Outline

- Types of data collection in WSNs
- Directed Diffusion model
- Evaluation results
- Diffusion variants
Types of Data Collection

- Streaming
- In-Network Processing
- Triggering
- Tasking

Streaming

- Most typical method in current deployments
- Great Duck Island deployment
- Correlation, processing, and logic is largely done outside of the network
In-Network Processing

- Routing
  - Diffusion Matching
- Storage
  - Data Distribution
  - Lookup
- Aggregation
  - TinyDB
- Compression
- ESS
- Loose Model

Triggering

- Event Detection
  - Data may be routed for different reasons
  - Relevance of data is defined by distributed model
Tasking

- Network control and configuration
- Query languages
- Programming “scripts”
- Binary-reprogramming

```
if (0 == iTemp)
{
    Sense(&var1);
}
```

Directed Diffusion

- New routing model
  - Named data
  - Publish and subscribe
  - Diffusion protocols
Publish and Subscribe

- Data sources need not know anything about data sinks
- Network can scale independent of data consumption
- Data can be routed dynamically, according to its value
  - Matching

Named Data

- Data names are attributes
  - Key
  - Operation
  - Value
- Routes are known as Gradients
- Attributes (data) is matched along gradients as it flows through the network
Attributes Implementation (1)

- Each attribute implemented as a key-type-value-operator tuple
- Key
  - Indicates the semantics of the attribute
  - Latitude, frequency, etc
- Type
  - Indicates the primitive type that the key will be
  - Available types
    - INT32_TYPE // 32-bit signed integer
    - FLOAT32_TYPE // 32-bit
    - FLOAT64_TYPE // 64-bit
    - STRING_TYPE // UTF-8 format
    - BLOB_TYPE // un-interpreted binary data

Attributes Implementation (2)

- Operator
  - Describes how the attribute will match when two attributes with the same key and type are compared
  - Available operators are: IS, EQ, NE, GE, GT, LE, LT, EQ_ANY
  - IS operator specifies a known, actual value
  - Other operators specify a condition that must be satisfied

- Examples
  - LATITUDE IS 12.3
  - LONGITUDE LE 45.2
  - TEMPERATURE IS 72.7
  - CONFIDENCE GT 0.80
Matching Example (1)

Each formal must match some actual:

- type EQ detection
- $x \geq -100$
- $x \leq 200$
- $y \geq 100$
- $y \leq 400$
- confidence $> 0.5$

- type IS detection
- $x \equiv 10$
- $y \equiv 150$
- confidence IS $0.7$

**valid match**

Matching Example (2)

Each formal must match some actual:

- type EQ detection
- $x \geq -100$
- $x \leq 200$
- $y \geq 100$
- $y \leq 400$
- confidence $> 0.5$

- type IS detection
- $x \equiv 10$
- $y \equiv 10$
- confidence IS $0.7$

**no match**
Basic Directed Diffusion

Setting up gradients

Source

Sink

Interest = Interrogation
Gradient = Who is interested

Basic Directed Diffusion

Sending data and Reinforcing the best path

Source

Sink

→ Low rate event

⇒ Reinforcement = Increased interest
Directed Diffusion and Dynamics

Recovering from node failure

Source

Sink

→ Low rate event

→ High rate event

→ Reinforcement

Stable path

→ Low rate event

→ High rate event
**Initial simulation study of diffusion**

- Key metric
  - Average Dissipated Energy per event delivered
    - indicates energy efficiency and network lifetime

- Compare diffusion to
  - flooding
  - centrally computed tree (omniscient multicast)

**Diffusion Simulation Details**

- Simulator: ns-2
- Network Size: 50-250 Nodes
- Transmission Range: 40m
- Constant Density: 1.95x10^-3 nodes/m^2 (9.8 nodes in radius)
- MAC: Modified Contention-based MAC
- Energy Model: Mimic a realistic sensor radio [Pottie 2000]
  - 660 mW in transmission, 395 mW in reception, and 35 mw in idle
**Diffusion Simulation**

- Surveillance application
  - 5 sources are randomly selected within a 70m x 70m corner in the field
  - 5 sinks are randomly selected across the field
- High data rate is 2 events/sec
- Low data rate is 0.02 events/sec
- Event size: 64 bytes
- Interest size: 36 bytes
- All sources send the same location estimate for base experiments

**Average Dissipated Energy (Standard 802.11 energy model)**

Standard 802.11 is dominated by idle energy
**Average Dissipated Energy (Sensor radio energy model)**

Diffusion can outperform flooding and even omniscient multicast. **WHY?**

**Impact of In-network Processing**

Application-level suppression allows diffusion to reduce traffic and to surpass omniscient multicast.
Impact of Negative Reinforcement

Reducing high-rate paths in steady state is critical

Diffusion Variants

- Diffusion family of protocols includes variants optimized for specific scenarios
  - No one single protocol is optimal for all cases
- Two-Phase Pull Diffusion
  - Interests have global scope (are sent to the network)
  - Data has local scope (does not leave a node if there is no matching interest)
- PUSH Diffusion
  - Interests have local scope (do not leave the sink node)
  - Data has global scope (is flooded to the network)
  - After rendezvous at a sink node, a positive reinforcement is sent on the reverse path, establishing a reinforced path
  - Local subscriptions not possible (as interest messages do not leave sinks)
One Phase Pull

- Subscriber based system that avoids one of the two phases of flooding in two-phase pull
- Only floods interests
- No exploratory messages

Tiny Diffusion

- Directed Diffusion’s younger brother
- Implemented on Mica motes
- Special component to filter out asymmetric links