Exploiting Mobility for Energy Efficient Data Collection in Wireless Sensor Networks (Data Mule)

Mobile Networks and Applications 2006

Presented by: Matthew Tan Creti and Madalina Vintila

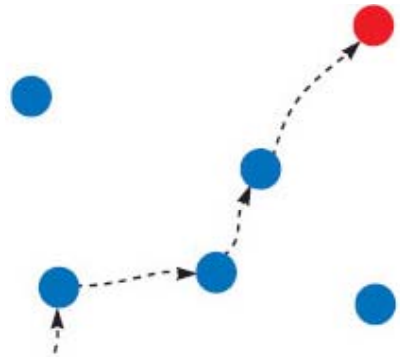
10/28/09

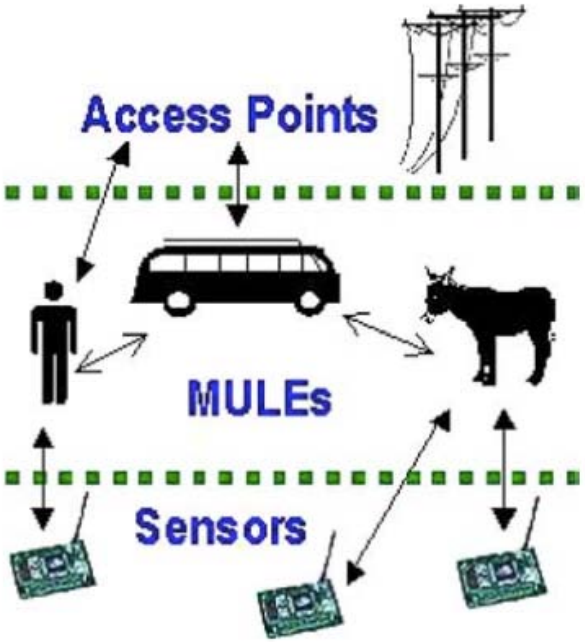
Data Collection in Sparse Sensor Networks

- Zebranet
  - Track zebra movement in remote locations
  - Used flooding to get data back to base station
- Other examples
  - Remote weather conditions
  - Habitat monitoring
  - City traffic
- Sensors should last a long time (e.g. months or years) unattended so the principle constraint is energy

Using Data MULEs

- One approach to data collection is to build ad-hoc networks
  - Increased radio power to reach neighboring nodes in sparse network
  - Nodes have to spend energy forwarding other nodes data
  - Network has hop spots near access points





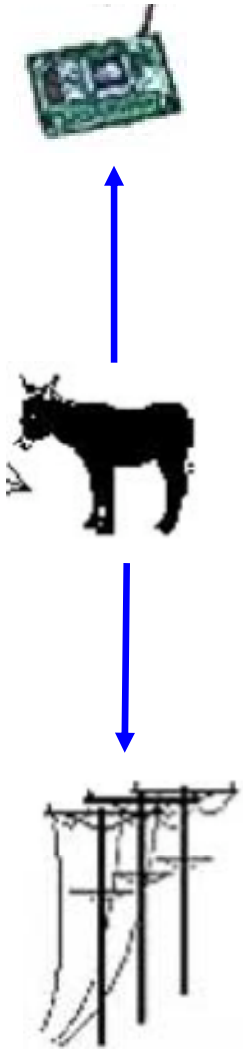
- The data MULE approach exploits mobile actors to carry data to access points
  - Examples of possible MULEs are animals, people, vehicles
  - Low power radios and no routing overhead
  - Tradeoff is high latency (usually acceptable in data collection applications)

Paper Outline

- Contributions
- MULE architecture
- Analytical Model
  - Performance measured in data transfer rate, latency, energy requirements
- Results
  - Order of magnitude energy savings and longevity of sensor nodes over ad-hoc networks
- Conclusion

MULE Architecture

- Lower tier - *sensors*: resource and energy constrained devices
- Middle tier - *MULEs*:
  - Large storage capacity and renewable energy
  - Responsible for discovering sensor nodes and access-points (why?)
  - MULE-to-MULE communication possible enhancement to basic architecture
- Upper tier - *access-points*:
  - Internet connected
  - Enhanced power
- Authors claim tiers can be collapsed into single device (e.g. Zebranet) (is this true?)



