A Max Flow based approach for discovering network building blocks

Md. Tanzirul Azim

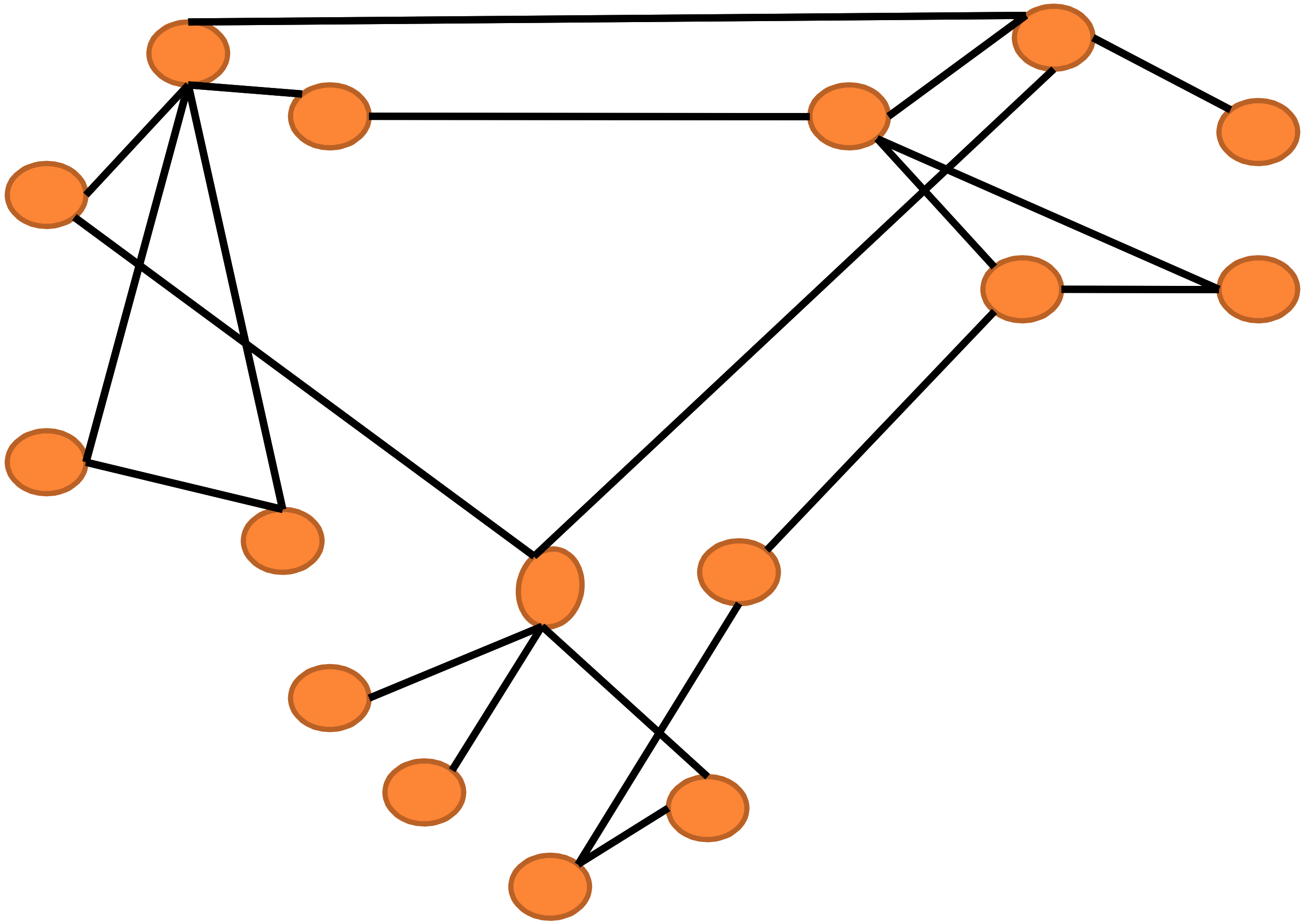
Kishwar Ahmed

Bangladesh University of Engineering and Technology

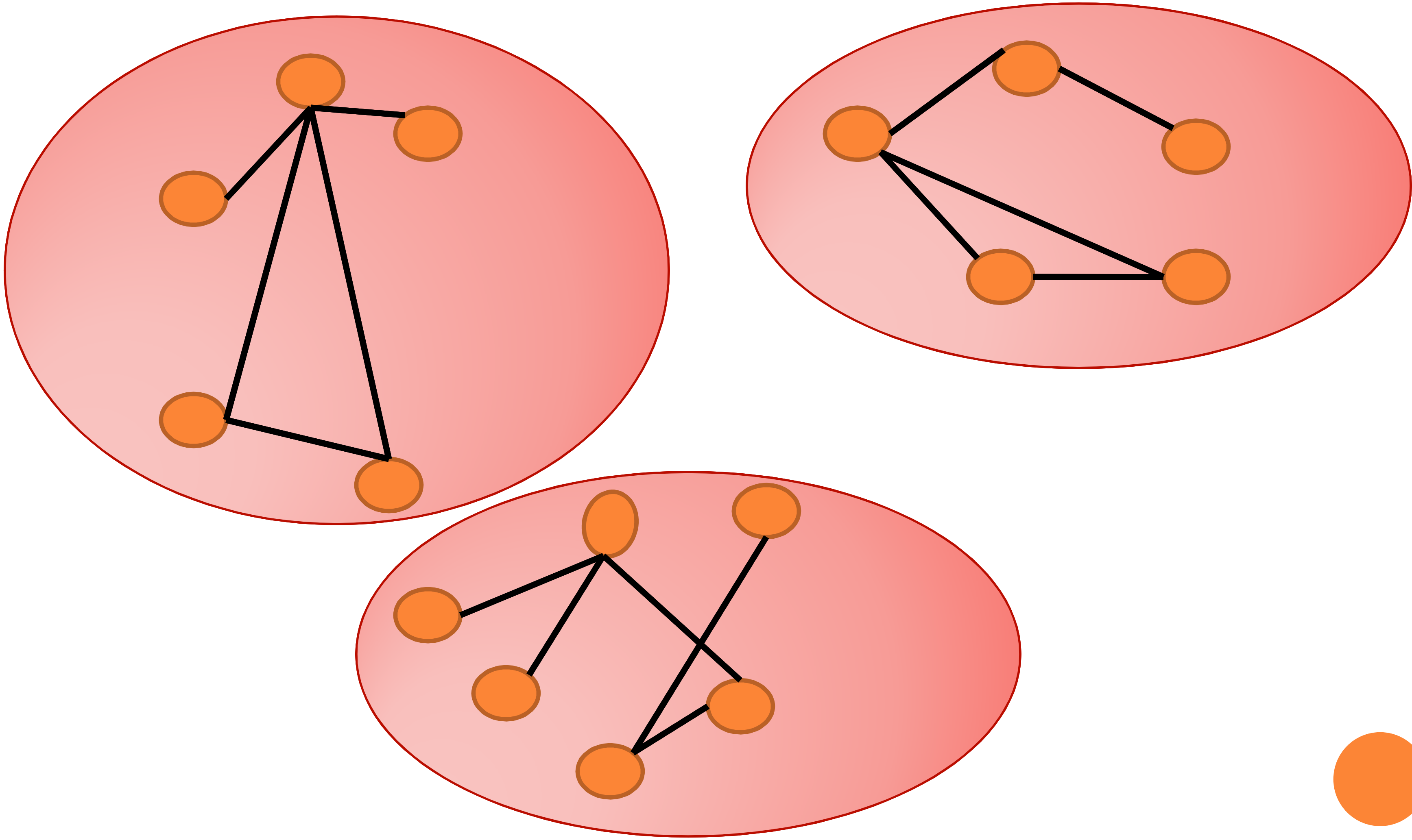
Network building blocks

- Collection of small subgraphs that creates a disjoint partitioning of nodes in the Graph.

A SAMPLE NETWORK



After finding network blocks for node size 5



Current approaches

- Subgraph finding techniques (SCNM)[1]
  - To find out network motifs
- An edgebetweenness based approach (GDNM)[2]
  - To find out network modules

