A Communication Non-intrusive Middleware for Resource Management in Sparse MANETs

Ovidiu Valentin, DRUGAN
Department of Informatics, University of Oslo, Norway

Overview

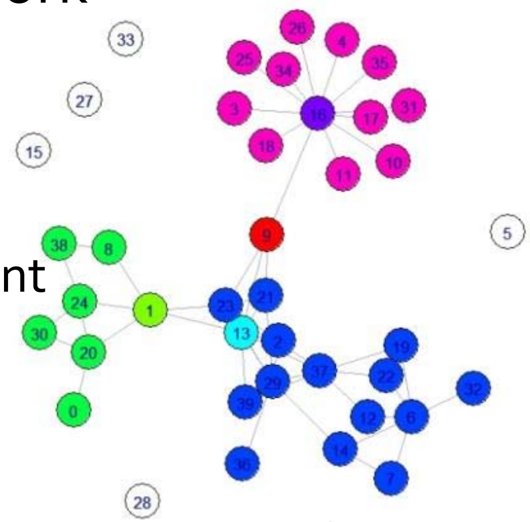
- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Rescue Operations and Emergency Intervention

- Red Cross, 2005: " … the flow of information throughout the disaster cycle is crucial for effective humanitarian operations …"
- Resource Operation Specific:
  - Diverse scenes: landscape, area size, number of people, status of infrastructure, and time span
  - Heterogeneity: organizations, devices, and configurations
  - Cooperative personnel hierarchically organized
- Middleware for Rescue Operations and Emergency Intervention → Ad Hoc InfoWare middleware

Ad Hoc InfoWare Middleware for Rescue Operations

- **Goal**: Improve the information sharing in rescue operations and emergency interventions
  - Increase the efficiency of collaborative work
- **Solution**: use data networks
  - MANETs + DTNs ≈ Sparse MANETs
  - Advantages:
    - Infrastructure independent and fast deployment
    - Mobile and heterogeneous devices
  - Disadvantages
    - Mobile and heterogeneous devices
    - Unpredictable scene layout, environment, and movements of users → unpredictable connectivity

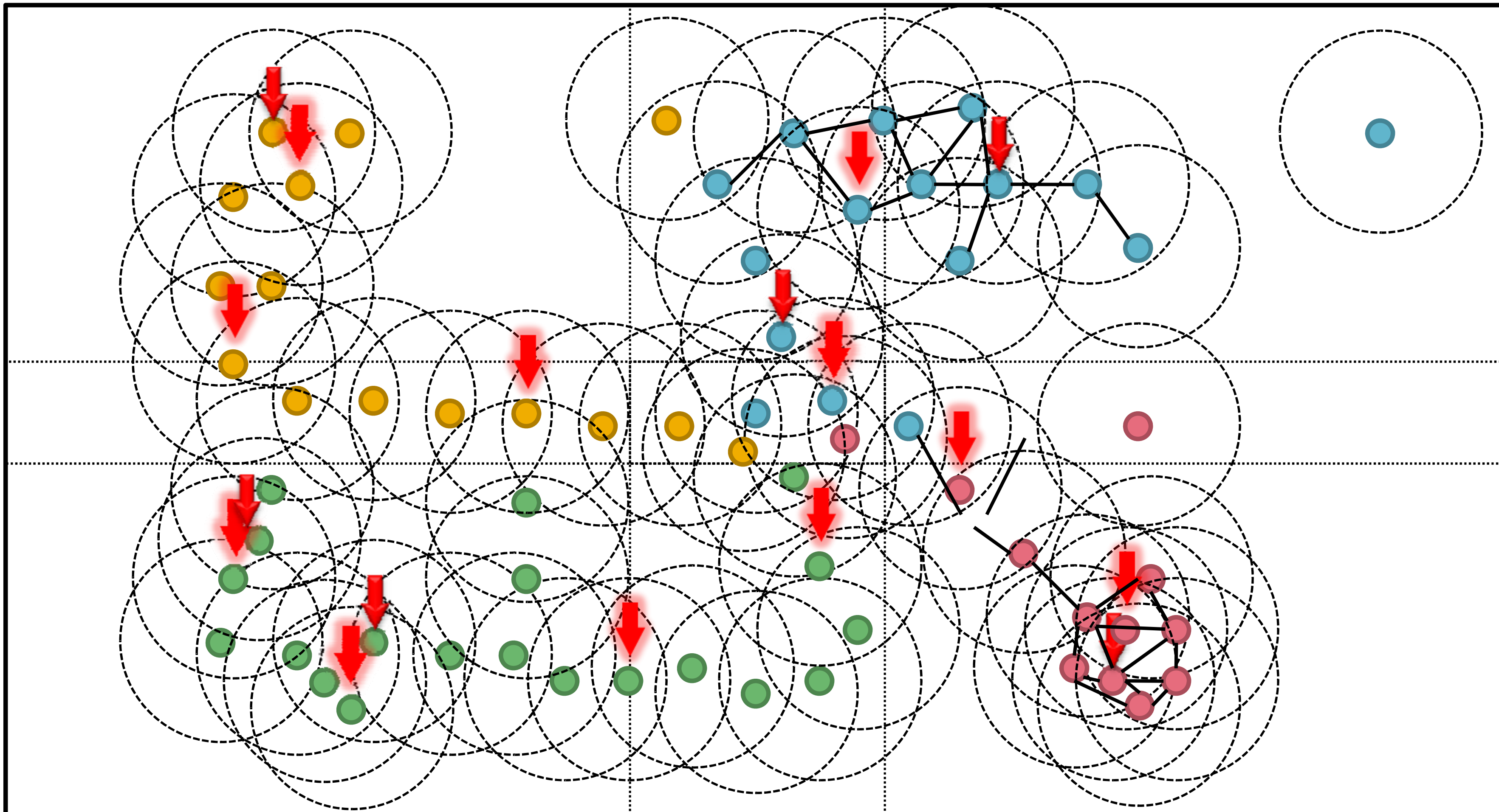


Ad Hoc InfoWare Middleware – Resource Management

- Provide a network resource aware view → well-connected regions and possible partitions of the network
- *Resource Management*
  - *Balance* the use and allocation of resources
  - *Eliminate* communication overhead and dependence on external services
  - Provide applications with information on remote resource *availability*

Enable adaptation to changing network configuration and capabilities of devices.