Advantages/Disadvantages of MULE Architecture

- Benefits
  - Energy efficient: sensors communicate over short range, no hot spots in network (are you sure about this?)
  - Spatial reuse: short range of communication allows spatial reuse of bandwidth, avoids radio complexities such as collisions (for sensors at least!)
  - No routing overhead
  - Robustness: Performance degrades gracefully as MULEs fail, any single MULE failure does not lead to a disconnected network (makes assumption about mobility!)
  - Scalable: adding new sensors and MULEs requires no network reconfiguration
  - Simplicity: No routing protocol to worry about, sensor's radio stack can be very lightweight, no need for synchronization or localization
  - Simple security: MULEs authenticate themselves, sensors encrypt data (can be done it advance), and access points decrypt; do not need to worry about buffer overflow?
- Limitations
  - Latency: limits realtime applications
  - Best-effort delivery: what happens when MULE fails, radio communication errors

Performance Metrics

- *Data success ratio* (DSR): ratio of total amount of data transferred to the access-points to the total amount generated, ideal =1, <1 caused by MULE failure, communication errors, or buffer overflow
- *Latency*: average time for data to reach access-point from time of generation
- *Communication energy*: both average energy per sensor and worst case (dictates network lifetime)

Parameter Space

- Sensor related:
  - λ := average amount of generated be a sensor
  - SB := sensor buffer size
- MULE related:
  - Model assumes knowledge of inter-arrival distribution (MULE arrival is a discrete event)
  - Assume MULE buffers are sufficiently large
- Access point related:
  - Modeled by parameter characterizing distribution of time interval between visits to access-point by a MULE
- Radio related:
  - r := MULEs can communicate within this distance
  - B := rate of data transfer from sensor to MULE
  - K := amount of data transferred between a MULE and a sensor

Model

- Queue of generated (undelivered) data at each sensor
- Queue is served whenever a MULE is in a sensor's rage (modeled as discrete event)
- Amount of data transferred on a MULE arrival is a random variable that depends on factors such as the time the MULE is in communication rate (for analytical tractability, this is taken as a fixed quantity)

Assumptions

- The MULEs arrival process at a sensor is a renewal process $\{S(t), t \geq 0\}$, where $S(t)$ is the total number of MULEs that have visited the sensor up until time t. This makes assumption that inter-arrival times of MULEs are independent and identically distributed denoted by random variable X^S. With average μ and variance σ_{ms}
- At a given time only one MULE interacts with a given sensor, and sensors are not close enough to each other to contend for service
- Sensors are identical and not mobile
- The data generation process at a sensor is a renewal process $\{U(t), t \geq 0\}$, where $U(t)$ is the total amount of data generated till time t. Average data generation rate is λ
- The queuing discipline is FCFS
- Without loss of generality, $SB \geq K$. If $SB < K$ then the maximum amount of data that is available at a sensor buffer to transfer is K
- Data transmission does not incur any loss, the only loss is due to sensor buffer overflow
- The queuing system is stable and only the stationary (time independent) probabilities are considered

Results

- **Result 1. The system is stable (the queue reaches a unique stationary regime) iff**

$$\frac{\lambda}{K\mu} \leq 1$$

- **Result 2. Data Success Ratio (DSR) is given by:**

$$DSR = \frac{\mu E[\min(K, Q)]}{\lambda} = \frac{\mu\left(\sum_{j=0}^{K} j P_j + \sum_{j=K+1}^{SB} K P_j\right)}{\lambda}$$

- **Result 3. Average queuing delay (Wb) is given by:**

$$W^q = \frac{\mu^2 \sigma_{ms} + 1}{2\mu} + \frac{E[B^{no}]}{\mu}$$

Scenarios for distribution analysis

1. *The MULE arrival distribution and the data generation process is Poisson*

- *In this case, Pj depends only on the ratio of miu and lambda… thus the specific values are not important*

