Compactness-Preserving Mapping on Trees

Jan Baumbach, Jiong Guo, Rashid Ibragimov

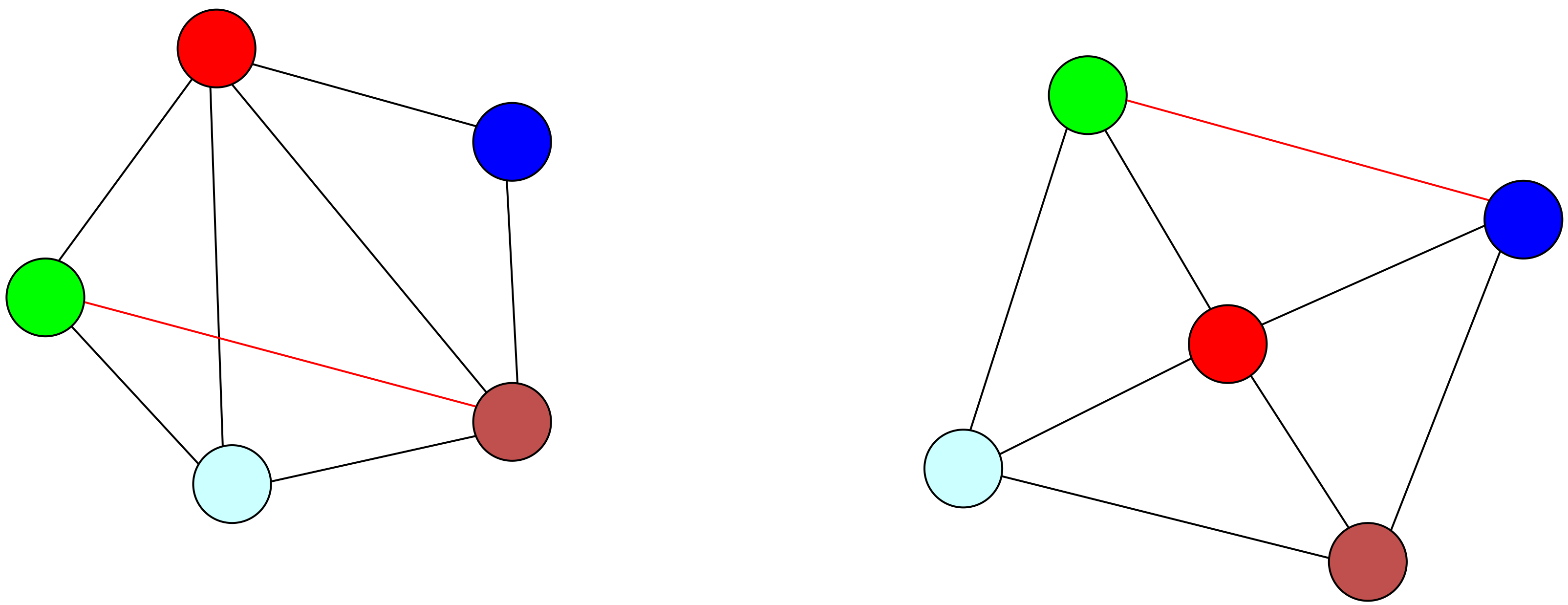
UNIVERSITÄT DES SAARLANDES

max planck institut informatik

M2CI CLUSTER OF EXCELLENCE

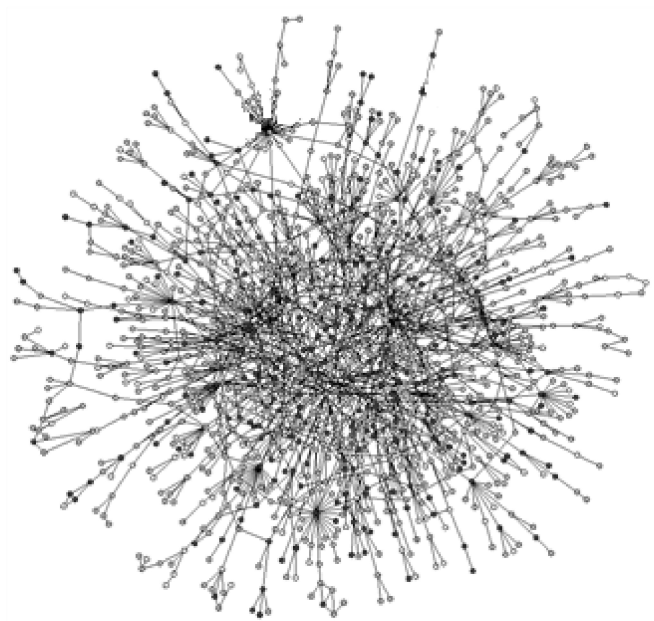
Network Alignment

Comparison of networks/graphs



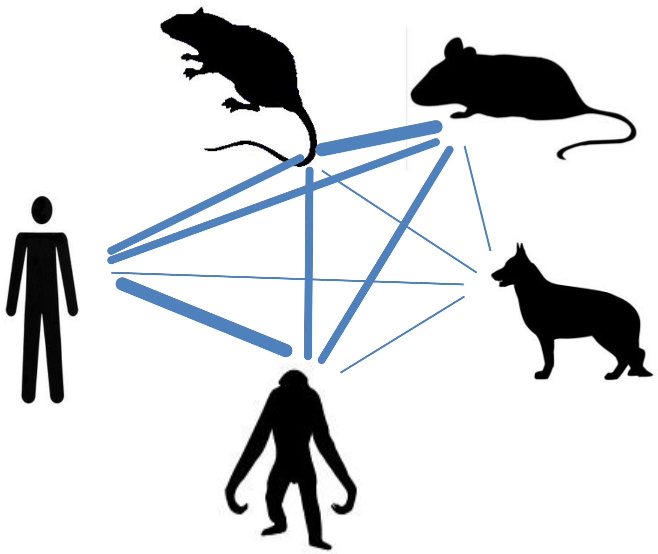
Motivation

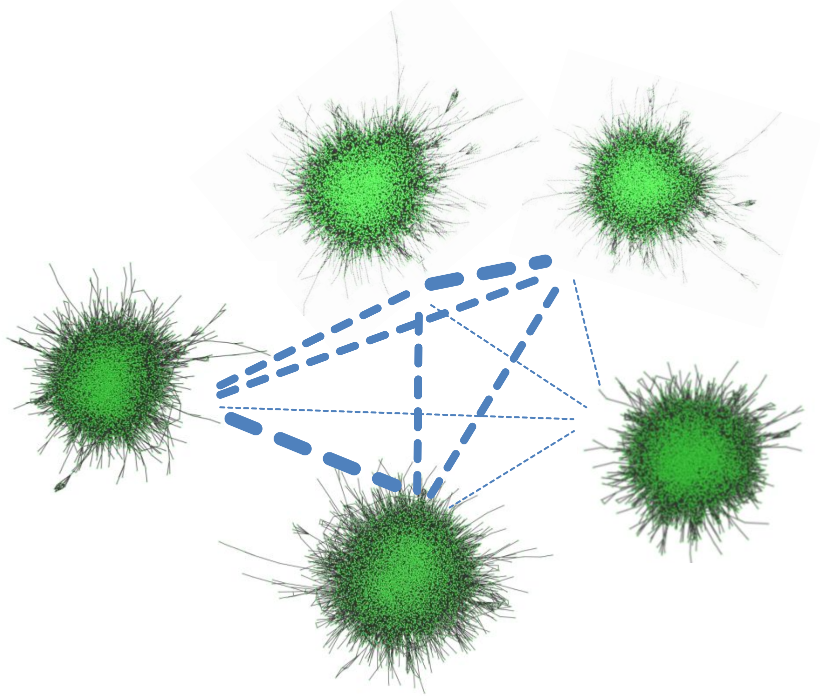