1. An efficient algorithm for detecting frequent subgraphs in biological networks

Mehmet Koyutürk∗, Ananth Grama and Wojciech Szpankowski

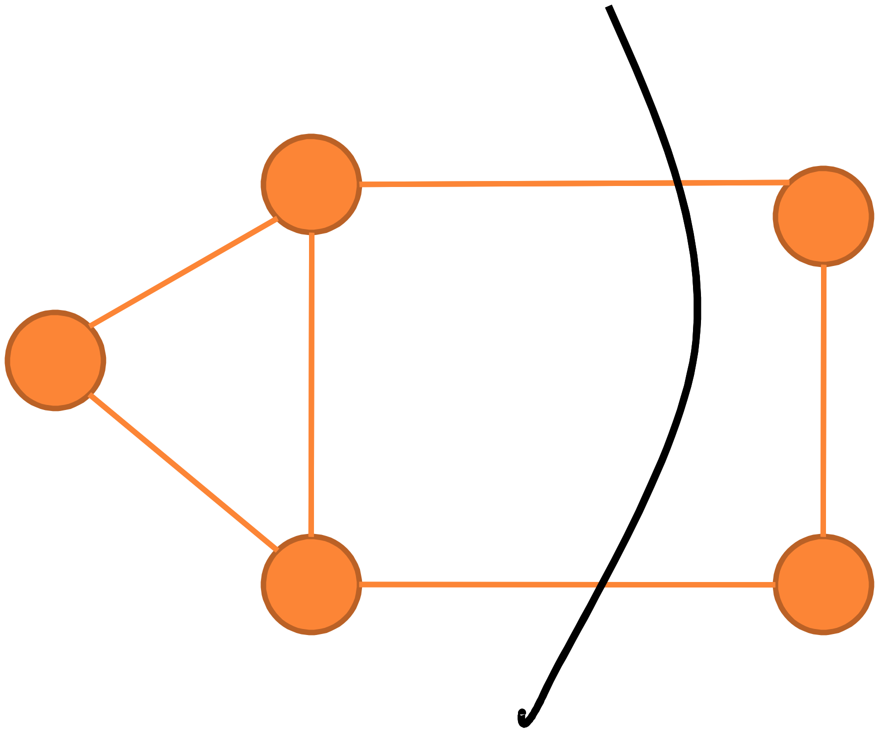
2. A decomposition approach for discovering network building blocks

Qiaofeng Yang, Stefano Lonardi

TERMINOLOGY

- Cut
  - A cut is a partition of the vertices of a graph into two sets.
    - The size of a cut is the total number of edges crossing the cut
- Minimum cut
  - A cut is minimum if the size of the cut is no larger than the size of any other cut.

Minimum cut

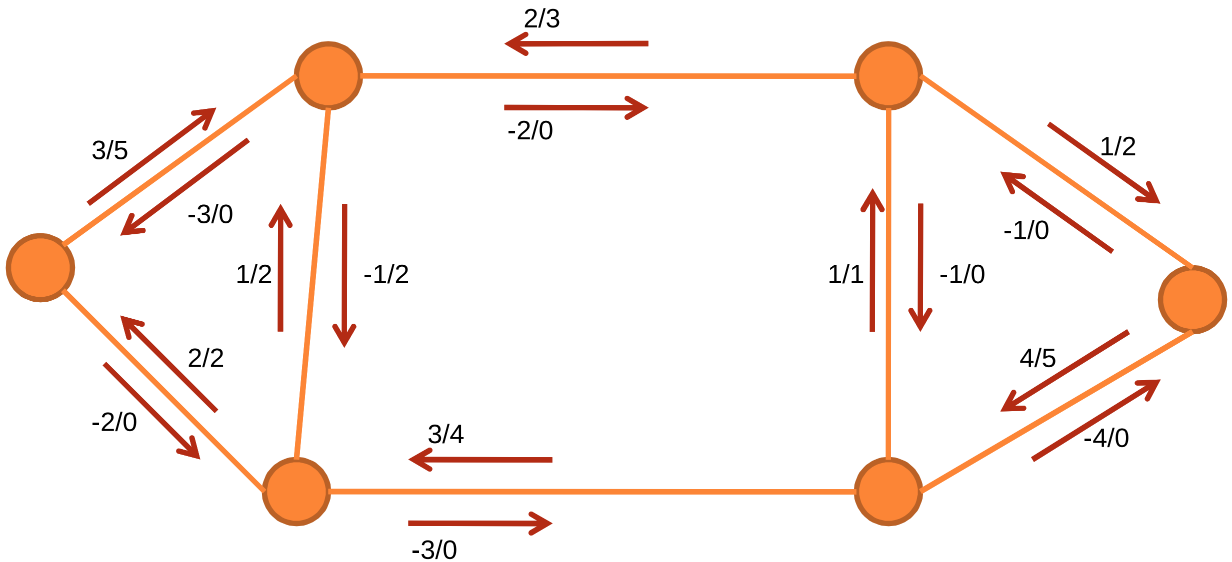


- This is a minimum cut with size 2.