2. K is large (K >= SB)

$$E[Q] = E[U[E[X^s]] = \frac{\lambda}{\mu}$$

Determining K

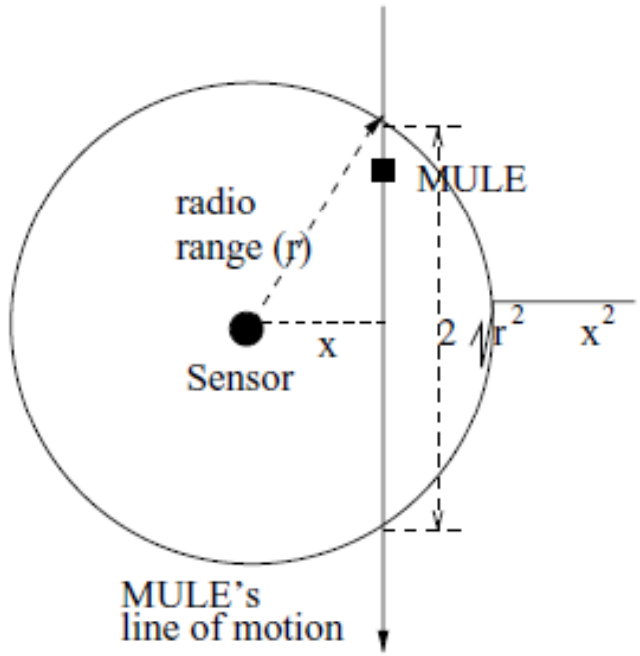


Figure 3. Amount of time a sensor is in contact with a MULE.

$$K = \left(\frac{\pi}{2}\frac{r}{v}\right) B$$

Impact of sensor duty cycle

- Duty cycle = discovery time / beacon time
- When duty cycle < 100%, performance suffers because:
  - A MULE may not be discovered at all because the sensor was asleep during the time the MULE was in communication range of sensor

$$\mu^* \text{ is } \mu(CT - DT)/BT.$$

$$\text{if } CT - DT \geq BT \; \mu^* = \mu$$

  - The amount of data that can be transferred (*K) in one* contact may decrease if the MULE is discovered in the middle of the duration it is in the communication range of the sensor

