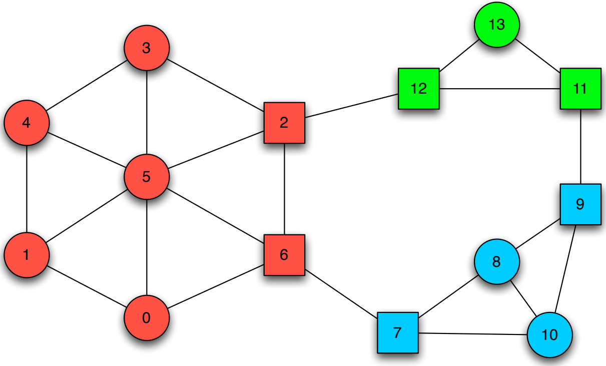
Communities and Roles in Wireless Networks

Florian Reith
florian.reith@stud.tu-darmstadt.de

TECHNISCHE UNIVERSITÄT DARMSTADT



Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

- Introduction and basic definitions
- Realistic network model for wireless networks
- Universal roles in wireless networks
- Roles for wireless networks
- Distributed community detection
- Role detection
- Evaluation

Graphs

- Graphs are denoted as $G = (V, E)$
- V is a set of vertices
- E is a set of edges
- Edges are unordered pairs $\{u, v\} : u, v \in V$
- Edges are undirected
- Edges are not weighted
- For an edge $\{u, v\}$ the vertices u and v are called ends of the edge
- At most 1 edge between vertices u and v

What are communities and roles?

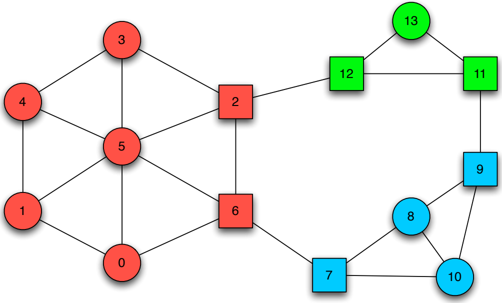
TECHNISCHE UNIVERSITÄT DARMSTADT

Communities

- Groups of vertices stronger connected internally than externally

Roles

- Classes for vertices with similar properties



Why are communities and roles interesting?

- Network nodes need information to make good decisions (eg routing)
- Bandwidth is limited in wireless networks
- Energy is limited for mobile devices (Battery)
- Nodes should gather information with minimal message overhead
- Communities were already used to optimize routing
- Roles can probably be used for optimization

Community definition

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (Strong community)

The subgraph V is a community in a strong sense, if

$$k_i^{in}(V) > k_i^{out}(V), \forall i \in V$$

Definition (Weak community)

The subgraph V is a community in a weak sense, if

$$\sum_{i \in V} k_i^{in}(V) > \sum_{i \in V} k_i^{out}(V)$$

where k_i^{in} is the number of neighbors of vertex i in community V and ${k_i}^{out}$ the number of neighbors, that do not belong to V.

Modularity

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (Modularity)

The modularity Q is a metric for the quality of a community structure.

$$Q = \sum_{c \in C} \left(\frac{|I_c|}{|E|} - \left(\frac{2 \cdot |I_c| + |E_c|}{2 \cdot |E|} \right)^2 \right),$$

where C is the set of communities,
$|I_c|$ the number of internal edges with both ends in community c,
$|E_c|$ the number of external edges with only one end in c,
$|E|$ denotes the total number of edges of the graph.

Relative within-module degree z

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (relative within-module degree)

z_i is a measure for how well a vertex i is connected within its own community

$$z_i = \frac{k^i_{s_i} - \langle k^j_{s_i} \rangle_{j \in s_i}}{\sqrt{\langle (k^j_{s_i})^2 \rangle_{j \in s_i} - \langle k^j_{s_i} \rangle^2_{j \in s_i}}},$$

where k^i_s is the number of edges between vertex i and vertices belonging to community s and s_i is the community of vertex i.

Participation coefficient P

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (participation coefficient)

P_i is a measure for how well vertex i is connected to all existing communities.

