Self-Organization in Autonomous Sensor/Actuator Networks

Dr.-Ing. Falko Dressler

Computer Networks and Communication Systems
Department of Computer Sciences
University of Erlangen-Nürnberg

http://www7.informatik.uni-erlangen.de/~dressler/
dressler@informatik.uni-erlangen.de
Overview

- **Self-Organization**
  Introduction; system management and control; principles and characteristics; natural self-organization; methods and techniques

- **Networking Aspects: Ad Hoc and Sensor Networks**
  Ad hoc and sensor networks; self-organization in sensor networks; evaluation criteria; medium access control; ad hoc routing; data-centric networking; clustering

- **Coordination and Control: Sensor and Actor Networks**
  Sensor and actor networks; coordination and synchronization; in-network operation and control; task and resource allocation

- **Bio-inspired Networking**
  Swarm intelligence; artificial immune system; cellular signaling pathways
Data-Centric Communication

- Flooding / Gossiping / WPDD
- Rumor routing
- Directed Diffusion
- Data aggregation and data fusion
Overview and classification

- Data dissemination – forwarding of data though the network
- Network-centric operation – data manipulation and control tasks
  - Network-centric pre-processing, e.g. data aggregation and fusion
  - In-network operation and control, e.g. rule-based approaches

Data-centric networking

- Data dissemination
  - Flooding
  - Gossiping
  - Agent-based approaches
  - Reverse path techniques

- Network-centric operation
  - Network-centric pre-processing
    - Aggregation
  - Data fusion
  - Rule-based data processing
  - GRID approaches

[SelfOrg]
Flooding

- **Basic mechanism:**
  - Each node that receives a packet re-broadcasts it to all neighbors
  - The data packet is discarded when the maximum hop count is reached

---

Step 1  
Step 2  
Step 3
Flooding

- **Advantages**
  - No route discovery mechanisms are required
  - No topology maintenance is required

- **Disadvantages**
  - **Implosion**: duplicate messages are sent to the same node
  - **Overlap**: same events may be sensed by more than one node due to overlapping regions of coverage \(\rightarrow\) duplicate report of the same event
  - **Resource blindness**: available energy is not considered and redundant transmissions may occur \(\rightarrow\) limited network lifetime
Topology assisted flooding

- Exploiting overhearing in wireless networks

while Receive a new flooding packet P do
  Start a process on packet P
  Wait for $T$ time units – overhearing period
  if Each one-hop neighbor is already covered by at least one broadcast of P then
    terminate process on packet P
  else
    Re-broadcast packet P
  end if
end while
Simple gossiping

- GOSSIP($p$) – Probabilistic version of flooding
- Packets are re-broadcasted with a gossiping probability $p$

**for each** message $m$

\[
\text{if random}(0,1) < p \text{ then message } m
\]
Simple gossiping

- **Advantages**
  - Avoids packet implosion
  - Lower network overhead compared to flooding

- **Disadvantages**
  - Long propagation time throughout the network
  - Does not guarantee that all nodes of the network will receive the message (similarly do other protocols but for gossiping this is an inherent “feature”)
Optimized gossiping

- Two-threshold scheme
  - GOSSIP($p$, $k$) – Flooding for the first $k$ hops, then gossiping with probability $p$
    - GOSSIP(1, $k$) $\rightarrow$ flooding
    - GOSSIP($p$, 0) $\rightarrow$ simple gossiping

- Destination attractors
  - Weighted gossiping probability according to the distance of the current node to the final destination
    
    \[
    P_{R_i} = \begin{cases} 
    (1 + k)P_{R_{i-1}} & \text{closer to destination} \\
    (1 - k)P_{R_{i-1}} & \text{further to destination} \\
    P_{R_{i-1}} & \text{same or indeterminate} 
    \end{cases}
    \]

  $P_{R_i}$ is the gossiping probability for a packet at the $i^{th}$ node $R_i$ in its path to the destination, $k$ can be used to scale the probability.
Weighted Probabilistic Data Dissemination (WPDD)

