



西安电子科技大学  
XIDIAN UNIVERSITY

# 复杂网络与群体智能

吴建设

Email: [jshwu@mail.xidian.edu.cn](mailto:jshwu@mail.xidian.edu.cn)

西安电子科技大学人工智能学院



- 最小生成树
- 基于最小生成树的社区检测



## 1. 什么是最小生成树 (minimum spanning tree) ?

所谓最小生成树，就是在一个具有 $N$ 个顶点的加权连通图 $G$ 中，如果存在某个树形子图 $T$ ，其包含了图 $G$ 中的所有顶点和一部分边，且不形成回路，并且子图 $T$ 的各边权值（距离）之和最小，则称 $T$ 为图 $G$ 的最小生成树。

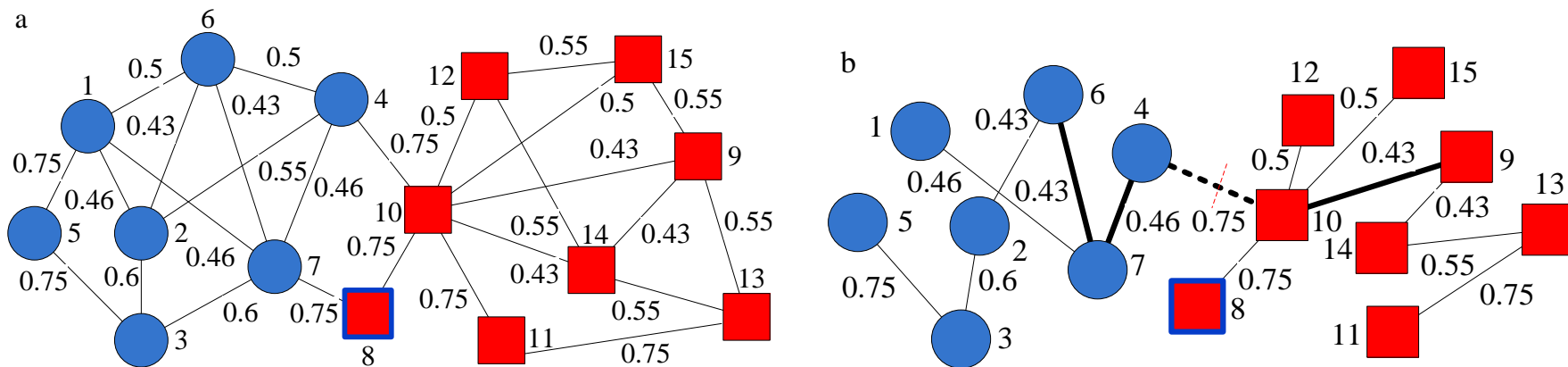


图1. 一个网络与它的最小生成树



## 2. 最小生成树的特性

- 无环，不能有回路
- 节点数和网络节点数相等
- 边数比节点数少一个
- 最小生成树可能有一个，也可能有多个

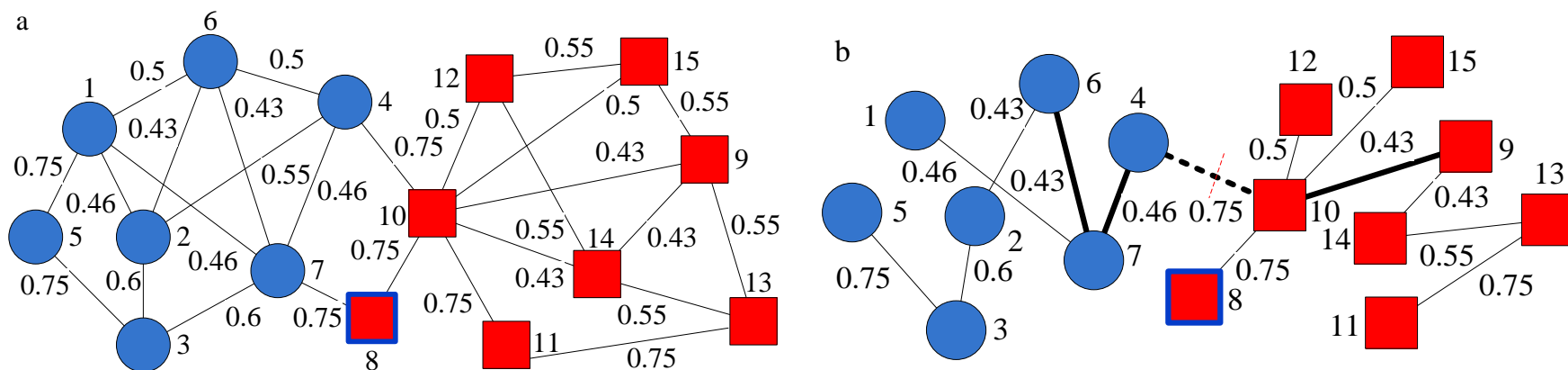


图1. 一个网络与它的最小生成树



## 3. 最小生成树算法

- 无环，不能有回路
- 节点数和网络节点数相等
- 边数比节点数少一个
- 最小生成树可能有一个，也可能有多个

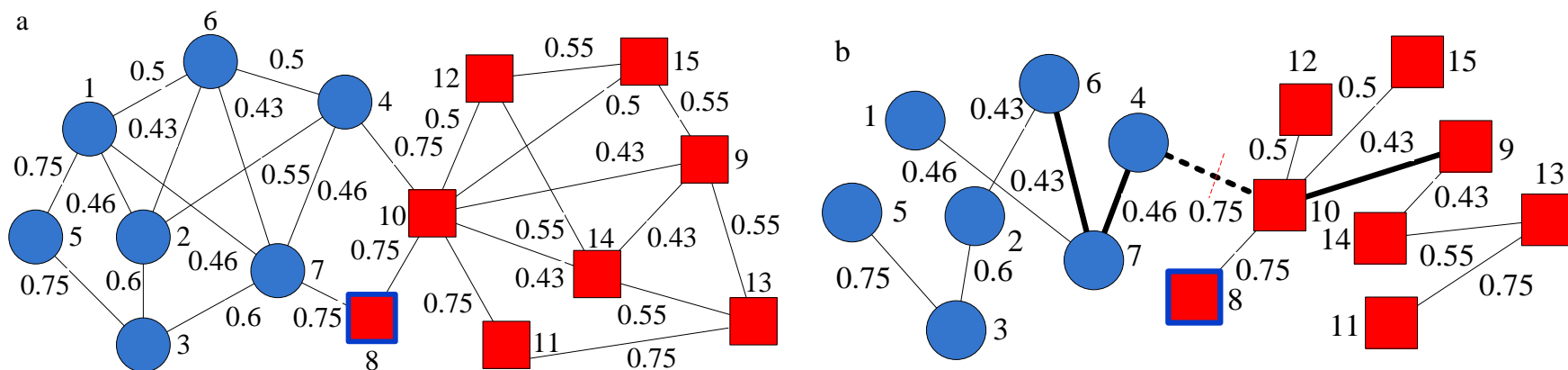


图1. 一个网络与它的最小生成树



## 3. 克鲁斯卡尔算法(Kruskal Algorithm)

第一步：在带权连通图中，将边的权值排序；