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{k_s^i}{k_i} \right)^2 ,$$

where N_M is the number of existing communities,
k_i is the degree of vertex i,
k_s^i is the number of edges between vertex i and vertices from community s.

Role definition by Guimerá et al.

TECHNISCHE UNIVERSITÄT DARMSTADT

Role	Name	P	z
R1	Ultra-peripheral node	$P \leq 0.05$	$z < 2.5$
R2	Peripheral node	$0.05 < P \leq 0.62$	
R3	Satellite connector	$0.62 < P \leq 0.80$	
R4	Kinless node	$P > 0.80$	
R5	Provincial hub	$P \leq 0.30$	$z \geq 2.5$
R6	Connector hub	$0.30 < P \leq 0.75$	
R7	Global hub	$P > 0.75$	

Network model

- Combination of connection model and placement model
- Connection model creates edges between vertices
- Placement model used to place vertices on a 2-dimensional field
- Combination of models is used to create graphs
- Relatively realistic model for wireless networks

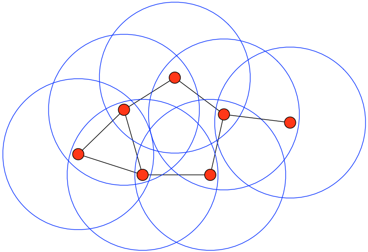
Connection models

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (Unit Disk Graph)

A unit disk graph is an undirected euclidean graph (V, E), where $V \subseteq \mathbb{R}^2$ and $\forall u, v \in V$:

- $d(u, v) \leq 1 \Leftrightarrow \{u, v\} \in E$



Definition (Quasi Unit Disk Graph)

A r-quasi unit disk graph is an undirected euclidean graph (V, E) with parameter $r \in [0, 1]$, where $V \subseteq \mathbb{R}^2$ and $\forall u, v \in V$:

- $d(u, v) \leq r \Rightarrow \{u, v\} \in E$
- $d(u, v) > 1 \Rightarrow \{u, v\} \notin E$

Placement models

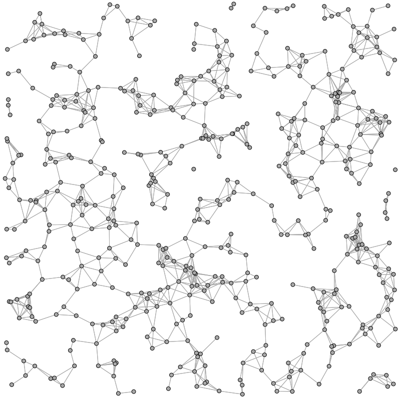
TECHNISCHE UNIVERSITÄT DARMSTADT

- Random
  - distributes vertices uniformly on the field
- Grid
  - vertices are placed on a grid
- Circle
  - vertices are placed on a circle
  - vertices are either placed on the radius or with a random deviation
  - vertices are either placed equally spaced or randomly in their slice
- Gaussian
  - vertices are placed normally distributed around a center
- Hotspot
  - combination of 2 models
  - first model is used to place hotspots
  - second model is used to place vertices at each hotspot