- Protein-Protein Interaction networks
  - nodes represent proteins
    - *0.1-10K nodes*
  - edges are interactions between proteins
    - *1-100K edges*
  - noisy and incomplete



Motivation

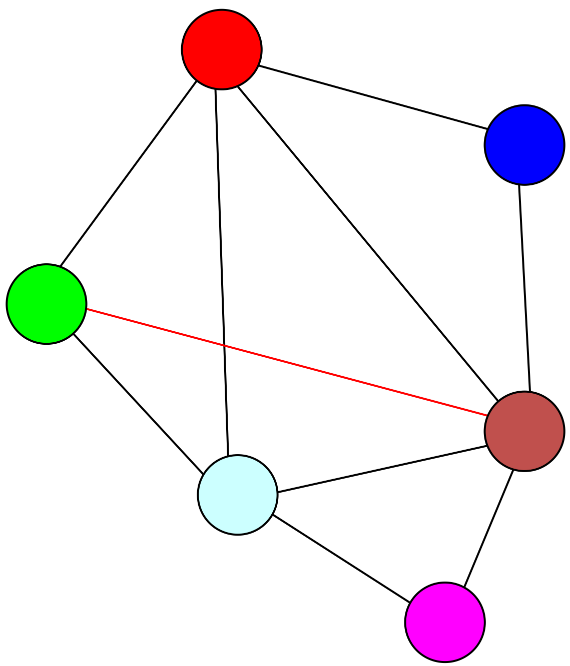
- find corresponding proteins between two networks
  - evolutionary conservation of many interactions
  - for every node, images of the neighboring nodes should be “close”