Resources & Service Placement – Example



Claims

1. *Non-intrusiveness*
   - Routing protocol holds updated information
2. Neighborhood *prediction*
   - Past adjacency information
   - No constraints
3. Network *clustering*
   - Current network layout
   - No constraints
4. *Feasible* on resource constrained devices

Overview

- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Requirements analysis for Resource Management

- General assumptions
  - Set-up phase for devices and personnel
  - Heterogeneous devices→ different capabilities and information
- Requirements
  - General solution (independent of routing protocol, device resource, and external service)
  - Local information
  - Preserve resources
  - Update resource information

- *Worst case scenario:* no external and additional information

Issues for Resource Management

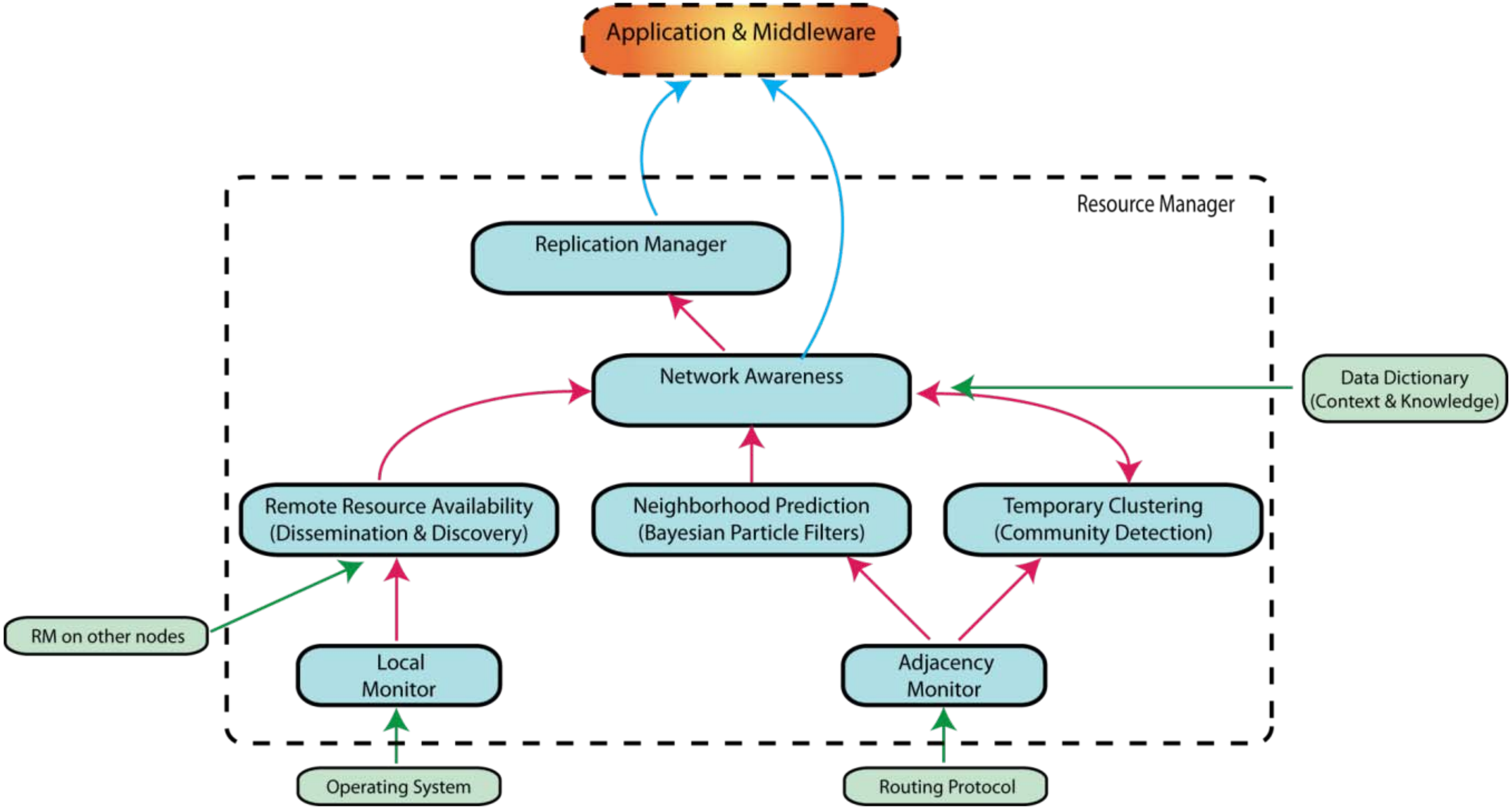
- **Goal**: Enable composability:
  - Remote resource sharing (*<u>Horizontal</u>*)
  - Local resources (Vertical)
  - Resource sharing profiles (Hierarchical)
- Remote resource sharing
  - Reliable communication channel → Discovering, Bookkeeping, Negotiation
  - Un-reliable communication channel → Resource Availability (e.g., link-lifetime)
    - Graceful degradation, Resource planning, Timed-soft-state-like reservations

Resource Management

- **Algorithm**: Link-lifetime Estimation
  - Classical: Global Positioning System (GPS)
    - Problems: imprecise measurements or lack of GPS coverage, physical proximity is not a guarantee of connectivity, needs information dissemination
  - Past Link-lifetime
    - Advantages: local information with no assumption on link-lifetime duration
- **Algorithm**: Network Clustering
  - Classical: Predefined number of clusters or cluster size
    - Problems: Dynamic of the network (number of nodes, connectivity, etc.)
  - Community Detection
    - Advantages: local information with no assumption on number of nodes in a cluster, number of clusters, network size and layout

- **Pros**: no assumption on location, no communication overhead
- **Cons**: more complicated algorithms, dependent on the accuracy of the information source

Information flow of Resource Management



Overview

- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions, and Future Work

Foundation for a non-intrusive RM

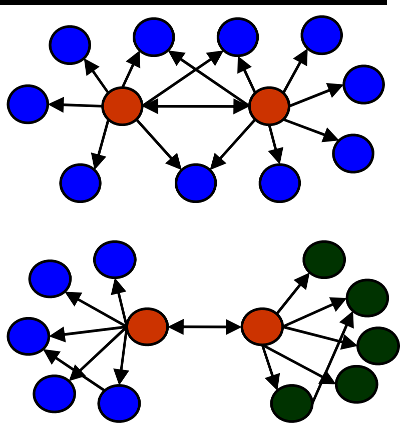
- Information Source: Routing table in the routing protocol
  - Advantages
    - Updated view of the network,
    - Location independent
  - Disadvantages
    - Sensitive to existence of communication
    - Sensitive to mobility and communication patterns
    - Partial topology of the network
  - Information type:
    - Neighborhood
    - Topology
  - Issues: accuracy and consistency

Test environment

<u>Mobility Model</u>	*Single mobility models*: Random Direction Mobility, Random Walk Mobility, Steady-State Random Waypoint *Group mobility models*: Random Point Group Mobility *Other simulators*: <u>***RoboCupRescue***</u> Simulation Project	
<u>Network simulator</u>	*GloMoSim*	Reactive Routing protocol: ***AODV***
		Proactive Routing Protocol: ***OLSR***
<u>Communication</u>	*Flat*: a random subset of the nodes are transmission sources *Structured*: follows the group structure	

General
- 40 nodes (4 groups)
- A=800x600u, T=18000s
- radius=100u,
- item=2048b,
- v=1÷7u/s

Communication
- Flat: tr_src>25%, per=50s
- Structured:
  - Leaders: tr_src=10%, per={30s,60s}
  - Group Nodes: tr_src>50%, per=30s