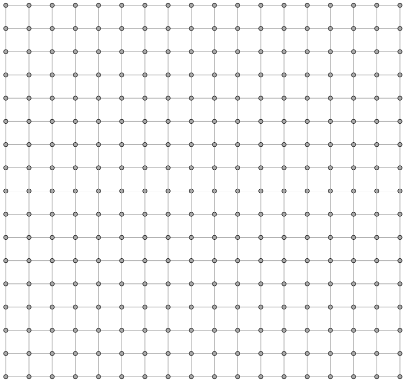
Placement model examples

TECHNISCHE UNIVERSITÄT DARMSTADT

Random



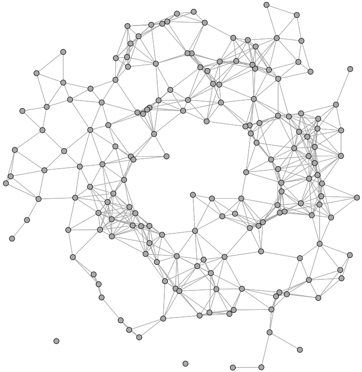
Grid



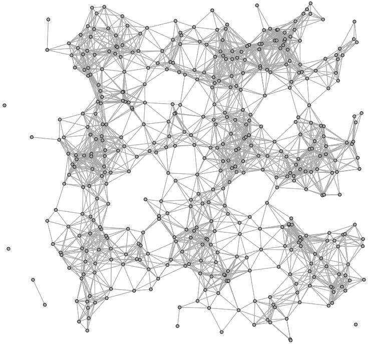
Placement model examples

TECHNISCHE UNIVERSITÄT DARMSTADT

Circle with normally distributed α and *d*



Normal distribution around hotspots on a grid



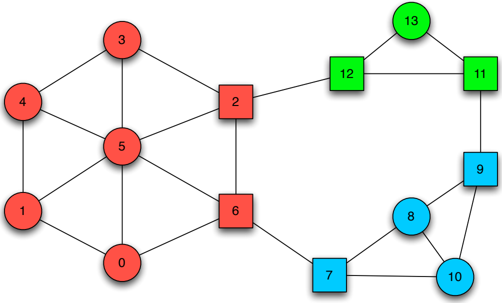
Roles by Guimerá et al. in wireless networks

TECHNISCHE UNIVERSITÄT DARMSTADT

- Unit disk graph is used to model wireless network
- Two implications:
  1. neighbors of neighbors are likely connected
     - difference between $k^i_{s_i}$ and $\langle k^j_{s_i} \rangle_{j \in s_i}$ is reduced
     - z_i is reduced
  2. vertices can not be connected over great distance
     - P_i is influenced
     - vertex in center of community is unlikely connected to foreign community
- Edge of community best location for high P
- Center of community best location for high z
- Hard to achieve both

Roles by Guimerá et al. in wireless networks

Vertex	P	z	Role
0	0.0	-0.409	R1
1	0.0	-0.409	R1
2	0.375	-0.409	R2
3	0.0	-0.409	R1
4	0.0	-0.409	R1
5	0.0	2.45	R1
6	0.375	-0.409	R2
7	0.444	-1.0	R2
8	0.0	1.0	R1
9	0.444	-1.0	R2
10	0.0	1.0	R1
11	0.444	0.0	R2
12	0.444	0.0	R2
13	0.0	0.0	R1



Number of adjacent communities

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (Number of adjacent communities)

The number of adjacent communities AC_i of a vertex i denotes the number of distinct communities occuring in the neighborhood of vertex i.

$$AC_i = \sum_{j \in C} \delta_{ij} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i \text{ has at least one neighbor in } j \\ 0 & \text{otherwise} \end{cases},$$

where C is the set of all existing communities.

Upper bound on adjacent communities

TECHNISCHE UNIVERSITÄT DARMSTADT

Theorem

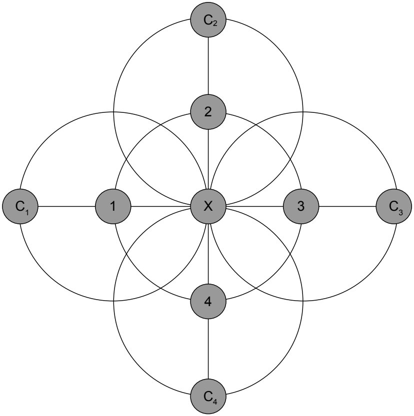
For r-quasi unit disk graphs and communities in a strong sense the number of communities adjacent to a vertex is bound from above by $\left\lfloor \frac{360^\circ}{2 \cdot \arcsin(\frac{r}{4})} \right\rfloor$*, where r is given by the r-quasi unit disk graph.*

Constructive proof

TECHNISCHE UNIVERSITÄT DARMSTADT

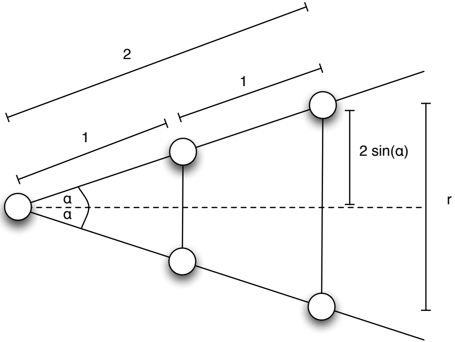
Construction

1. Neighbors i equidistant on unit disk around x
2. Community C_i on unit disk around i, so that $d(C_i, x) = 2$
3. C_i and C_j are not connected, as long as $d_{ij} > r$
4. C_i and C_j are one community for $d_{ij} \leq r$ (Strong community)
5. Add neighbors until $d_{ij} \leq r$