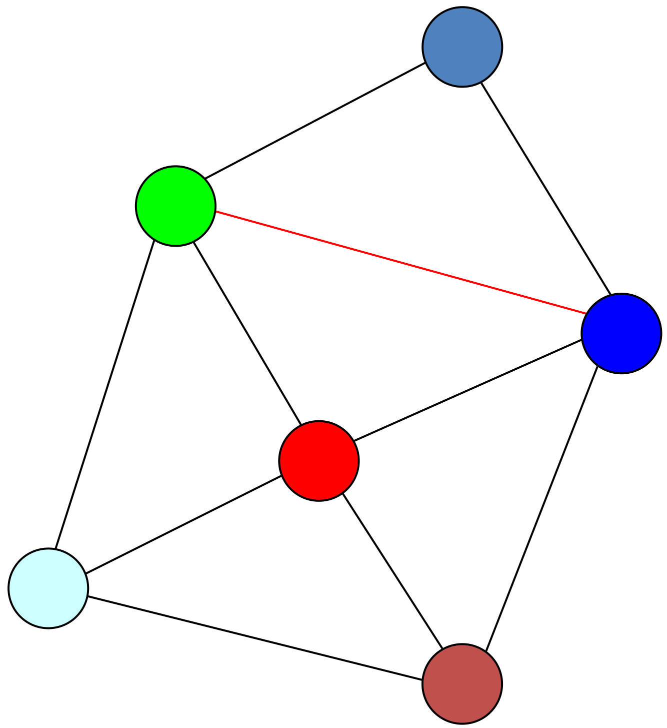




Network Alignment

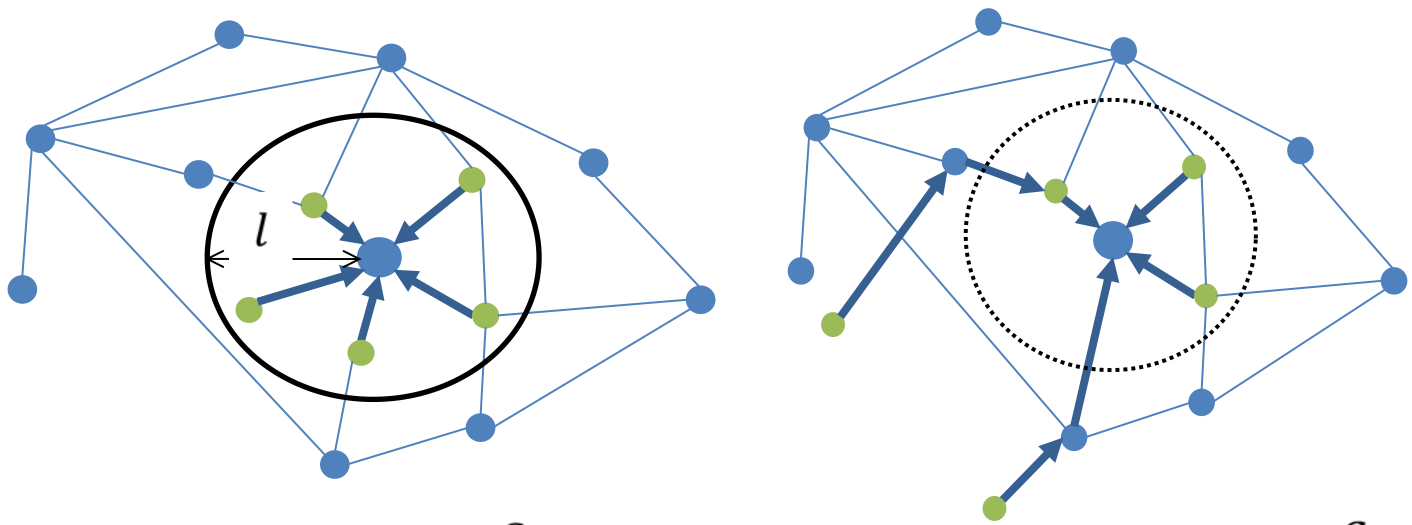
- Graph isomorphism $\rightarrow$ *perfect match*
- Graph edit distance