FLOW NETWORK

- A flow network is a directed graph where each edge has a capacity and each edge receives a flow.
- A flow must satisfy the restriction that the amount of flow into a node equals the amount of flow out of it, except when it is a source, which has more outgoing flow, or sink which has more incoming flow.
- Max flow between two nodes is the maximum amount of flow in a graph.

A flow network



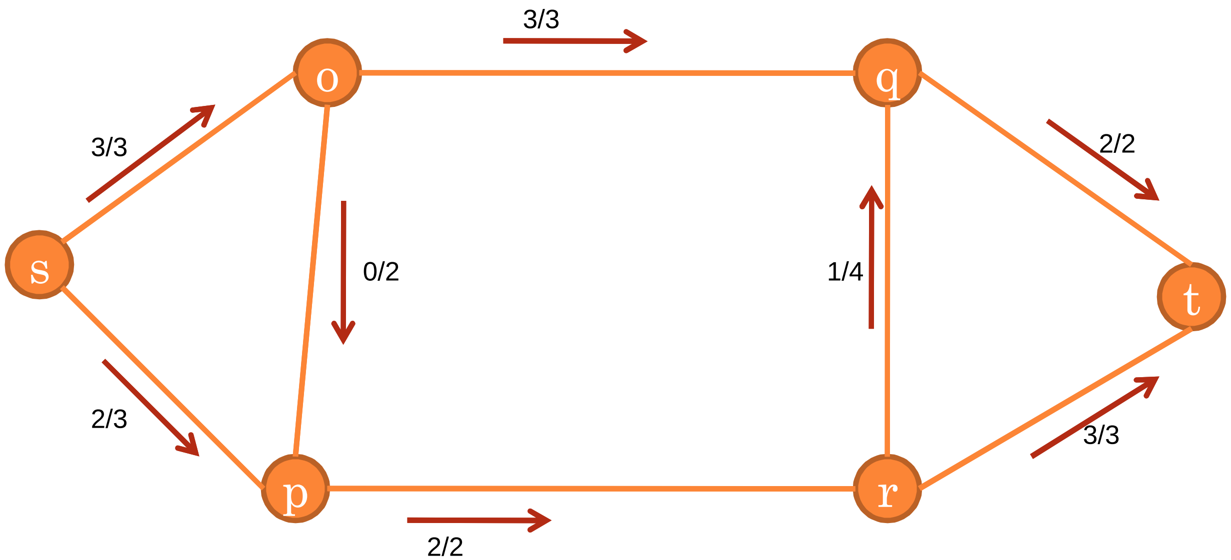
- Format:
  - f(u,v)/c(u,v); where
    - f(u,v)= amount of flow through an edge
    - c(u,v)= capacity of an edge

MAX-FLOW MIN-CUT THEOREM

- The max-flow min-cut theorem is a statement in optimization theory about maximum flows in flow networks. It derives from **Menger's theorem**. It states that

  the maximum amount of flow is equal to the capacity of a minimum cut.

Max-Flow Min-Cut Theorem example



Cut	Capacity
$S = \{s,p\}, T = \{o,q,r,t\}$	$c(s,o) + c(p,r) = 3 + 2 = 5$
$S = \{s,o,p\}, T = \{q,r,t\}$	$c(o,q) + c(p,r) = 3 + 2 = 5$
$S = \{s,o,p,q,r\}, T = \{t\}$	$c(q,t) + c(r,t) = 2 + 3 = 5$

