MaD-WiSe: Management of Data in Wireless Sensor Networks

ISTI-CNR
Giuseppe Amato  giuseppe.amato@isti.cnr.it
Stefano Chessa  stefano.chessa@isti.cnr.it
Antonio Caruso  antonio.caruso@isti.cnr.it
Paolo Baronti  paolo.baronti@isti.cnr.it
Several applications on WSN produce and process huge amount of data
- Data are continuously produced (data streams)
- Data produced by different sensors might need to be compared/matched
- Behaviour of sensors might need to be adjusted/refined over time
- Environmental situation can change so new strategies might need to be used
- Use of gathered data is not always known a priori
Declarative Queries

- Programming Apps is Hard
  - Limited power budget
  - Lossy, low bandwidth communication
  - Require long-lived, zero admin deployments
  - Distributed Algorithms
  - Limited tools, debugging interfaces
- Queries abstract away much of the complexity
  - Burden on the database developers
  - Users get:
    - Safe, optimizable programs
    - Freedom to think about apps instead of details
Data management: trivial approach

All sensed data go to a PC to be stored and processed

Wireless sensor network
Wireless Sensor Network as a distributed database
MaDWiSe approach

- Existing approaches do not distinguish among data acquisition, data transfer, data processing phases
- Our approach -> layered architecture:
  - Network layer
  - Stream system
  - Stream query processing
- All nodes ow the WSN have these layers
Network layer (1)

- Localization and Routing:
  - Virtual Coordinate assignment protocol (VCap):
    - Large sensor networks, not equipped with localization devices such as GPS
    - Distributed protocol which defines a coordinate system unrelated to the sensor location
    - VCap selects three anchors in the network boundary and assigns to each node a triplet of coordinates which represents the hop distance of the node from the three anchor nodes
    - The coordinate system can efficiently support greedy geographic routing
Network layer (2)

- Energy-efficient, application-driven communication:
  - Many applications use channels at fixed data rate
    - E.g.: directed-diffusion paradigm or data-base oriented applications
  - Connection-oriented communication protocol
  - Estimation of the next packet arrival time and turn on/off the radio accordingly
    - Minimization of the packet losses due to radio off
    - Minimization of energy consumption
  - Sensors not involved in the communication channels turn off the radio
Stream system (1)

- Wireless sensor network mainly produce and process streams of data
- Tree types of data sources
  - Transducers -> *Sensor streams*
  - Local applications -> *Local streams*
  - Network -> *Remote streams*

- Stream system: the equivalent of the “file system” for WSN applications
  - open, close, read, write like operations on various type of streams
  - Streams are n -&gt; 1 (n can write, 1 can read)
    - This limit is easily manageable
Stream system (2)

- Sensor Streams
  - A transducer writes, any local application can read
  - You can specify
    - the transducer
    - The acquisition rate
      - E.g. every x ms, or on demand
    - The lifetime
    - The buffer size
Stream system (3)

- Local streams
  - A local application writes
  - A local application read
  - You can specify
    - The data rate
    - The lifetime
    - The buffer size
Stream system (4)

- Remote streams
  - A possibly remote application writes
  - A possibly remote application reads
  - You can specify
    - A global name for the stream
    - The data rate (useful for energy efficient communication)
    - The lifetime
    - The buffer size
    - A reliability strategy
WSN as a database: Architecture

Diagram showing the architecture of a wireless sensor network (WSN) integrated as a database system. The process involves a query that goes through a parser and translator, generating a relational algebra expression. This expression is then optimized and translated into an execution plan. The execution plan is used to interact with the wireless sensor network (depicted as a graph of interconnected nodes) and provide statistics about the data. The output is then directed back to the query output.
WSN as a database:
Stream query processor

- In traditional databases queries access relations
- In our case queries access data streams
  - We use stream query processing techniques
WSN as a database: Simple query example

```
select t, l
from T, L
Where T.t_s=L.t_s
  t > 20 and
  l > 10
```
WSN as a database: Power aware query optimisation

- Cost of a query is estimated considering energy consumption
  - Processing data consumes energy
  - Sensing data consumes energy
    - Different sensors have different energy consumption
  - Transmitting data consumes energy
  - Receiving data consumes energy
WSN as a database: Execution on a single node of a WSN

Cost:
\[(C_l + C_{pr}) \times n + P(l > 10) \times (C_{pr} + C_t) \times n + P(l > 10) \times C_{pr} \times n\]
WSN as a database:
Execution on two nodes of a WSN

\[
\text{Cost:} (C_l + C_{pr}) \cdot n + (C_{tx} + C_t + P(l>10)(C_{pr} + C_{rx})) \cdot n + P(l>10) \cdot C_{pr} \cdot n
\]

Receive what needed
Transmit all
Expansive sensing
Low energy sensing

Data Processing
Data Processing
Data Processing
MaDWiSe: Current Status

- **Network layer**
  - Design ready
  - Simple prototype ready (no smart energy management, no multi-hop)
  - Full functional prototype in progress
- **Stream System**
  - Design ready
  - Advanced prototype ready
  - Full functional prototype in progress
- **Query execution**
  - Query processor ready and incrementally updated
  - Graphical interface for query execution plan definition ready
  - SQL query parser and optimizer in progress