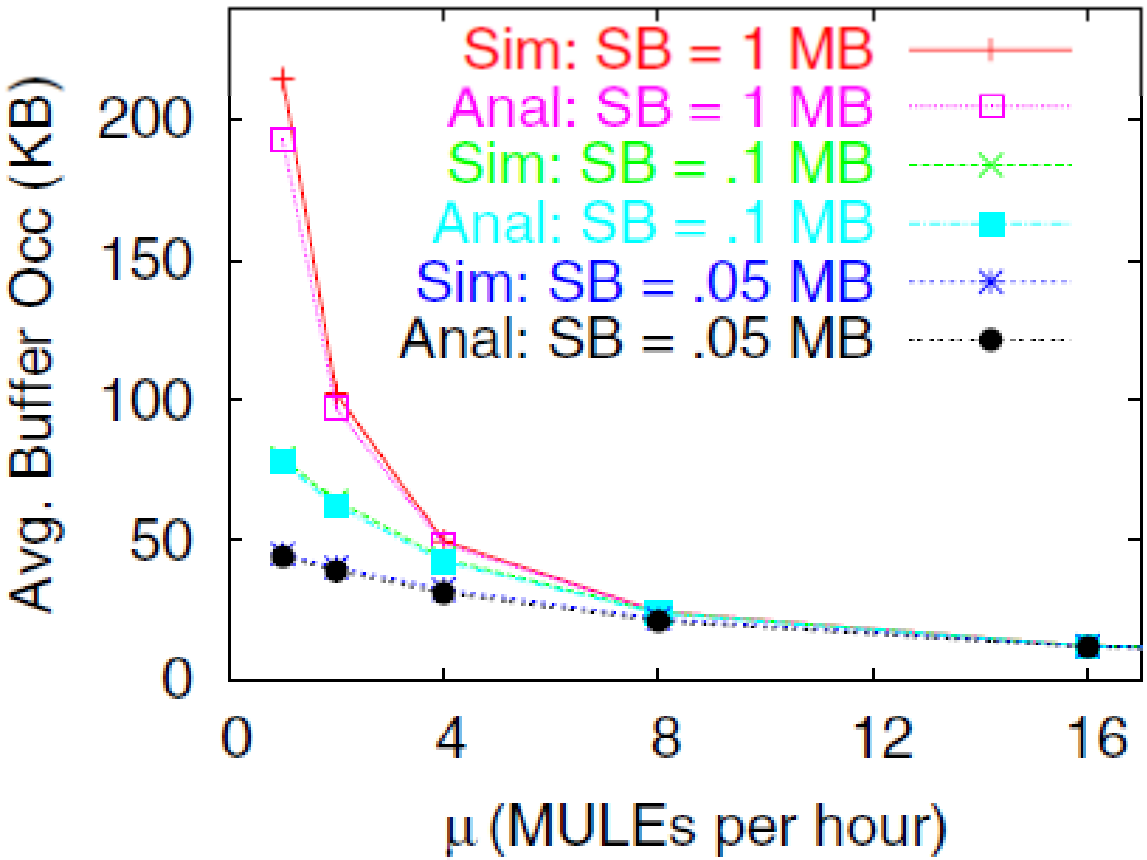
$$K^* \text{ is } \frac{K}{2}(1 + DT/CT)$$

Simulation setup

- Two-dimensional grid, sensors and access points randomly placed in grid
- MULEs: described by initial position and mobility pattern (Random waypoint, random walk, deterministic arrivals (fixed route and velocity), and Poisson arrivals)
- Data generation either uniform or Poisson distribution
- Parameters:
  - K is varied by varying radio bandwidth B
  - μ is varied by adding more MULEs
  - 2 km x 2 km grid, 100 sensors and 1 access point at corner
  - MULE buffer was 10 MB, velocity 10 m/s, radio range 25 m
  - Λ fixed at 90 KB / hour, sensor duty cycle 1/200