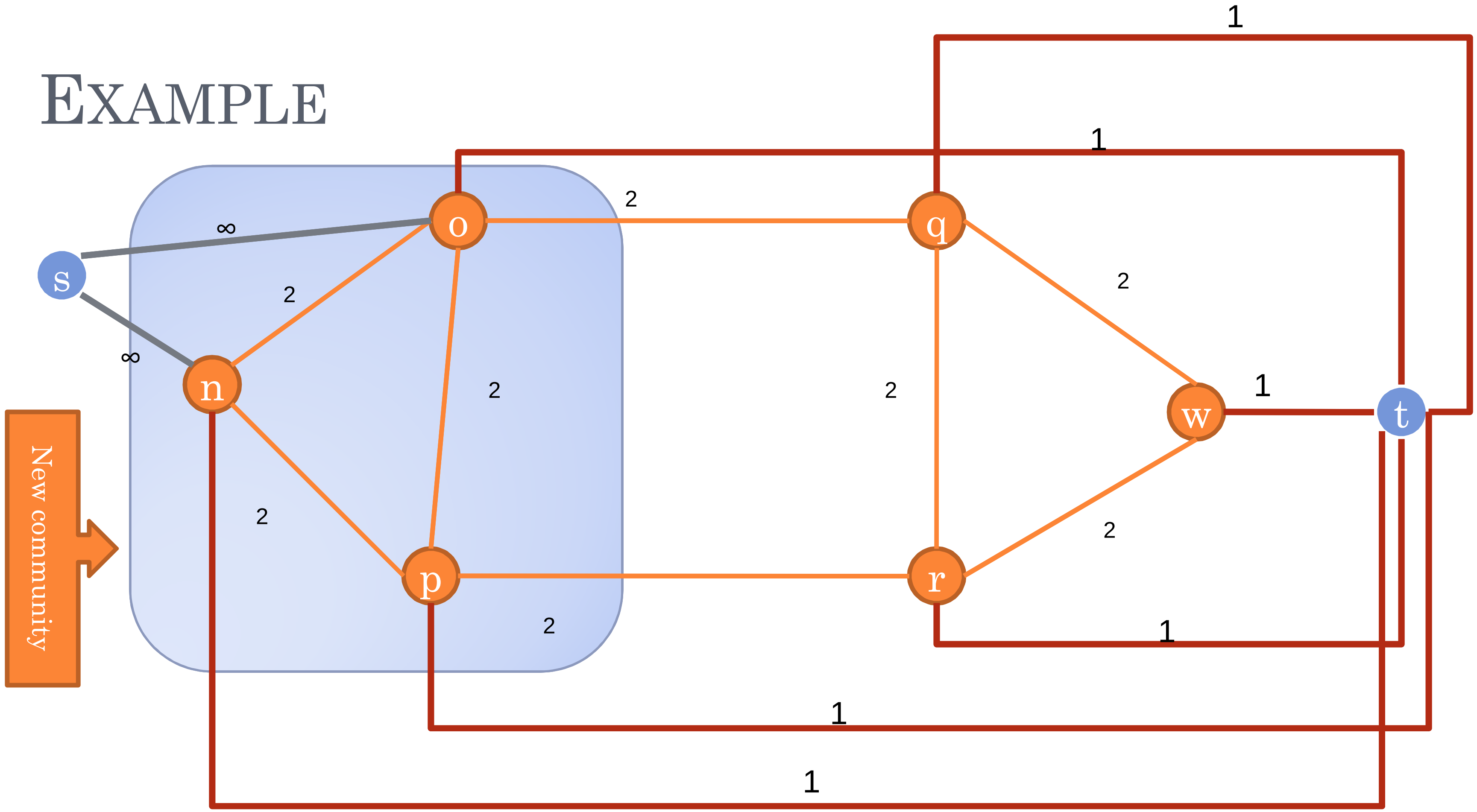
A MAX-FLOW BASED APPROACH FOR DETECTING NETWORK BUILDING BLOCKS

- Let the required nodes in a community is μ.
- Construct a network graph G(V,E). The vertex set V is a union of Seed vertices set S and P,Q,…
- Add to G a virtual source s and a virtual sink t. and add to G a virtual edge (s,x :x∈S) and a virtual edge (x,t: x∈V). Let G_1'(V',E') be the resulting network.
- The capacity c(e) for each edge e∈E' is given as follows
  - C(e)=k for c∈E is equal to the number of seed vertices.
  - C(e)=∞ for e=(s,x))
  - C(e)=1 for e=(x,t)