Compactness-Preserving Mapping (CPM)

- Neighbors of a node v – within distance l
- Closeness:
  - L_v is the total distance to neighbors of v
  - L'_v is the total distance to the corresponding nodes
  - the difference $L'_v - L_v$ should be small ($\le d$)



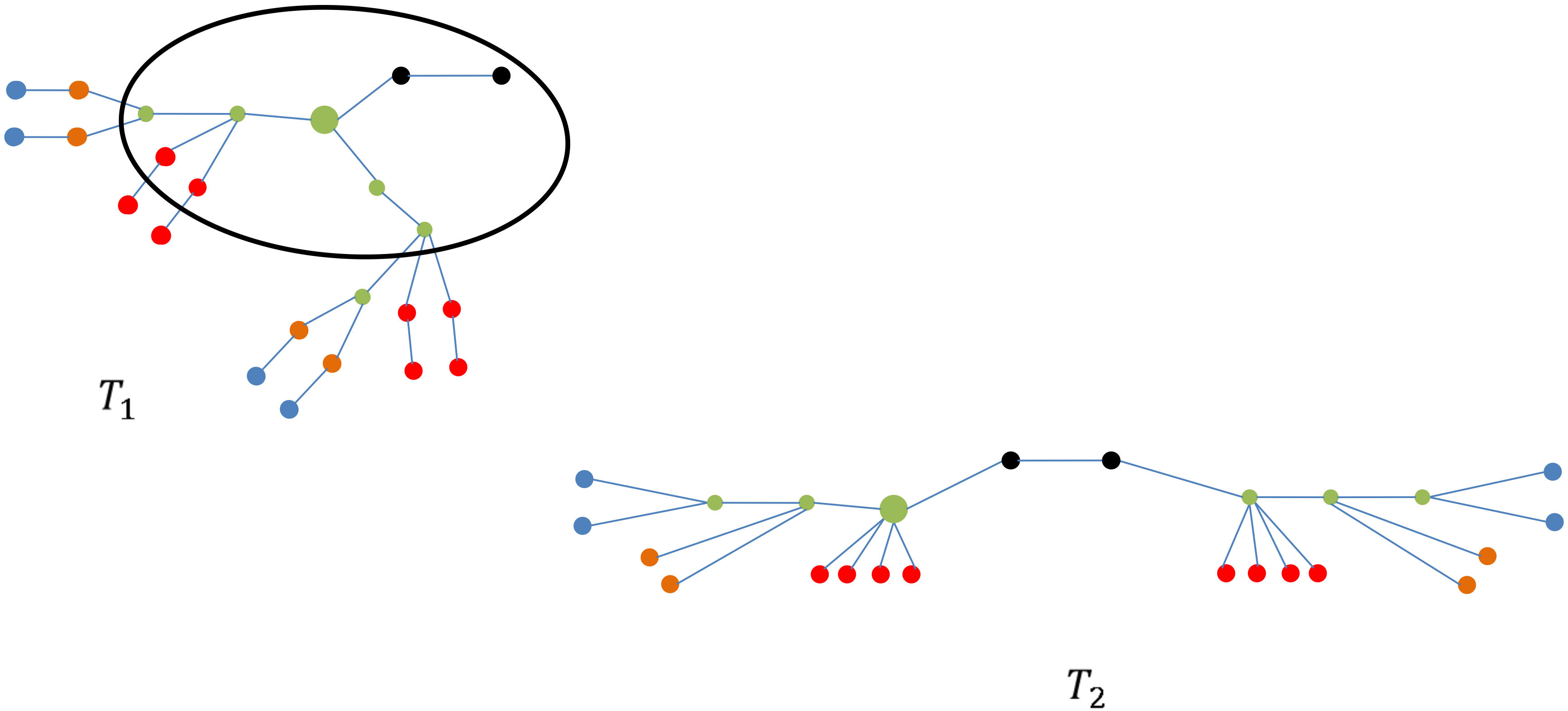
G_1 G_2

$L_v = 5$

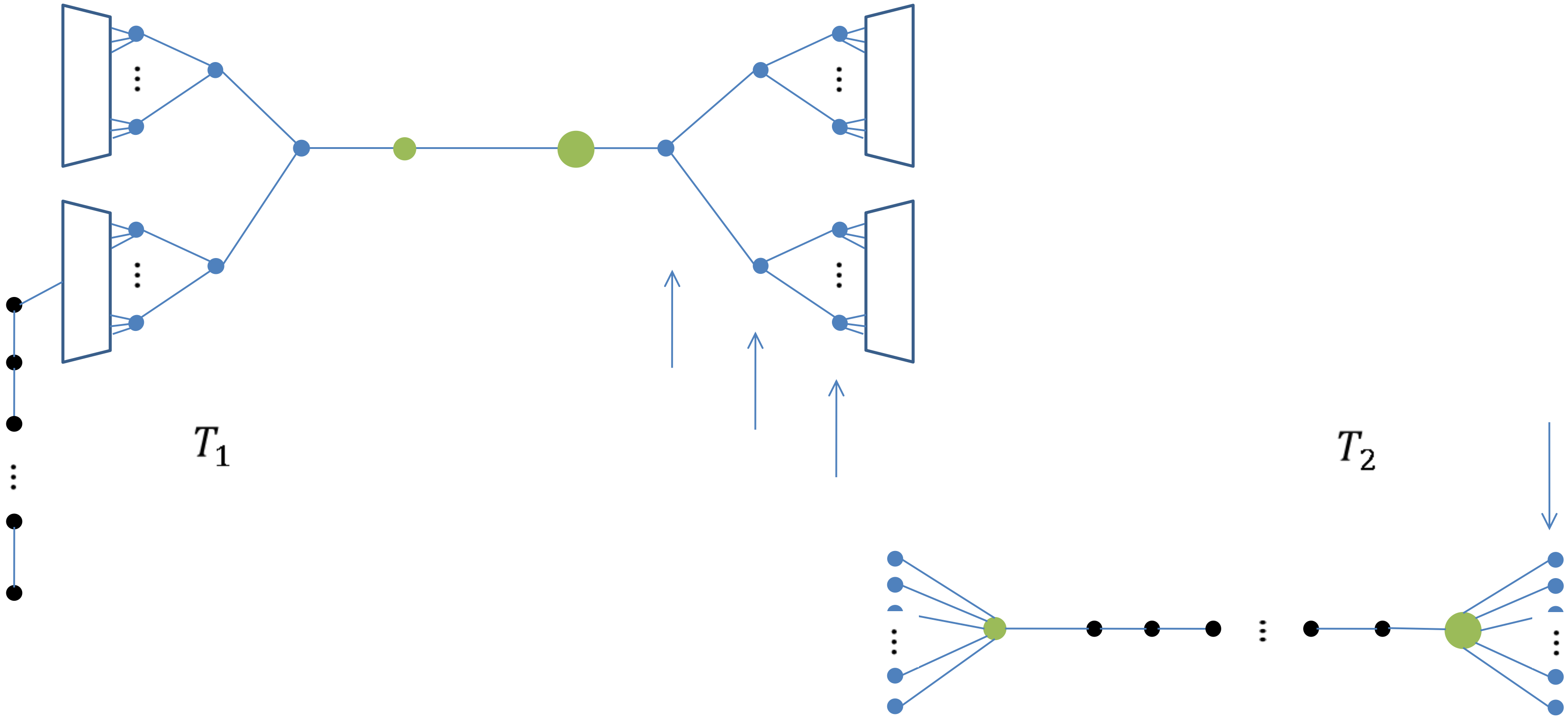