Performance Metrics

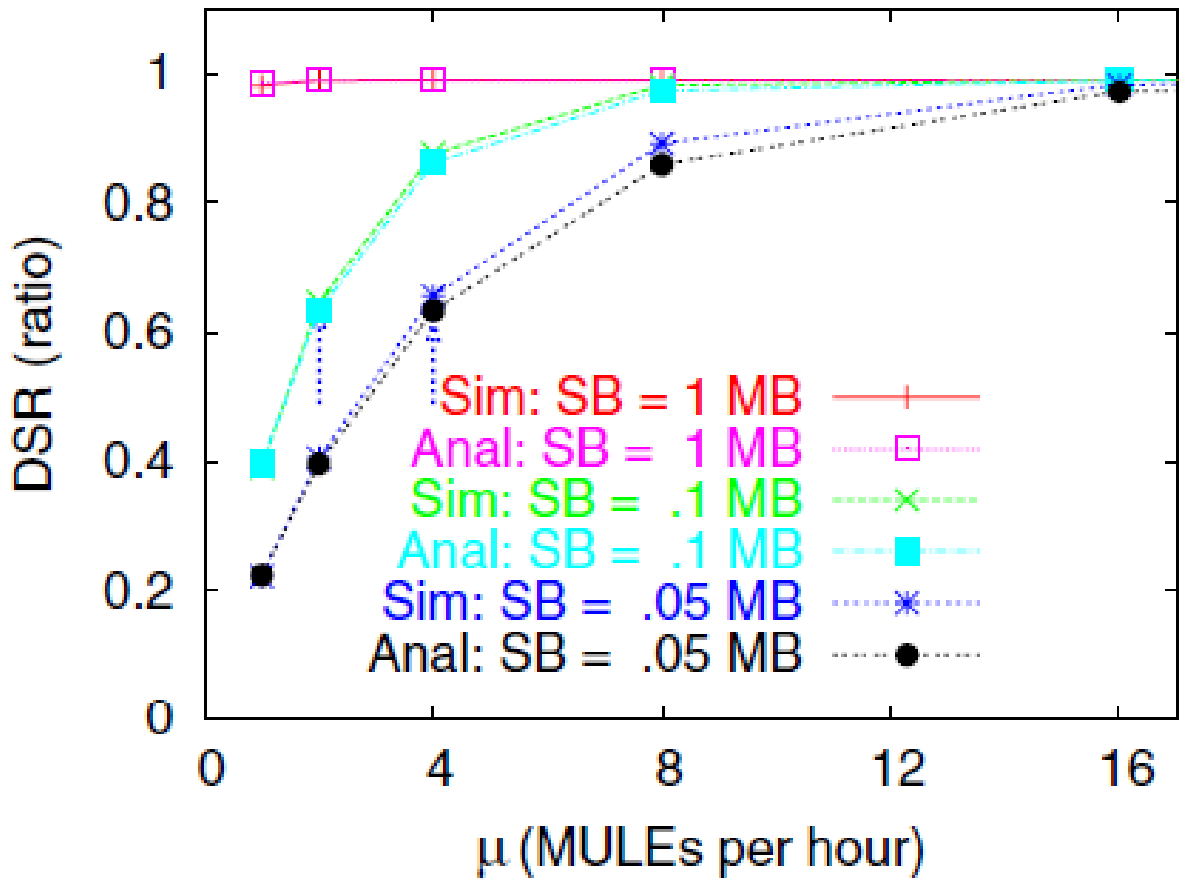
- Scaling μ and SB



(a) Average Buffer Occupancy

Performance Metrics

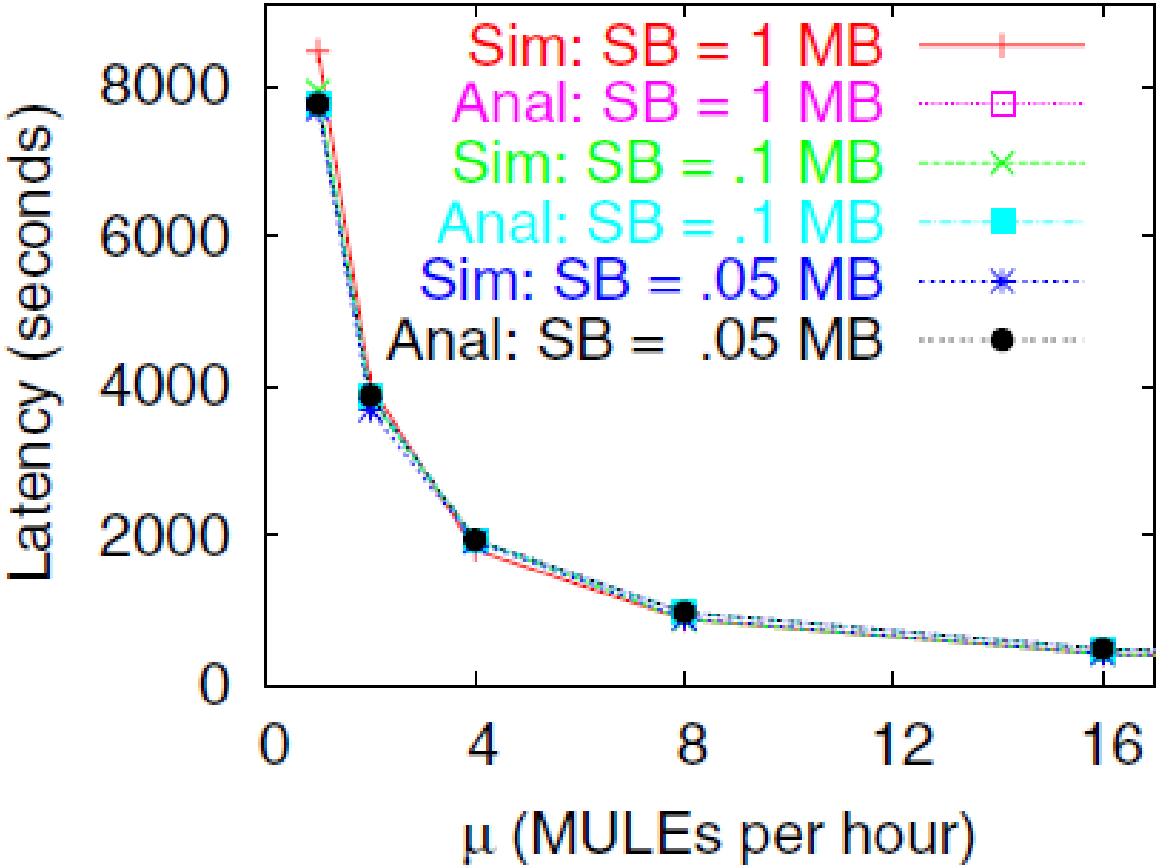
- Scaling μ and SB



(b) DSR

