ Σημαντικά οφέλη, αλλά μετάδοση από κάθε κόμβο

26/08/2005

Αντόνιος Δεληγιαννάκης

14

## Προσεγγιστικές Συναθροιστικές Ερωτήσεις

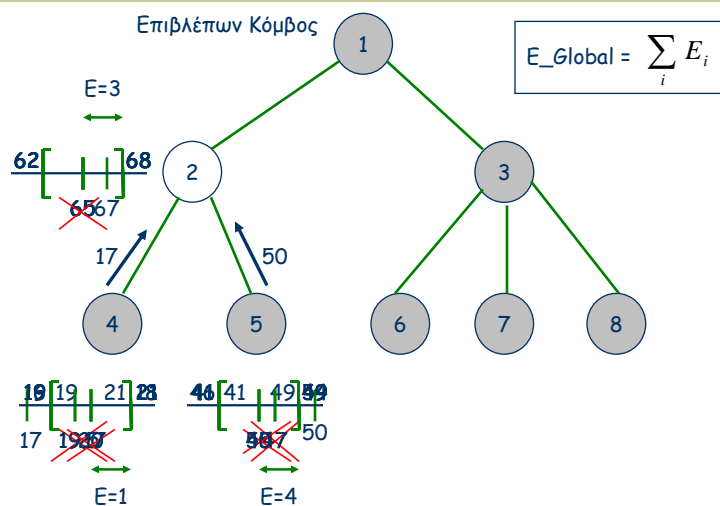
- ♦ **Ιδέα: Μετρήσεις συχνά μεταβάλλονται λίγο**
  - Μη μεταδίδεις μικρές μεταβολές σε συναθροιστικές τιμές και μετρήσεις
- ♦ **Αισθητή μείωση μηνυμάτων με μικρά λάθη**
- ♦ **Εφαρμογή καθορίζει μέγιστο σφάλμα  $E\_Global = |V - \hat{V}|$**
- ♦ **Αλγόριθμος εφαρμόζει φίλτρα σφάλματος στους κόμβους**
  - Στόχος: Ελαχιστοποίηση μεταδιδόμενων μηνυμάτων
- ♦ **Κάθε κόμβος υπολογίζει συναθροιστική τιμή υποδέντρου**
  - Μετάδοση μόνο αν νέα τιμή είναι εκτός του φίλτρου σφάλματος
  - Σε κάθε μετάδοση, νέος ορισμός κέντρου του φίλτρου
- ♦ **Φίλτρα περιοδικά μεταβάλλονται**
  - Προσαρμογή σε αλλαγές στα χαρακτηριστικά των δεδομένων

26/08/2005

Αντόνιος Δεληγιαννάκης

15

## Παράδειγμα Φίλτρων Σφάλματος



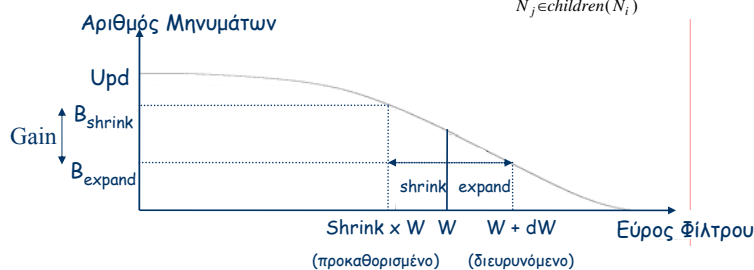
26/08/2005

Αντόνιος Δεληγιαννάκης

16

## Διαδικασία Μεταβολής Φίλτρων

- ♦ Φίλτρα μικραίνουν περιοδικά σε  $W \times \text{Shrink}$ 
  - Δημιουργία πλεονάσματος λάθους  $(1-\text{Shrink}) \times E\_Global$
- ♦ Πλεόνασμα λάθους μοιράζεται ξανά
  - Χρειάζεσαι εκτίμηση οφέλους όταν αυξάνεις φίλτρα
    - Ανά κόμβο και σε ολόκληρα υποδέντρα
  - Πιθανό όφελος: Μηνύματα που περιμένουμε να γλιτώσουμε με μεγαλύτερο φίλτρο σε σύγκριση με το προκαθορισμένο φίλτρο την επόμενη περίοδο
  - Cumulative gain:  $CumGain_i = Gain_i + \sum_{N_j \in children(N_i)} CumGain_j$