$L'_v = 7$

$L'_v - L_v = 2$

Example: CPM on Trees, $l = 2, d = 2$



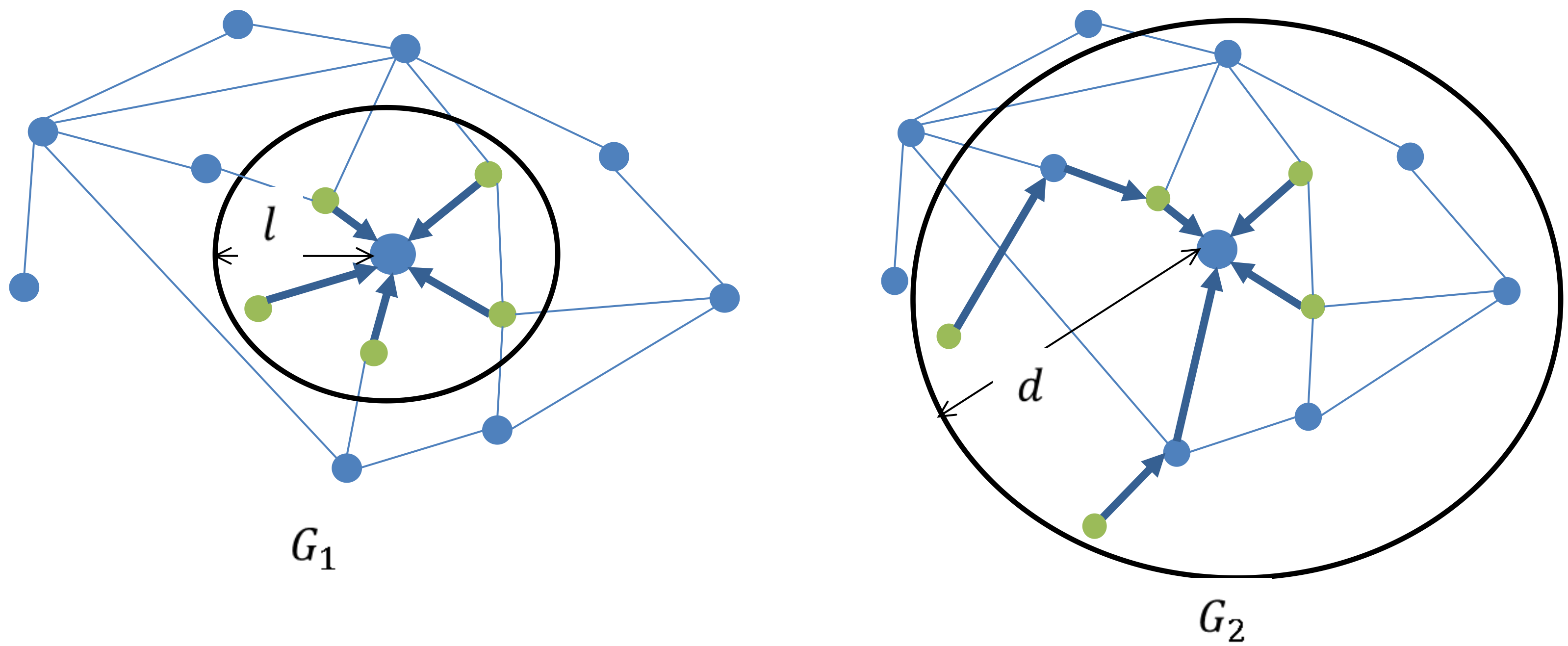
Example: CPM on Trees, $l = 3, d \geq 0$



- with larger l there are more neighbors to take into account
- some neighbors can be mapped closer, such that others very far away (greater than d)

Related problem: NPM

- Neighborhood-Preserving Mapping