Performance Metrics

- Scaling μ and SB



(c) Latency

Performance Metrics

- Scaling K

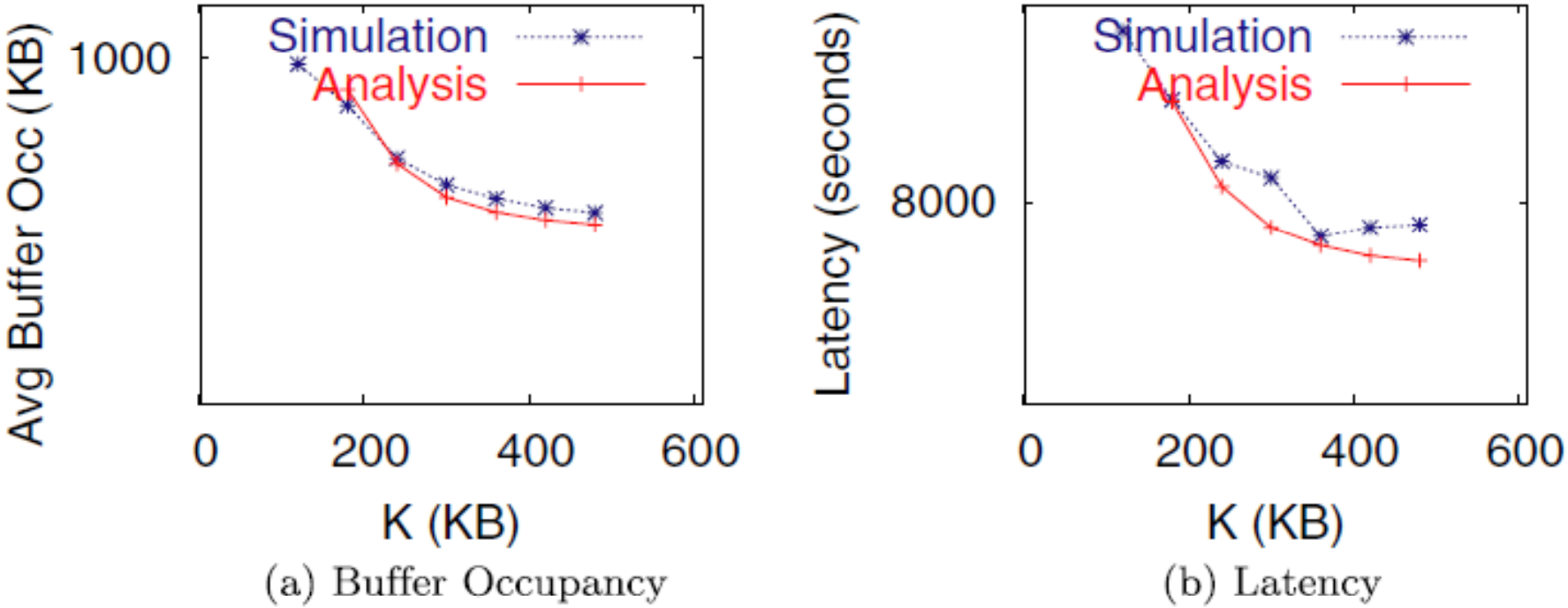


Figure 5. Effect of scaling K. (a) buffer occupancy (b) latency.