- Optimized gossiping
  - Each message (data value) to be sent is given a priority \( I(msg) \)
  - The message is processed according to the message-specific gossiping probability \( G(I(msg)) \) and a node-specific weighting \( W(S_i) \) for each node \( S_i \)

- Message forwarding condition: \( G(I(msg)) > W(S_j) \)
Rumor Routing

- Agent-based path creation algorithm
  - Agents, or “ants” are long-lived entities created at random by nodes
  - These are basically packets which are circulated in the network to establish shortest paths to events that they encounter
Rumor Routing

- Agent-based path creation algorithm
  - Can also perform path optimization at nodes that they visit
  - When an agent finds a node whose path to an event is longer than its own, it updates the node's routing table
Directed Diffusion

- Diffusion routing protocol
- Improves on data diffusion using interest gradients

- Basic behavior
  - Each sensor node *names its data* with one or more attributes
  - Other nodes *express their interest* depending on these attributes
  - The sink node has to periodically refresh its interest if it still requires data to be reported to it
  - Data is propagated along the reverse *path of the interest propagation*

- Optimizations
  - Nodes are allowed to cache or locally transform (aggregate) data
    - increases the scalability of communication and reduces the number of required transmissions
Directed Diffusion

- **Interest propagation**
  
  ```
type = four-legged animal
interval = 1s
rect = [-100, 200, 200, 400]
timestamp = 01:20:40
expiresAt = 01:30:40
  ```

- **Data transmission**
  
  ```
type = four-legged animal // type of animal seen
instance = elephant // instance of this type
location = [125, 220] // node location
intensity = 0.6 // signal amplitude measure
confidence = 0.85 // confidence in the match
timestamp = 01:20:40 // event generation time
  ```
Directed Diffusion

(a) Interest propagation

(b) Gradient setup

(c) Data delivery
Directed Diffusion – Performance Aspects

Average Dissipated Energy

Node Failures – Event Delivery Ratio

[Graphs showing average dissipated energy and event delivery ratio with network size as the x-axis and energy/event delivery ratio as the y-axis, with different node failure scenarios represented by different markers and line styles.]
Improving directed diffusion

- **Node mobility**
  - Aggressive diffusion – improved *timeout* handling
  - Handoff and proxies – similar to handoff in mobile communication
  - Anticipatory diffusion – setting up paths **before** node movements

- **Energy efficiency**
  - Based on passive clustering techniques

Diagram:

- Gradient setup w/o clustering
- Gradient setup w/ clustering
Data aggregation – Motivation

- **Energy constraints and network congestion**
  - Data transmission in sensor networks is much more energy expensive compared to local computation efforts
  - The reduced number of transmitted messages towards the base station helps reducing network congestion (especially near the base station)

- **Redundancy and correlation**
  - A certain degree of overlap and redundancy is created as measured sensor data is often generated by nearby nodes
  - Measured data can be expected to be highly correlated allowing further improvements of the information quality by using data fusion approaches (possibly exploiting further available meta information)
Data aggregation – Terminology

- **Data aggregation** – Data aggregation is the process of combining multiple information particles (in our scenario, multiple sensor messages) into a single information that is representing all the original messages. Examples of aggregation methods are statistical operations like the mean or the median.

- **Data fusion** – Data fusion is the process of annotating received information particles with meta information. Thus, data from different is combined to produce higher quality information, e.g. by adding a timestamp or location information to received sensor readings.
Aggregation techniques

Chain-based aggregation

Tree-based aggregation

Grid-based aggregation

[SelfOrg]
Limitations

- Optimization latency vs. efficiency
  - High aggregation ratios require long aggregation delays $\Delta t$
  - Large $\Delta t$ will obviously lead to increased message transmission delays
Summary (what do I need to know)

- **Data-centric communication**
  - Main ideas and principles

- **Data dissemination techniques**
  - Principles and limitations of
    - Flooding / Gossiping / WPDD
    - Rumor routing
    - Directed Diffusion

- **Data aggregation and data fusion**
  - Differentiation aggregation vs. fusion
  - Advantages and limitations
References