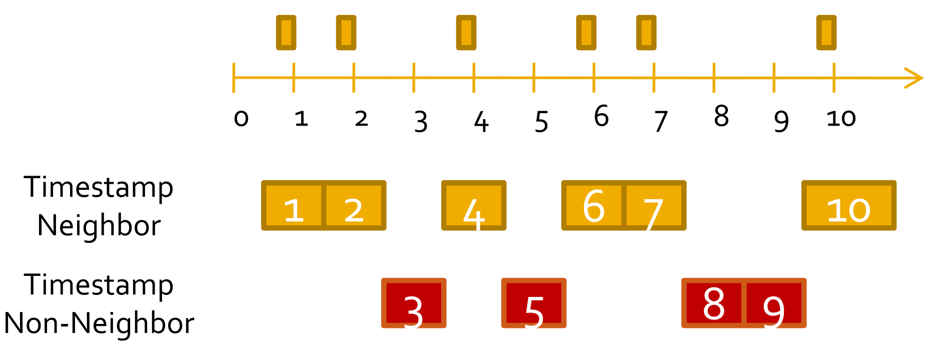
Neighborhood Connectivity Periods

- **Issue**: Accuracy of neighbor information in route table
- **Question**: Are neighbor periods similar in the scene (mobility model) and the routing protocol?
- **Solution**: Discrete time (routing protocol) approximate continuous time to time periods: Neighbor or Non-Neighbor (n_1, n_2)
  - Aggregate timestamps of direct contact between a pair of nodes
    - Interval $[0, T_{sim}]$ where $0 \le t_1 < \ldots < t_k \le T_{sim}$

      $T_{\sup} = \{t_{i+1} \mid t_{i+1} - t_i > \sigma\} \cup \{T_{sim}\} \quad T_{\inf} = \{0\} \cup \{t_i \mid t_{i+1} - t_i > \sigma\} \quad \sigma - time\ \ treshold$
    - Neighbor time periods

      $$C = \{[t_i, t_j] \mid t_i \in T_{\inf}, i = \{1, \dim(T_{\inf}) - 1\}, t_j \in T_{\sup}, j = \{2, \dim(T_{\sup})\}\}$$
    - Non-neighbor time periods

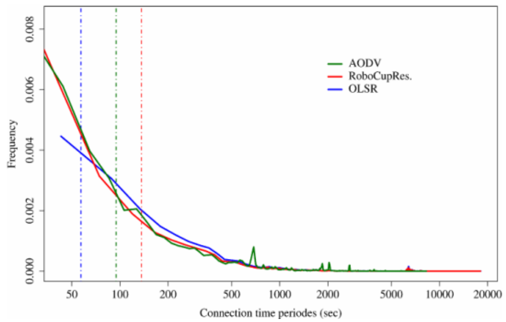
      $$D = \{(t_i, t_j) \mid t_i \in T_{\inf}, t_j \in T_{\sup}; i, j = \{1, \dim(T_{\inf})\}\}$$



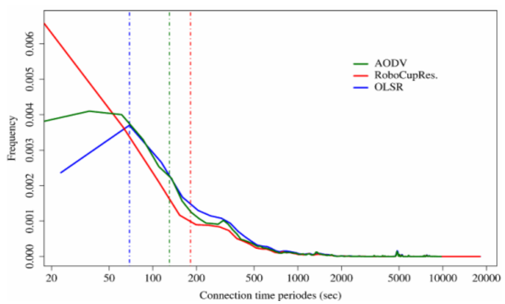
Neighbor Period $\{[1,2], 4, [6,7], 10\}$

Non-Neighbor Period $\{(0,1), (2,4), (4,6), (7,10)\}$

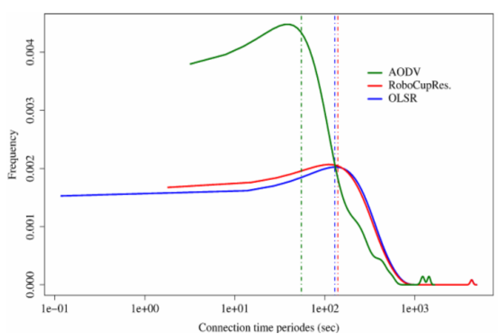
RoboCupRescue – Distribution of Connectivity Periods



Connections – Flat

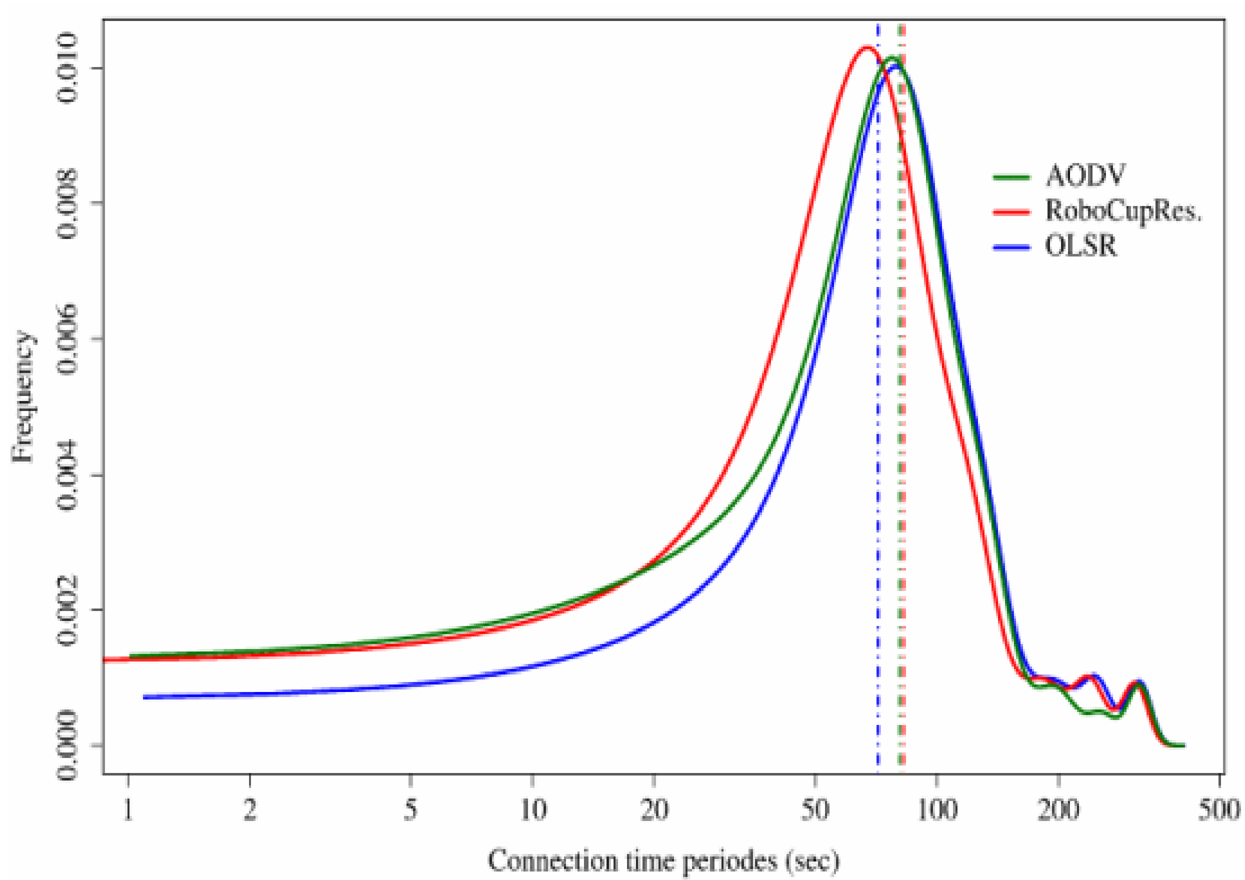


Connections – Struct.