Performance Metrics

Table 1
Coefficient of variation (CVR) for MULE inter-arrival distribution for different mobility models.

Mobility model	CVR
Poisson	1.0
Deterministic	0.0
Waypoint	0.75
Manhattan	2.1

Table 2
Effect of parameters on performance.

	Performance metrics		
Parameters	Buffer Occ	DSR	Latency
$\mu\uparrow$	$\downarrow$	$\uparrow$	$\downarrow$
$SB\uparrow$	–	$\uparrow$	–
$K\uparrow$	$\downarrow$	$\uparrow$	$\downarrow$
$\lambda\uparrow$	$\uparrow$	$\downarrow$	$\uparrow$

Note. $\uparrow$ indicates an increase in the quantity. $\downarrow$ indicates a decrease and — indicates no effect.

Performance Metrics

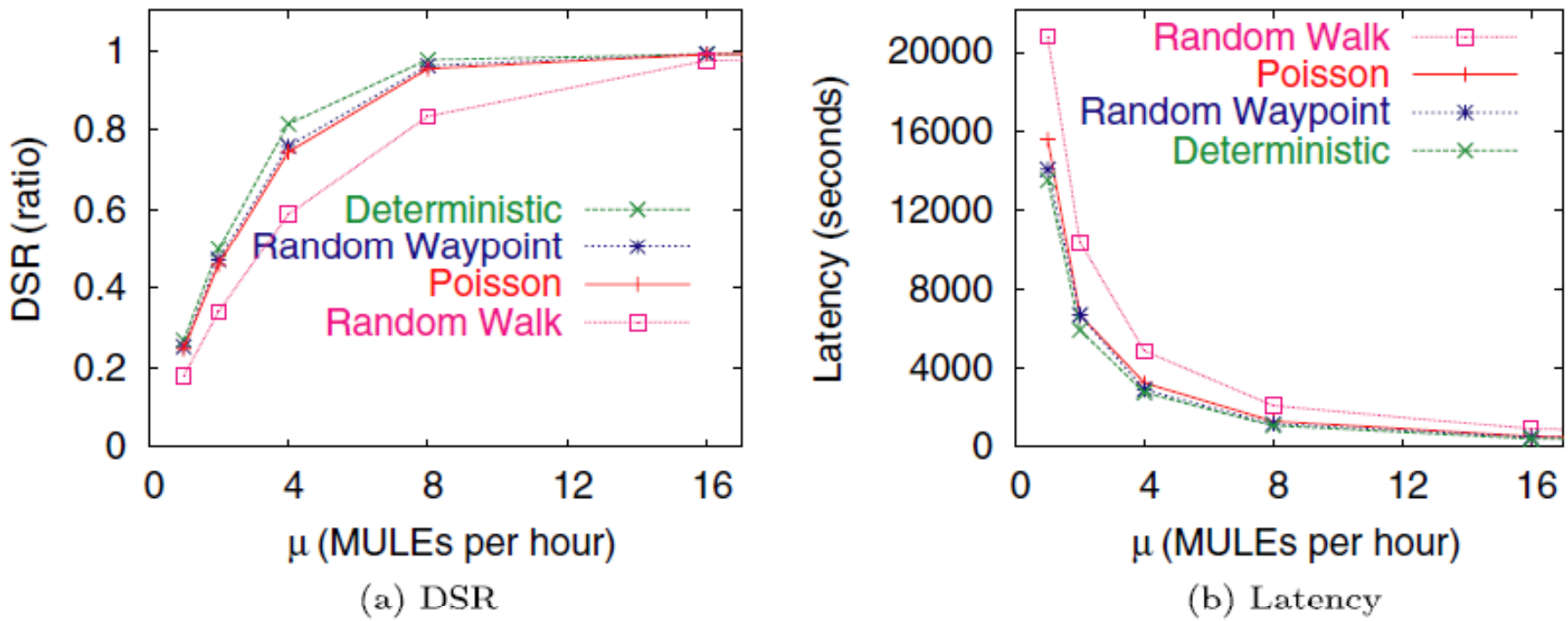
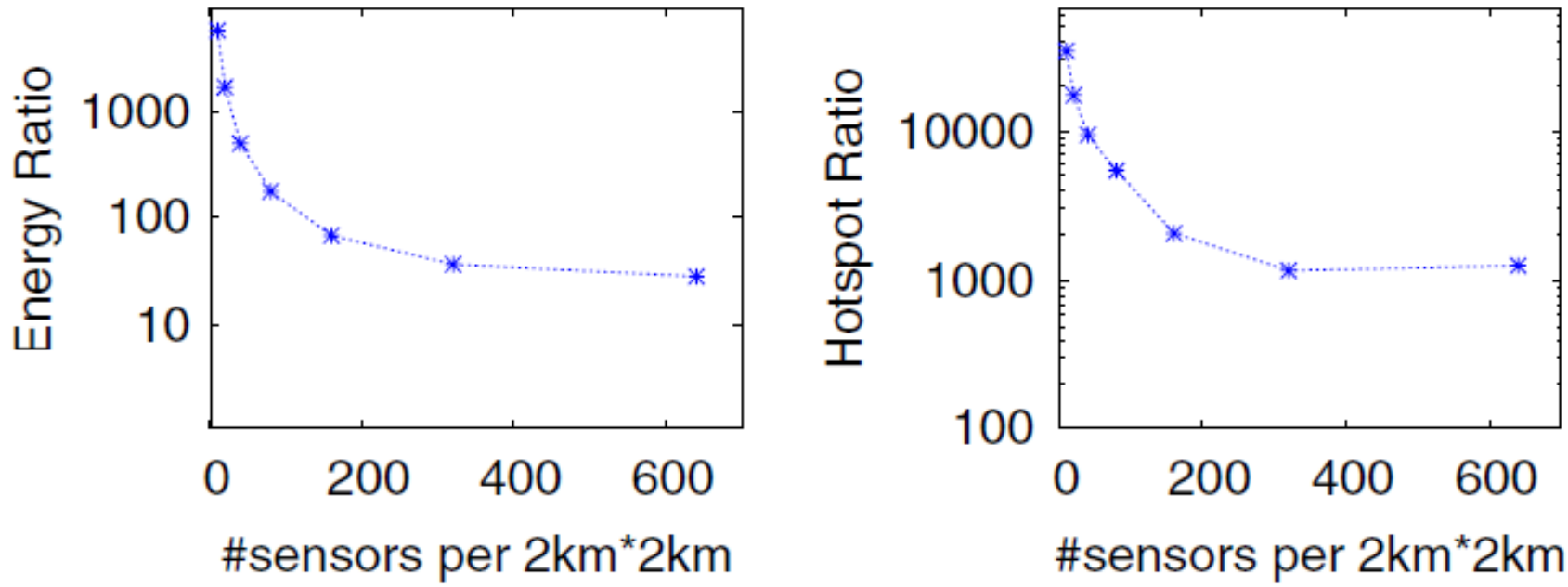


Figure 6. Effect of different mobility models.

Performance Metrics



(a) Avg. Energy Ratio (b) Hotspot Ratio

Figure 7. Energy comparison of MULE vs Ad-hoc network approach as a function of sensor density.

Conclusions

- Exploiting mobility for energy-efficient data collection in sparse sensor networks vs. forming ad hoc network
- Limited to non real-time communication and only when there are mobile nodes present in the network
- MULE to MULE and sensor to sensor communication could enhance data success ratio