Constructive proof

TECHNISCHE UNIVERSITÄT DARMSTADT



- Upper bound for AC_i is $\left\lfloor \frac{360^\circ}{2 \cdot \arcsin(\frac{r}{4})} \right\rfloor$. For Unit Disk Graph ($r = 1$): 12

Role definition for wireless networks

Role	Name	AC	z
C	Common	$AC < 3$	$z < 1$
B	Bridge	$AC = 3$	
S	Star	$AC > 3$	
HC	Common hub	$AC < 3$	$z \geq 1$
HB	Bridge hub	$AC = 3$	
HS	Star hub	$AC > 3$	

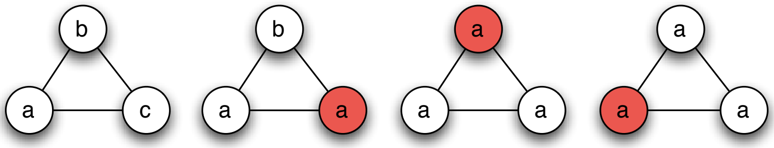
Label Propagation Algorithm

TECHNISCHE UNIVERSITÄT DARMSTADT

- Fast iterative algorithm
- Designed to have low runtime
- Results do not have optimal modularity
- Already partially distributed
- Label updates need only local knowledge
- Still global update order and break condition
- Asynchronous version to prevent label oscillation

Label Propagation Algorithm

1. Initialize the labels for all vertices. For each vertex x, $C_x(0) = x$.
2. Set $t = 1$
3. Arrange the vertices in a random order and set it to X
4. For each $x \in X$ in that order set $C_x(t)$ to the label that occurs with the highest frequency among the neighbors of vertex x. Ties are broken uniformly at random.
5. If every vertex has the same label as the maximum of its neighbors stop, else set $t = t + 1$ and go to 3.



Distributed Label Propagation Algorithm

TECHNISCHE UNIVERSITÄT DARMSTADT

Distributed update order

- Iterations are local time intervals of 60s
- Random delay $\in [0, 1[$ after each iteration

Label update mechanism

- Broadcast label updates to neighbors
- Update label based on last message from each neighbor, prefer own label

Local break condition

1. Label was not changed in 6 latest iterations
2. No neighbor changed label in 6 latest iterations

Role detection

- P_i can be calculated locally

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{k_s^i}{k_i} \right)^2$$

- AC_i can be calculated locally

$$AC_i = \sum_{j \in C} \delta_{ij} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i \text{ has at least one neighbor in } j \\ 0 & \text{otherwise} \end{cases}$$

- z_i needs average $k_{s_i}^j$ and $(k_{s_i}^j)^2$ from entire community

$$z_i = \frac{k_{s_i}^i - \langle k_{s_i}^j \rangle_{j \in s_i}}{\sqrt{\langle (k_{s_i}^j)^2 \rangle_{j \in s_i} - \langle k_{s_i}^j \rangle_{j \in s_i}^2}}$$

Gossiping

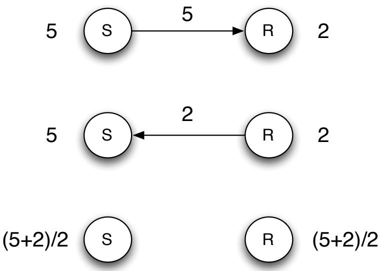
- Iterative push-pull algorithm
- Aggregates average *a*
- Conservation of true average

Sender (Push)

1. send a_s to random neighbor
2. on answer set $a_s = (a_s + a_r)/2$

Receiver (Pull)