Group Connections – Flat

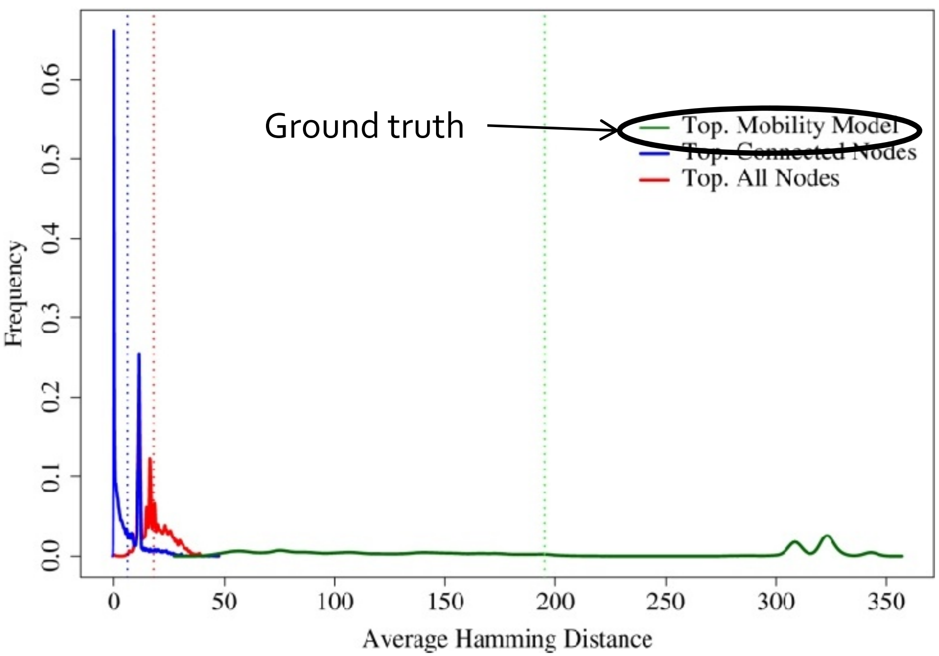
Result: The neighbor periods have the same distribution for mobility models and routing protocols.

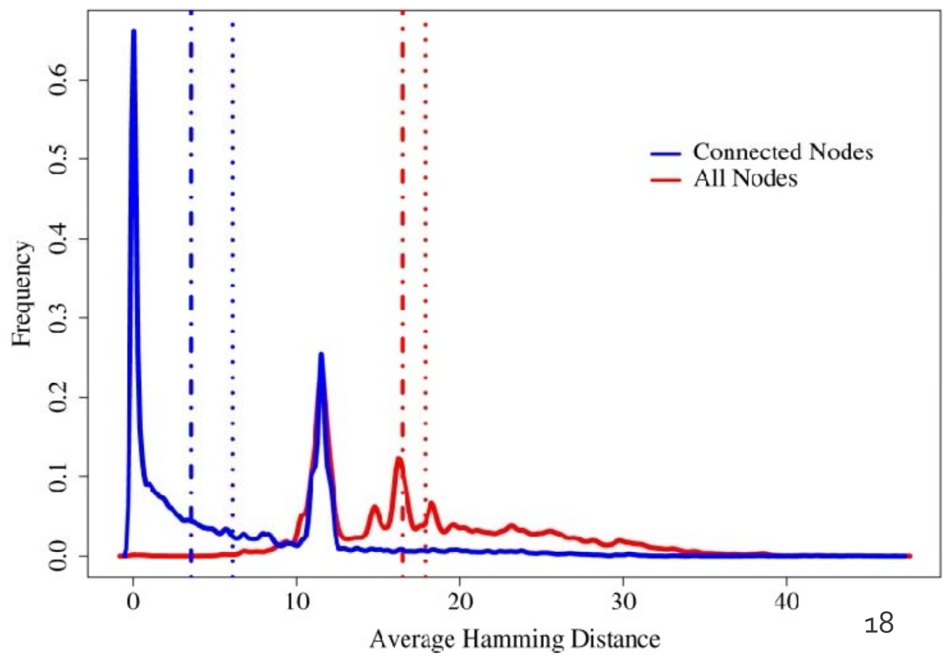


Group Connections – Structured communication

Topology Data Consistency

- **Issue**: Topology consistency (i.e., nodes may have different topology information)
- **Question**: How consistent is the topology information at the different nodes?
- **Solution**: Compare topology information on all the node in the network
  - Compute the Hamming Distance between topologies (count the differences between the adjacency matrixes of different nodes)
- **Result**: Similar topologies, if node are connected.





Overview

- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Prediction Algorithm

- **Question**: Is the neighborhood of a node predictable?
- **Idea**: Neighborhood estimation algorithm based on routing information
  - Active communication → routing protocol has good information on neighborhood
  - Discrete timestamps approximate continuous time → Neighbor or Non-Neighbor Periods
- **Algorithm:** *SIR Future Neighbor*
  - Window in a data stream (data process model)
    - Changes of direct connectivity between pairs of nodes
  - Sequential Monte Carlo (estimation method)
    - Estimates the future state of a node's neighborhood
    - No assumptions on: Links' lifetime and Nodes' movement

Window in a data stream

- **Solution**: Neighborhood window
  - Mapping from the time domain *T* to the domain of possible time intervals *Ts*

    $$\Psi : T_s \rightarrow \{[t_1, t_2] \mid t_1, t_2 \in T, t_1 \leq t_2\}$$

  - Aggregate timestamps of direct contact between two nodes
    - State of the pair (n_1, n_2) during an interval $[t_j, t_{j+1}]$ where

      $$\min(\psi(t)) \leq t_1 < \ldots < t_k \leq \max(\psi(t)) \text{ and } \psi(t) = [\min(\psi(t)), \max(\psi(t))]$$

  - Neighbor time period

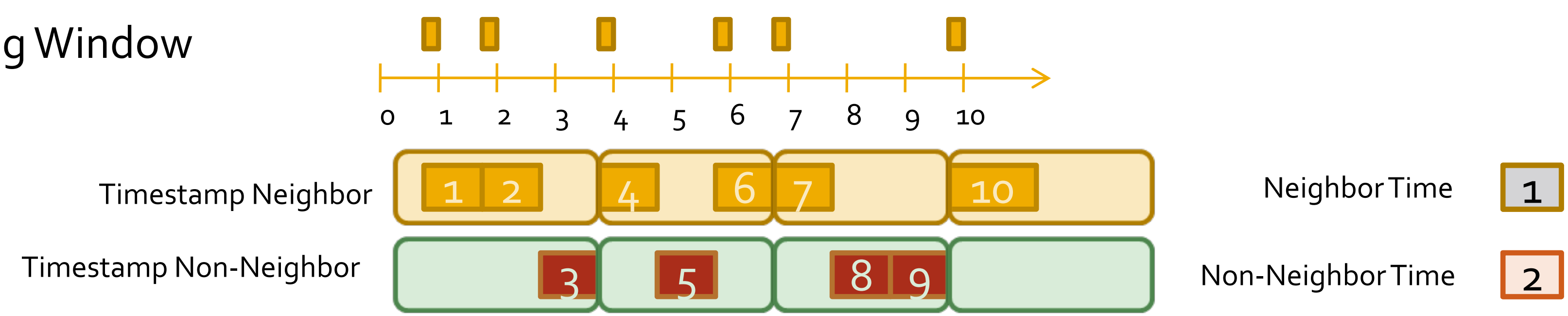
    $$C_{\Psi(t)}^{(n_1,n_2)} = \sum_{j=1}^{k-1} (t_{j+1} - t_j) \quad if \quad t_{j+1} - t_j \leq \sigma \quad \sigma - time\,treshold$$