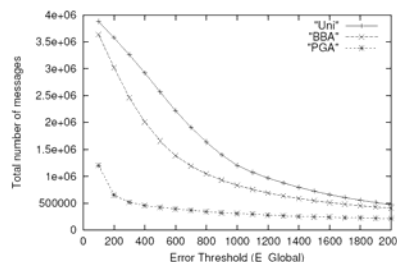
26/08/2005

Αντώνιος Δεληγιαννάκης

17

## Πειράματα – Απόδοση Αλγορίθμου

	$T_{leaves}$	$T_{all}$	$T_{random}$
PGA	423	479	677
BBA	1394	1426	2474
Uni	2568	2451	3906
TAG+ Caching	4176	4176	5142



- ♦ 3 τοπολογίες δέντρου (όνομα δείχνει θέση ενεργών κόμβων)
- ♦ 344-500 αισθητήρες
- ♦  $E\_Global = 500$   
( $\approx 2\%$  σχετικό σφάλμα)
- ♦ Μηνύματα μεταβάλλοντας  $E\_Global$
- ♦ Μέχρι 5 φορές βελτίωση

26/08/2005

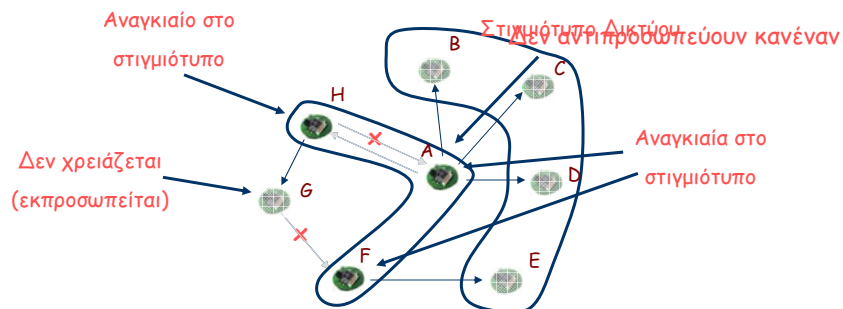
Αντώνιος Δεληγιαννάκης

18



## Τελικό Στάδιο

- ◆ Σπάσε ισοπαλίες, κύκλους
- ✓ Συνολική διαδικασία: 3-5 μηνύματα



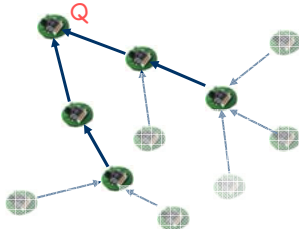
26/08/2005

Αντόνιος Δεληγιαννάκης

21

## Οφέλη σε Συναθροιστικές Ερωτήσεις

```
select AVG(temperature)
from Sensors
where loc in [ $x_{min}, x_{max}$ ] × [ $y_{min}, y_{max}$ ]
use snapshot
```



Query Area	Transmission Range	
	0.2	0.7
1%	11%	29%
10%	38%	77%
50%	52%	91%

26/08/2005

Αντόνιος Δεληγιαννάκης

22

## Συμπεράσματα

---

- ♦ Παρουσιάσαμε μεθόδους μείωσης δεδομένων σε δίκτυα αισθητήρων
  - Οφέλη σε bandwidth/ενέργεια
- ♦ Μετάδοση ιστορικών δεδομένων
  - Εκμετάλλευση συσχετίσεων, δημιουργία λεξικού
- ♦ Προσεγγιστικές συναθροιστικές ερωτήσεις
  - Ανάθεση/Μεταβολή φίλτρων σφαλμάτων σε κόμβους μέσω τοπικών στατιστικών
- ♦ Χρήση στιγμιότυπων δικτύου
  - Αντιπρόσωποι χρησιμοποιούνται σε απαντήσεις ερωτήσεων
  - Λιγότεροι ενεργοί κόμβοι, λιγότερα μηνύματα, μεγαλύτερη διάρκεια ζωής δικτύου

26/08/2005

Αντώνιος Δεληγιαννάκης

23

## Ευχαριστούμε - Ερωτήσεις?

---



26/08/2005

Αντώνιος Δεληγιαννάκης

24

## Localized Groups

- ◆ Sensors in a vicinity expected to obtain similar measurements
- ◆ Base signals may contain significant redundancy
- ◆ Not all nodes collect equally hard to approximate data
- ◆ Can localized groups improve quality of approximation?
  - By using at most as much bandwidth as non-localized