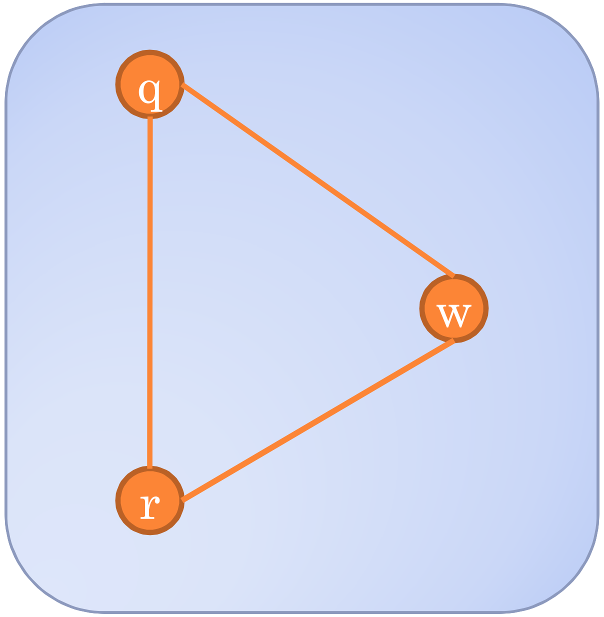
A MAX-FLOW BASED APPROACH FOR DETECTING NETWORK BUILDING BLOCKS

- For i=1 to l(some number l)
  - Compute the min-cut of the network G_i' by finding a max-flow for G_i'. Let E(X, V \X) be the set of edges in the min-cut, where X ⊆ V' is a vertex set containing s.
  - If i < l, then find a vertex u of maximum degree in X \ (S ∪ {s}), i.e. a vertex u such that d(u) ≥ d(u') for every vertex u' ∈ X \ (S ∪ {s}), where d(x) denotes the degree of vertex x ∈ V' in G_i', that is, d(x) is the sum of in-degree and out-degree of x. Regard u as a new seed, set S := S ∪{u}, and construct a new network G'_{i+1} for the new set S as in Steps 1 and 2. Increment i by one.
  - Otherwise, i.e. i = l , output X \ {s} as an approximate community.
  - Remove the community from the graph. Apply the algorithm for the resulting graph until there is no community having vertices > μ

Example

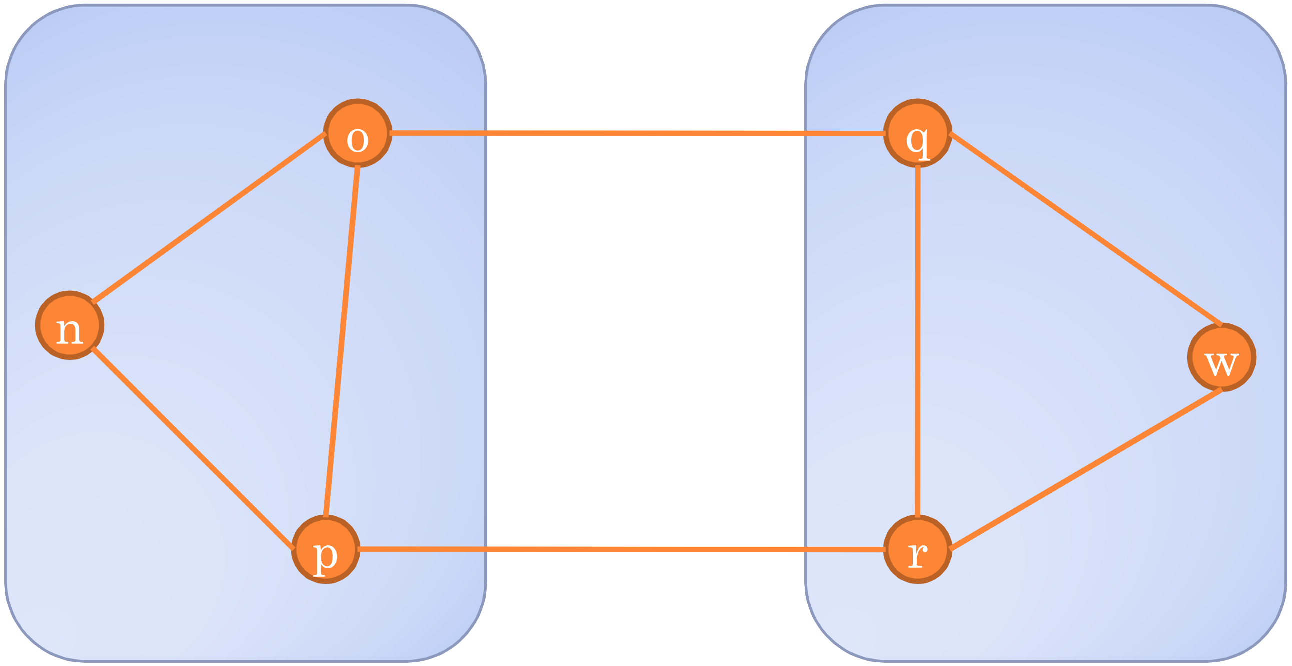


EXAMPLE



- New graph contains 3 nodes equal to μ.

OUTPUT



- Two communities separated