  - Non-neighbor time period

    $$D_{\Psi(t)}^{(n_1,n_2)} = (\max(\Psi(t)) - \min(\Psi(t))) - C_{\Psi(t)}^{(n_1,n_2)}$$

- **Assumptions**: Distributions of periods are non-linear / non-Gaussian

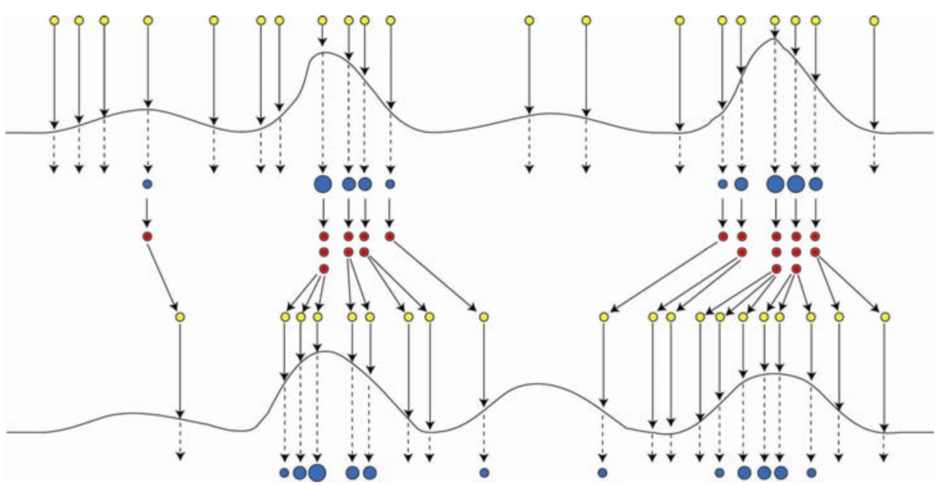
Tumbling Window



Sequential Monte Carlo (SMC)

- *Recursive Bayesian Tracking*: Represent the posterior density function by a set of random samples with associated weights and compute the estimates based on samples and weights
- SMC for the neighbor pair (n_1, n_2)
  - Recursively calculate the degree of belief in the estimation x_{τ_k} for the next window interval τ_k taking into consideration the values of the current measurements $z_{\tau_{1:k}}$
    - Predict: use previous measurements for prior $p(x_{\tau_k} \mid z_{\tau_{1:k-1}})$
    - Update: likelihood function $p(z_{\tau_k} \mid x_{\tau_k})$ → prior distribution $p(x_{\tau_k} \mid z_{\tau_{1:k}})$
      - Sampling Importance Resampling (SIR)

$i = 1, \dots, N_s = 20 \text{ particles}$



$\{x^i_{\tau_{k-1}}, N_s^{-1}\}$

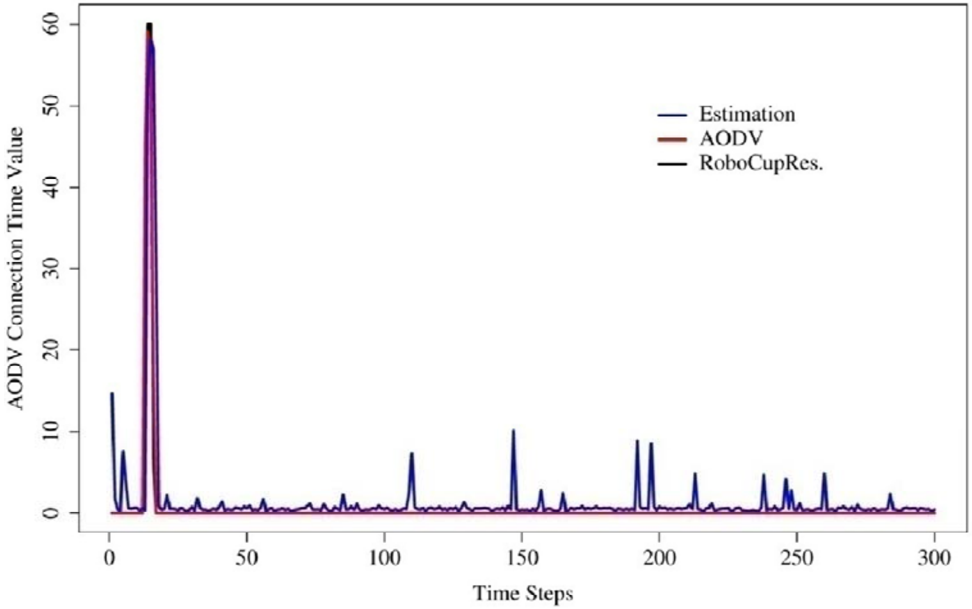
$\{x^i_{\tau_{k-1}}, w^i_{\tau_{k-1}}\}_{i=1}^{N_s}$

$\{x^i_{\tau_{k-1}}, N_s^{-1}\}$

$\{x^i_{\tau_k}, N_s^{-1}\}$

$\{x^i_{\tau_k}, w^i_{\tau_k}\}_{i=1}^{N_s}$

Tumbling Window – SIR Future Neighbor



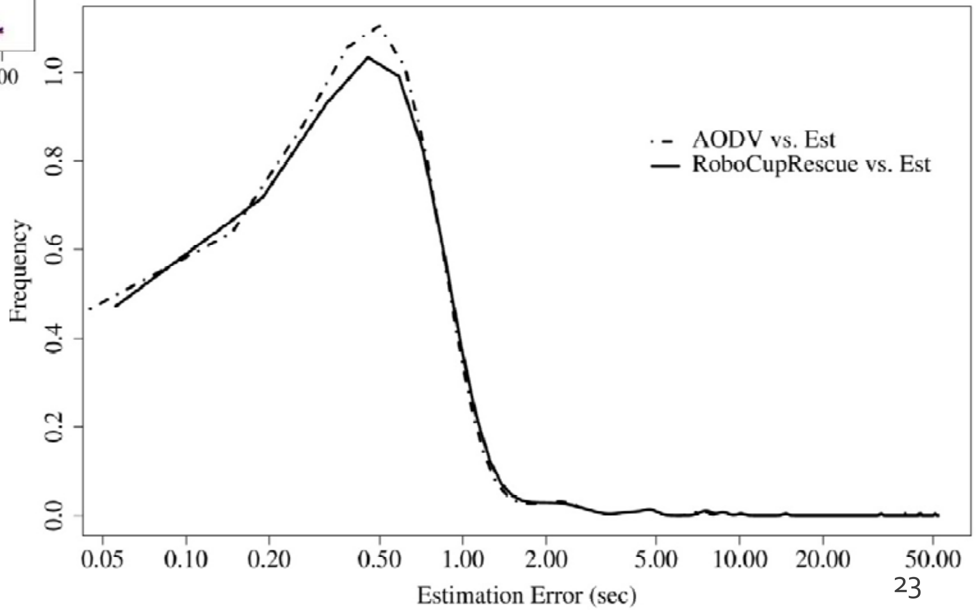
Connection time

Window type: Tumbling → Estimate a value between 0 and ϕ

$C_{\Psi(t)}^{(n_1,n_2)}, D_{\Psi(t)}^{(n_1,n_2)} \in [0,\phi]$

$\phi = 60$

Euclidian Distance



Result: The estimated values are close to the measured and “true” values.