  - neighbors of a node – within distance
  - images of the neighbors – within distance from

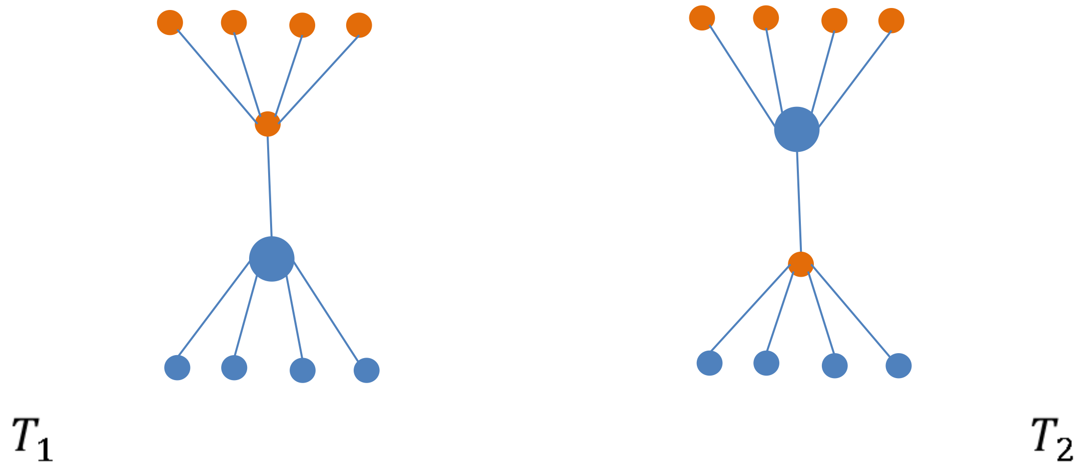


Result: CPM on Trees

l \ d	0	1	2	>2
1	P	P	NP	NP
2	P	P	NP	NP
3	NP	NP	NP	NP
>3	NP	NP	NP	NP

CPM on Trees: $l \leq 2, d = 0$

- Case *l=1, d=0*
  - Tree Isomorphism
- Case *l=2, d=0*
  - Tree Isomorphism – if diameter at least 3