1. on request send a_r to sender
2. set $a_r = (a_s + a_r)/2$



Gossiping issues

TECHNISCHE
UNIVERSITÄT
DARMSTADT

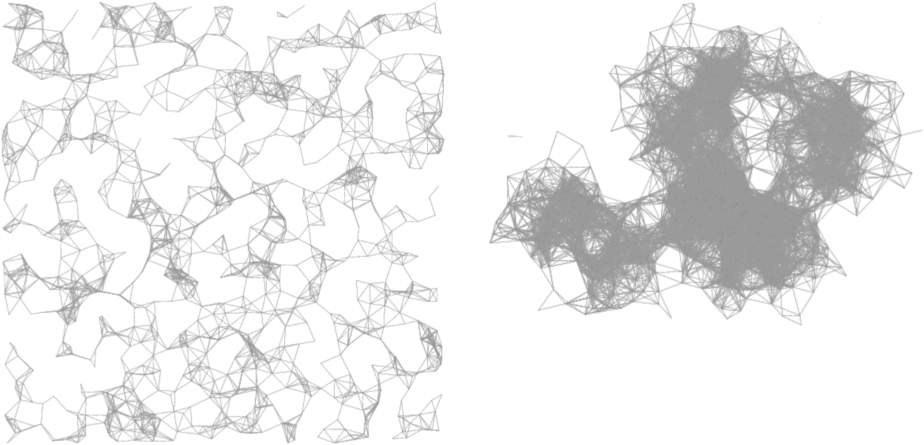
- Unicast communication
- High message overhead
- *a* is corrupted if sender receives request while waiting for answer
- Solution: do not answer while waiting and let him rerequest
- Even more message overhead
- Slow algorithm runtime
- Gossiping deactivated for evaluation

Evaluation setup

TECHNISCHE UNIVERSITÄT DARMSTADT

- SIDnet-SWANS simulator used for evaluation
- Unit disk graph connection model
- fixed field size
- 1000 to 10000 vertices in steps of 1000
- 2 configurations for placement models
- Random configuration
  - 1000 to 10000 vertices randomly placed on field
- Hotspot configuration
  - 10 to 100 hotspots placed normal distributed around center of field
  - for every hotspot 100 vertices placed normal distributed around the hotspot