Prediction Discussions

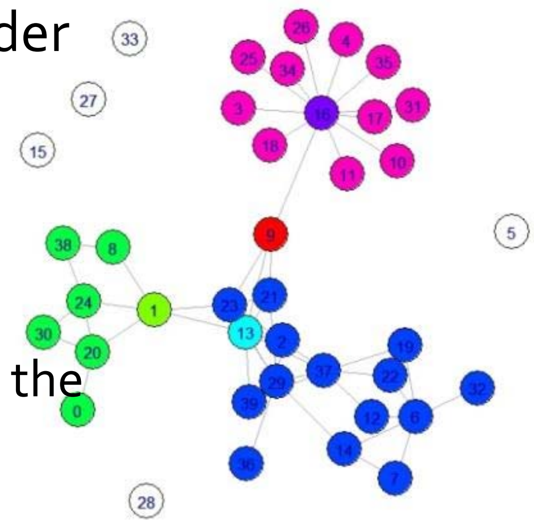
- Window Size (ϕ) vs. Time Step (β)
  - Gives the type of window, Impacts the performance and resources required, β does not influence accuracy
- Reactive routing protocol
  - No communication → No accuracy
  - Windowing techniques can filter small communication pauses
- Linear Cost with Number of Samples (Ns)
  - Number of Samples (Ns) and Time Step (β): Large β → Large Ns
  - Number of Samples (Ns) and Number of neighbors: High number of neighbors → High computation cost
- Run continuously
  - Adapts to the behavior of the neighborhood, Correct after a few estimations
  - No on-demand estimations

Overview

- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Temporary services

- **Question**: Where to place services in the network?
- **Issue**: Minimize the distance to resources in order to balance the use of resources
  - *Requirements*
    - Management overhead independence
    - Position independence
  - *Solution*: Temporary Clustering with dynamic clustering methods (i.e., consider the dynamic in the network)
    - Clustering which adapts to the current network layout
      - Adaptive number of clusters
      - Unconstrained number of nodes in a cluster
    - Temporary service positioning: Number of data replica and services, Data and service placement, Network partitioning
  - *Problem*: No methods to handle dynamics issues of MANETs



Passive Clustering

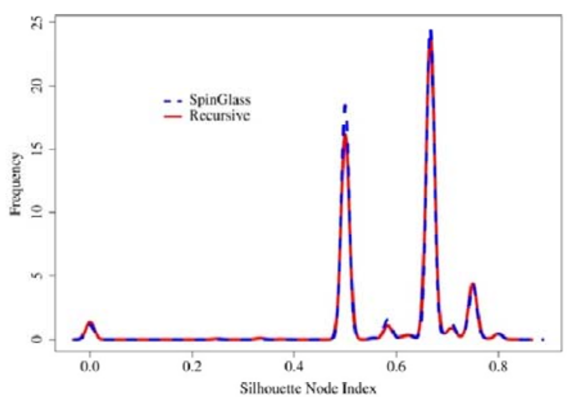
- **Idea**: Clustering based on the network topology in the route table
- **Methods**: use static networks algorithms
  - *Clustering*: Community Detection
    - Divide or agglomerate to detect the groups of nodes in the network with dense network connections, and sparse connections outside the groups
    - 1. Newman and Grivan 2004 (Recursive) where recursively eliminates the links with high weight in the network
    - 2. Reichard and Bornholdt 2006 (SpinGlass) where community membership of a node is determined by its neighborhood (i.e., number of neighbors and neighbors' membership)
  - *Placement*: Cluster head election based on centrality measures
  - *Coverage*: Network Voronoi Node Diagram
    - For nodes in a given set it defines: dominance set and bisector set nodes
  - *Evaluation*: measure quality, stability, similarity, and consistency

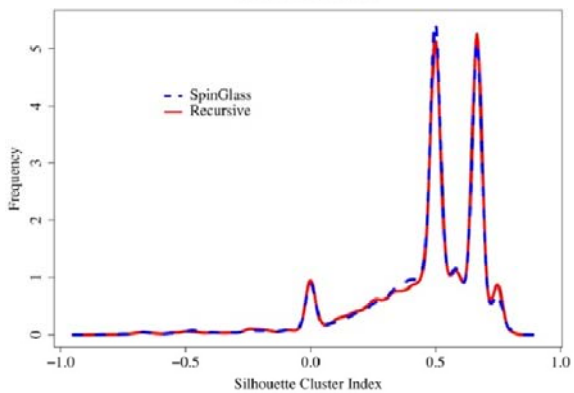
Clustering Evaluation: Quality

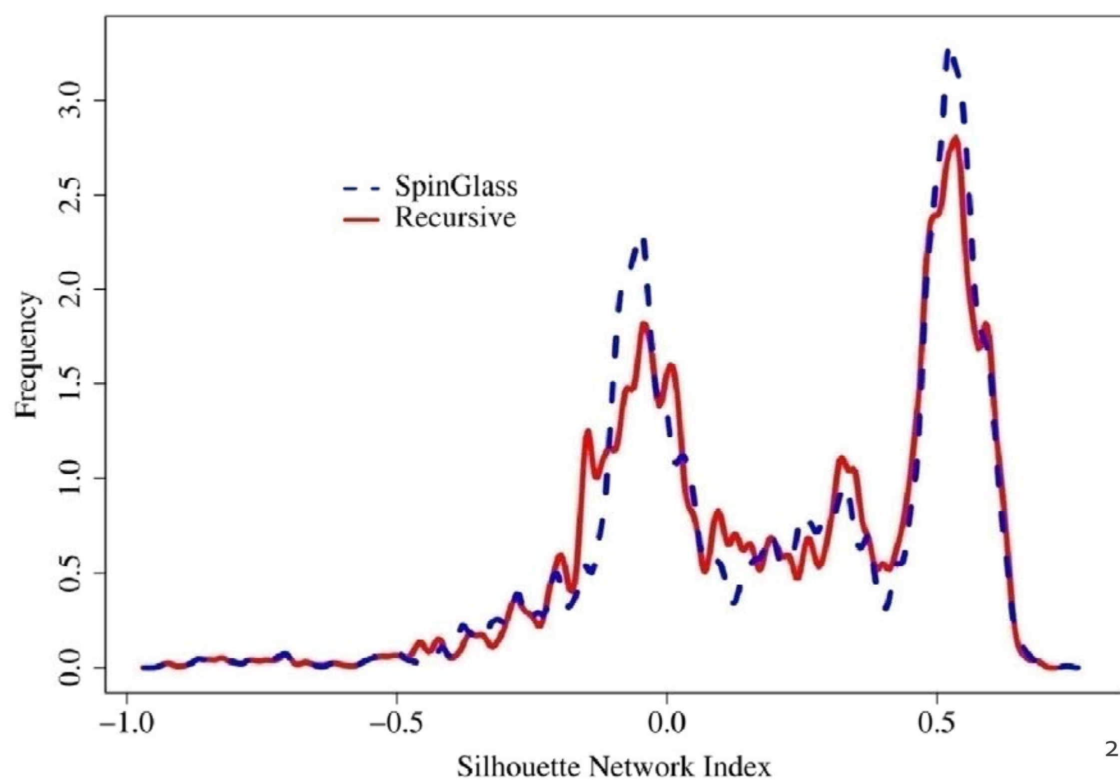
- **Question**: Is the clustering valid?
- **Measure**: Silhouette index (how well is a node clustered considering its distance to the center and of the center of the closest cluster)

Node: $s(p_i) = \frac{\bar{d}(p_i, C_h) - \bar{d}(p_i, C_j)}{\max(\bar{d}(p_i, C_h), \bar{d}(p_i, C_j))}$ Cluster: $S_j = \frac{\sum_{i=1}^{N_j} s(p_i)}{N_j}$ Network: $GS = \frac{\sum_{j=1}^{k} S_j}{k}$

- **Results**: The created clusters are good.