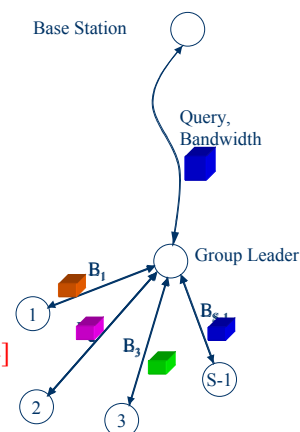
26/08/2005

Αντόνιος Δεληγιαννάκης

25

## Algorithm Overview

- ◆ Decide fraction of bandwidth for group transmission
- ◆ Allocate fraction of this bandwidth to nodes
- ◆ Nodes report error, NOT data
- ◆ More space to nodes with larger errors (handshake or 2-step)
- ◆ Compressed data to group leader
- ◆ Group leader recompresses
- ◆ Group leader election using HEED [Younis04]
- ◆ Optimization: Low-end sensors can use base signal of group leader
  - Requires 1 broadcast message



26/08/2005

Αντόνιος Δεληγιαννάκης

26

# Experiments

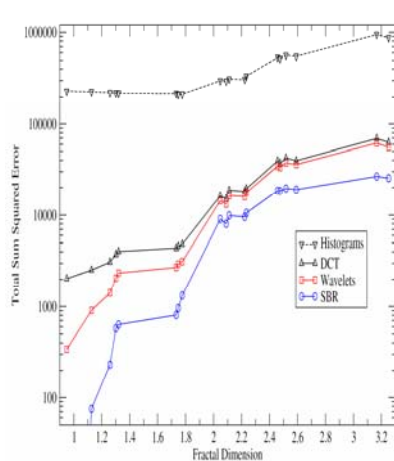
- ◆ Used real data sets
  - Stock data
  - Weather data
  - Phone data
- ◆ Comparison to Wavelets, DCT, Histograms
- ◆ Significantly smaller errors for SSE error
  - Even larger improvements in other error metrics
- ◆ Rare base signal updates after initial transmission

26/08/2005

Αντόνιος Δεληγιαννάκης

27

# Results – Synthetic & Localized



Synthetic data, N=4, M=2,048

k	Localized	Non-localized
5%	203,213	2,923,141
7.5%	99,342	812,026
10%	64,507	303,460
15%	35,784	84,907
20%	22,278	38,954
25%	14,876	23,706
30%	10,169	16,901

Total SSE Errors, H = 10

26/08/2005

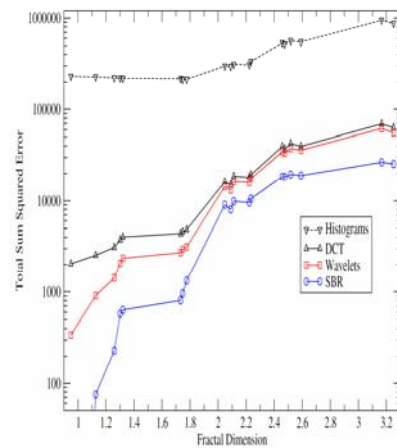
Αντόνιος Δεληγιαννάκης

28

## Results – SSRE Error & Synthetic

	Phone	Mixed
5%	9.78	186
10%	5.97	234
15%	4.31	284
20%	5.37	371
25%	5.21	1034
30%	6.22	526

Ratio of Errors of second best approximation to our Algorithm



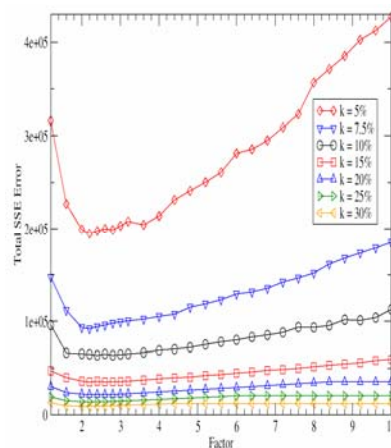
Synthetic data, N=4, M=2,048

26/08/2005

Αντόνιος Δεληγιαννάκης

29

## Localized Groups



k	Localized	Non-localized
5%	203,213	2,923,141
7.5%	99,342	812,026
10%	64,507	303,460
15%	35,784	84,907
20%	22,278	38,954
25%	14,876	23,706
30%	10,169	16,901

Total SSE Errors, H = 10

H = 10, modifying  $\text{avg}(k_1)/k_{g1}$

26/08/2005

Αντόνιος Δεληγιαννάκης

30