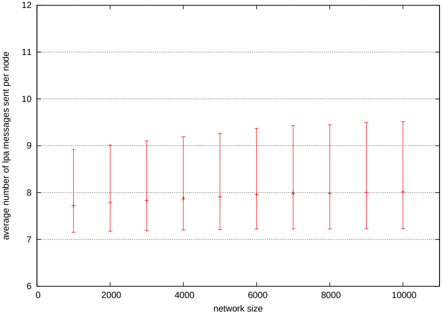
Examples with 1000 vertices



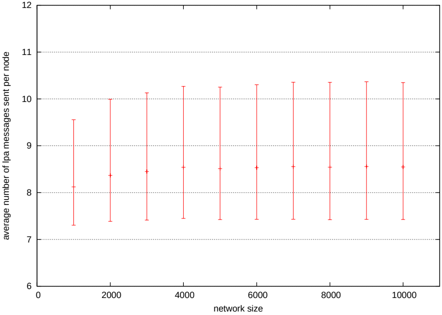
Label propagation message overhead

TECHNISCHE UNIVERSITÄT DARMSTADT

random configuration



hotspot configuration

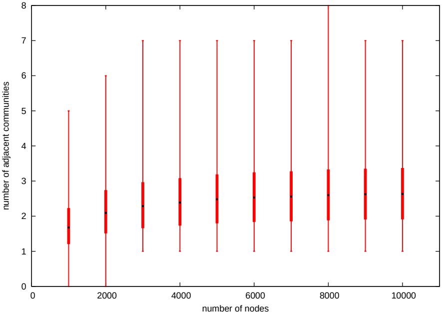


- Message overhead is independent of network size

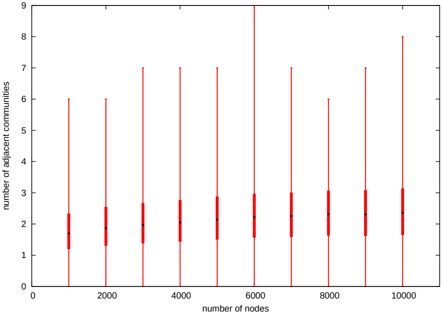
Maximum number of adjacent communities

TECHNISCHE UNIVERSITÄT DARMSTADT

random configuration



hotspot configuration

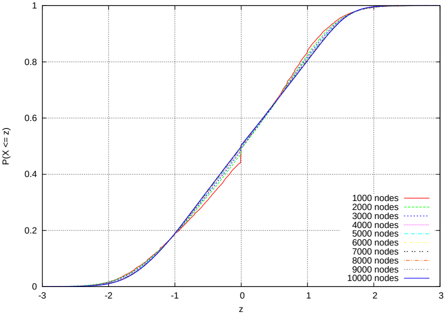


- Maximum number of adjacent communities is lower than theoretical limit of 12

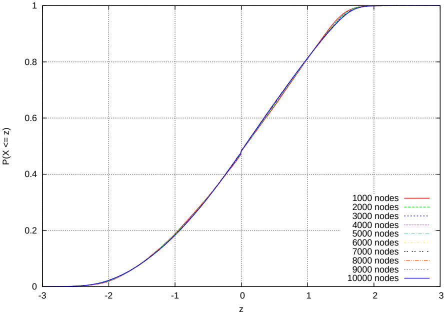
Relative within-module degree z

TECHNISCHE UNIVERSITÄT DARMSTADT

random configuration

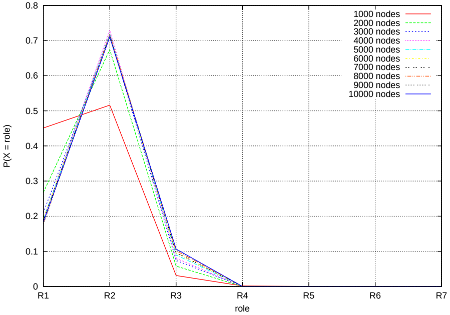


hotspot configuration

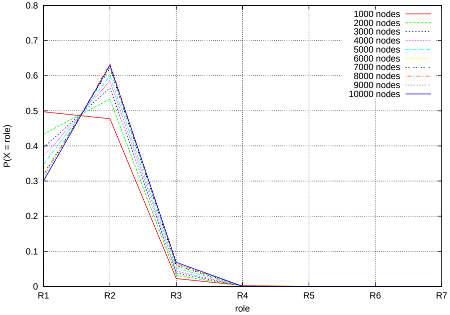


Frequency of roles by Guimerá et al.

random configuration



hotspot configuration

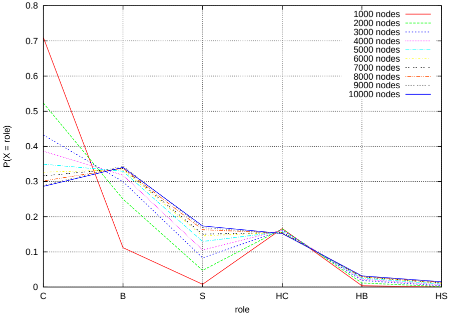


- In wireless networks roles by Guimerá et al. have limited information value.

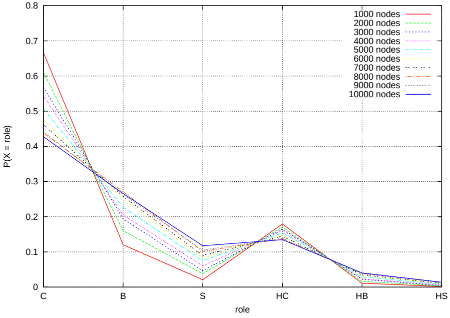
Frequency of roles for wireless networks

TECHNISCHE UNIVERSITÄT DARMSTADT

random configuration



hotspot configuration



Conclusion and Outlook

TECHNISCHE UNIVERSITÄT DARMSTADT

Conclusion

- Described realistic network model for wireless networks
- Found drawbacks of universal roles in wireless networks
- Found proof for upper bound of adjacent communities
- Introduced role definition for wireless networks
- Implemented distributed community detection
- Evaluated universal roles and roles for wireless networks

Outlook

- Evaluate which applications benefit from roles (eg routing)
- Improve role detection (gossiping) with broadcast algorithms