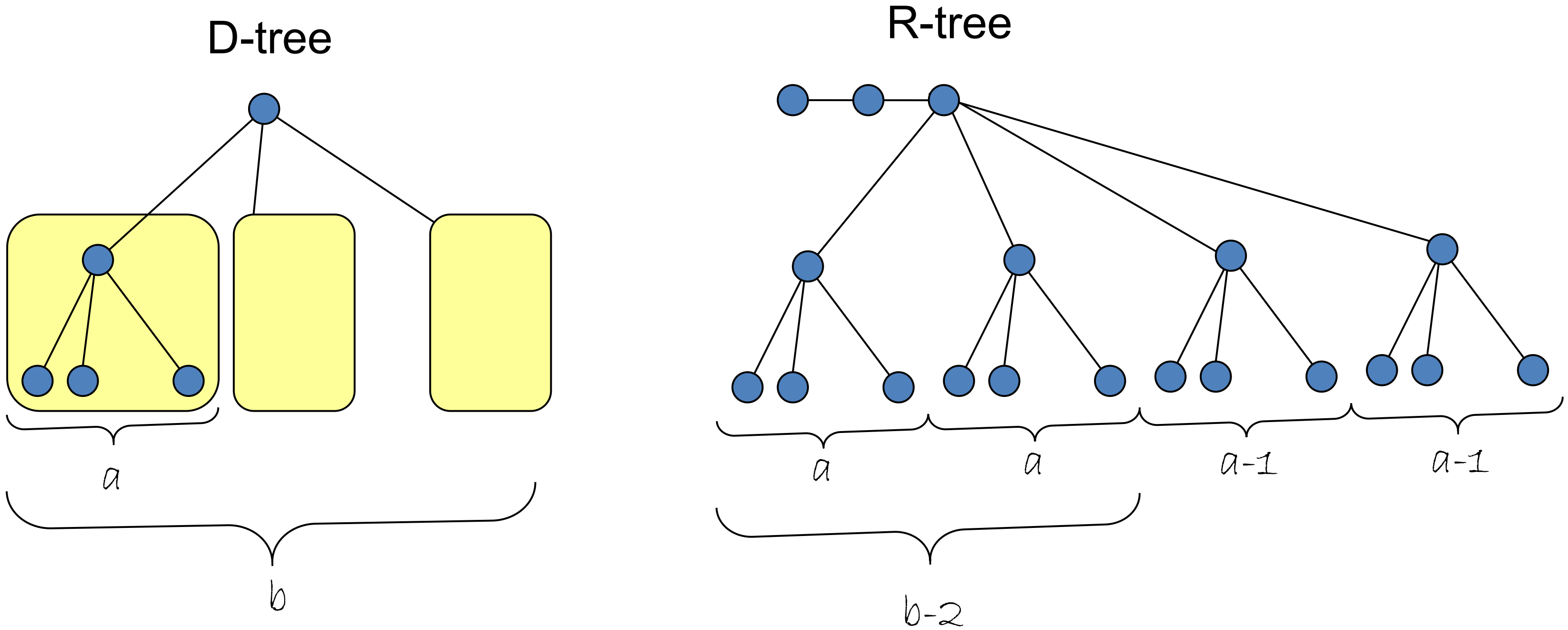
CPM on Trees: NP-Complete Cases

- Cases:
  - $l = 1, d = 2$
  - $l = 2, d = 2$
  - $l = 3, d = 0$
- All reductions from 3-PARTITION
  - case dependent subtree gadgets