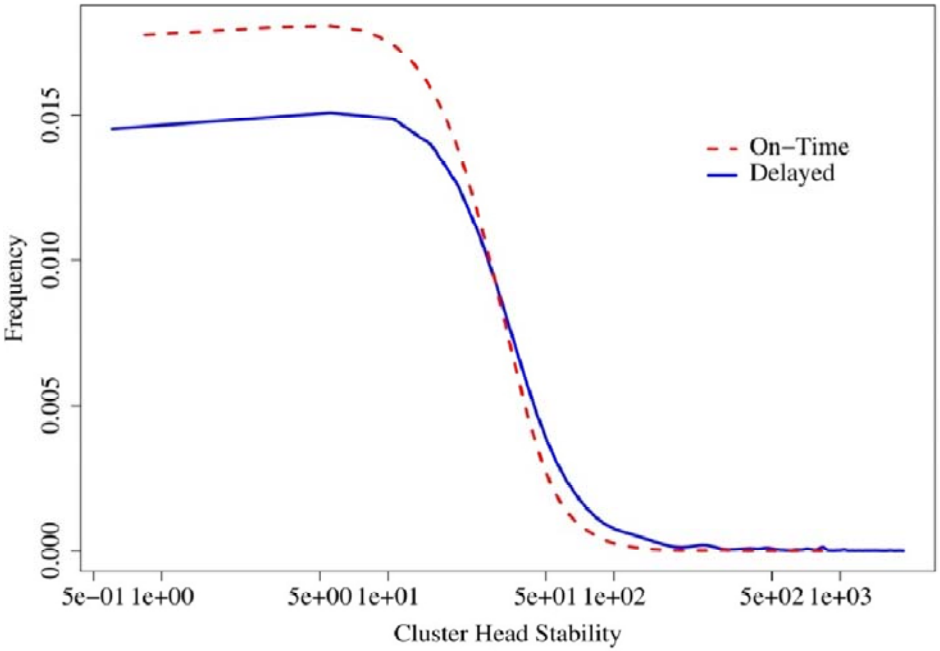


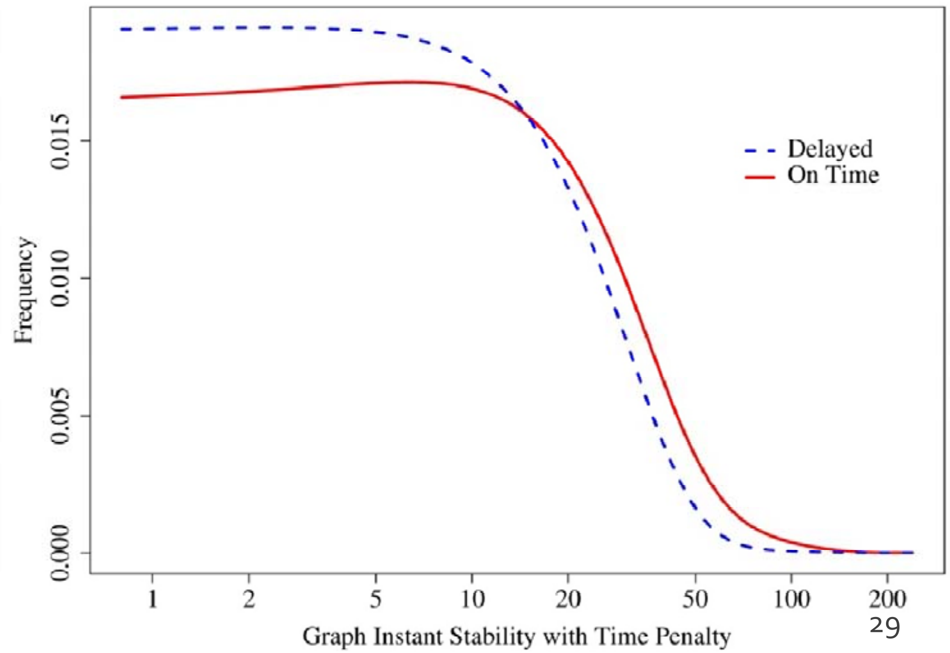




Clustering Evaluation: Stability

- **Question**: Is the clustering stabile?
- **Measure**: Stability quantifies the changes of the clustering with respect to the new network structure (i.e., establishment and breakage of links, and quality increase and decrease cluster membership)
  - Delay the clustering $S_i = \left[\frac{Q_i + F_i}{B_i + E_i} \right] * \frac{1}{\delta t}$
- **Results**: Delayed clustering can improve the stability of the clusters.



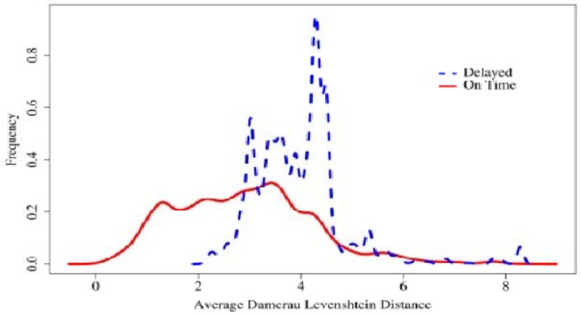


Clustering Evaluation: Similarity

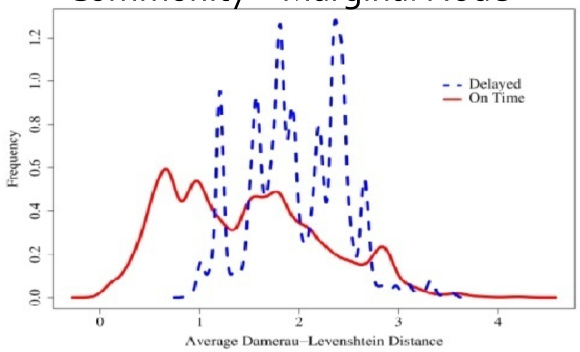
- **Question**: What is the difference between different clusterings?
- **Measure**: The similarity measures the variation of information between clustering over the same network
  - SpinGlass vs. Recursive
    - The methods return similar clustering, i.e., in the interval 0 and 0.8
  - SpinGlass vs. Voronoi and Recursive vs. Voronoi
    - The Voronoi Diagrams are also similar to the clustering's, i.e., in the interval 0.1 and 0.9
- **Result**: The different clustering techniques produce similar results.