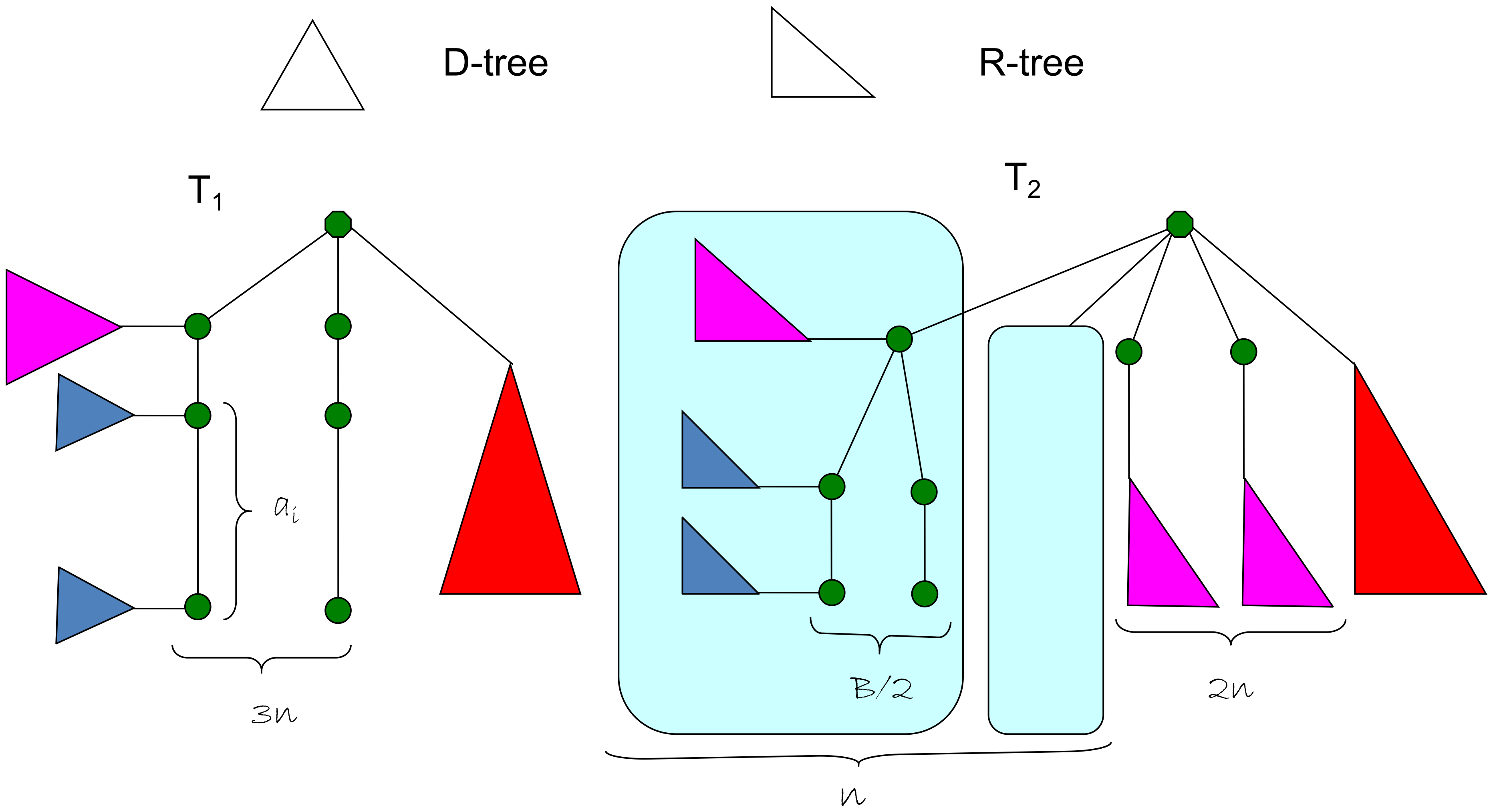
3-Partition

- 3-PARTITION
  - Given:
    - a set of integers $A := \{a_1, \ldots, a_{3n}\}$ and
    - B with $a_1 + \cdots + a_{3n} = nB$.
  - Question:
    - Is there a 3-partition $A_1, \ldots, A_n$ of A
      - i.e. $\bigcup A_j = A, |A_j| = 3, \sum_{a \in A_j} a = B, j := 1, \ldots, n$?
- For example,
  - $A := \{1, 2, 2, 3, 3, 4, 4, 6, 8\}, \; B := 11$
  - $A_1 = \{3, 4, 4\}, \; A_2 = \{2, 3, 6\}, \; A_3 = \{1, 2, 8\}$

NP-Case: $l=1,\ d=2$



NP-Case: $l=1, d=2$



CPM on Trees $l = 1, d = 1$

- 2-phase polynomial-time algorithm
- Phase 1: labeling
  - compute configurations
    - Can a node potentially be mapped to node?
    - distinguish several cases
    - reduce to a variant of 2-Path Packing in bipartite graphs
- Phase 2: resolving labels
  - bottom-up

Conclusion

- Optimization versions:
  - minimize max error over every node
  - minimize total error over all nodes
- Network Alignment
  - CPM-based models
    - heuristics
    - evaluation

Thank you for your attention!