Clustering Evaluation: Consistency

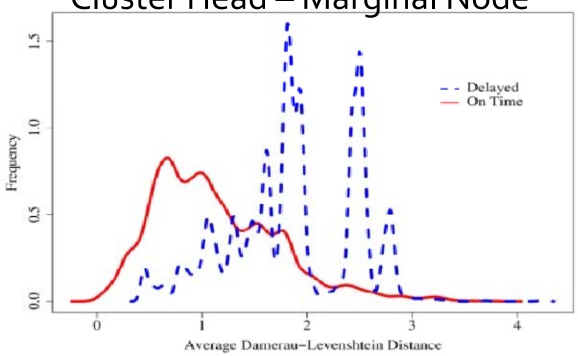
- **Question**: Are the clusters consistent in the network?
- **Measure**: Damereau-Levenshtein Distance between detected communities at different nodes (i.e., no. of insertions, deletions, substitutions of single characters, and transpositions between two sets)
- **Results**: Communities are similar at different nodes.



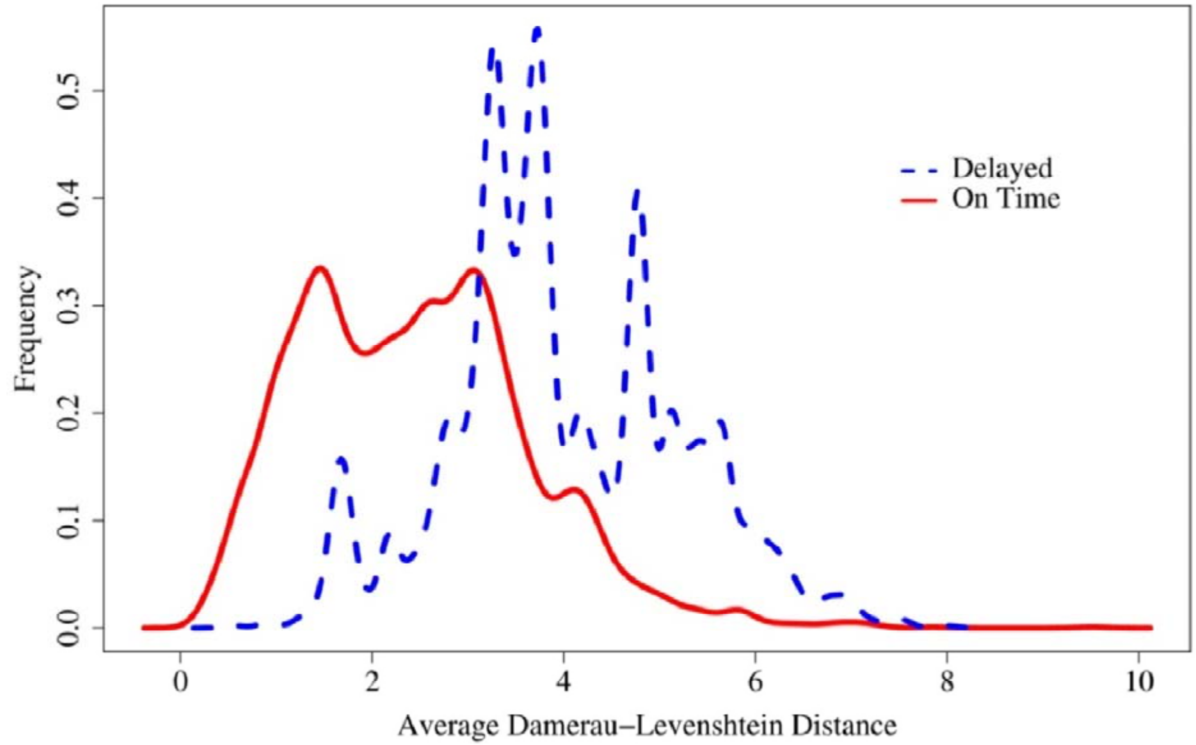
Community – Marginal Node



Cluster Head – Marginal Node



Cluster Head – Central Node



Community – Central Node

Clustering Discussion

- Topology information
  - Does not apply for reactive routing protocols
  - Requires a consistent view of the topologies at the nodes
- Clustering measures
  - Not for dynamic networks
  - Quality
    - Not a general accepted metrics
    - Different metrics may give contradicting conclusions
  - Stability
    - Does not consider the changes in the number of clusters
    - Does not consider the changes in the number of nodes in the network
  - Similarity:
    - Not applicable to clustering from different nodes in the network
  - Consistency
    - Detected communities are more consistent than elected cluster heads

Overview

- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Implementation – Nokia 770 Internet Tablet

- SIR Future Neighbor
  - Values
    - Footprint 28kb
    - Routing protocol: OLSRD
    - $\beta=\phi=\{30,60,120,300\}$
    - Ns= 20, T = 5200
    - Number of nodes = 5 – 25 , Simulated neighborhood change every 4 s
  - Computation time per step
    - Tumbling Window process ~ 0.03 – 2 ms
    - SIR Future Neighbor process ~ 2 – 40 ms

Image from www.nokia.com

- Network Voronoi Node Diagrams
  - Values
    - Footprint 38.5kb
    - Routing protocol: OLSRD
    - Number of nodes = 10 – 15 , Number of clusters = 2 – 4

Result: Possible to run on resource constraint devices.

Overview

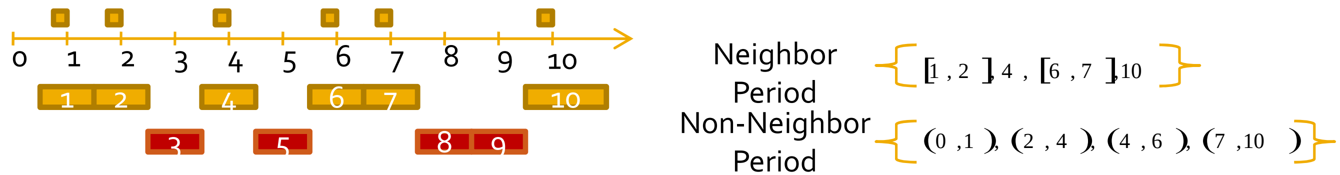
- Introduction
- Communication Non-intrusive Resource Management
  - A non-intrusive information source
  - Prediction
  - Clustering
  - Feasibility study
- Conclusions and Future Work

Conclusions

- Thesis contributions
  - Middleware architecture for RM in Sparse MANETs
  - Information service for worst case scenario to deliver updated availability of remote services and resources
  - Communication non-intrusive algorithms for prediction and clustering
- Thesis claims: Critical review
  1. Non-intrusiveness
     - Good if the node is involved in communication
  2. Neighborhood prediction
     - Non-intrusive is possible
     - Dependent on communication
     - Cannot perform timely predictions
  3. Network Clustering
     - Community detection gives good results
     - Can benefit from delaying the process
     - Only for proactive protocols
  4. Feasible on resource constrained devices

Future Work

- Integrate higher level information
- Timely predictions
- Clustering for reactive protocols
- Use predictions and adaptive clustering techniques (which support the dynamics of the network)
- Apply the algorithms to delay tolerant streaming
- Evaluate with real mobility and communication traces
- Evaluate during a rescue exercise



Thank you, for your attention!

Questions?