第二步：从权值最小的边开始，判断是否需要选择这条边。判断的依据是边的两个顶点是否已连通，如果连通则继续下一条；如果不连通，那么就选择使其连通。

连通时，要做如下判断：这两个点是否在已找到点的集合中出现过？①.如果两个点都没有出现过，那么将这两个点都加入已找到点的集合中；②.如果其中一个点在集合中出现过，那么将另一个没有出现过的点加入到集合中；③.如果这两个点都出现过，则不用加入到集合中。

第三步：循环第二步，直到图中所有的顶点都在子图中，即得到最小生成树。



## 4. Prim算法（加点法）

$G=(V,E)$ 为一图，其中 $V$ 为顶点的集合， $E$ 为边的集合

**第一步：**从某一顶点 $v_1$ 出发，选择与它相连的具有最小权值的边 $(v_1, v_i)$ ，将其顶点 $v_i$ 加入到生成树顶点集合 $U$ 中。 $U$ 用于存放 $G$ 的最小生成树中的顶点， $F$ 存放 $G$ 的最小生成树中的边。初始时刻：  
 $U=\{v_1\}$ ； $F=\{ \}$ 。

**第二步：**重复上述步骤

以后每一步从 $U$ 中选择一个顶点 $v_i$ 而另一个顶点 $v_j$ 属于 $V-U$ 的边中，选取具有最小权值的边 $(v_i, v_j)$ ，将顶点 $v_j$ 加入集合 $U$ 中，将边 $(v_i, v_j)$ 加入集合 $F$ 中；

如此不断重复，直到 $U=V$ 时，最小生成树构造完毕，这时集合 $F$ 中包含了最小生成树的所有边。



## 1. 社区网络最小生成树的特征

- 在最小生产树上，社区之间的边距离较长；社区内的边距离较短。利用这个特征就可以基于最小生成树进行社区检测。
- 叶子节点和其它节点的连接距离较长。

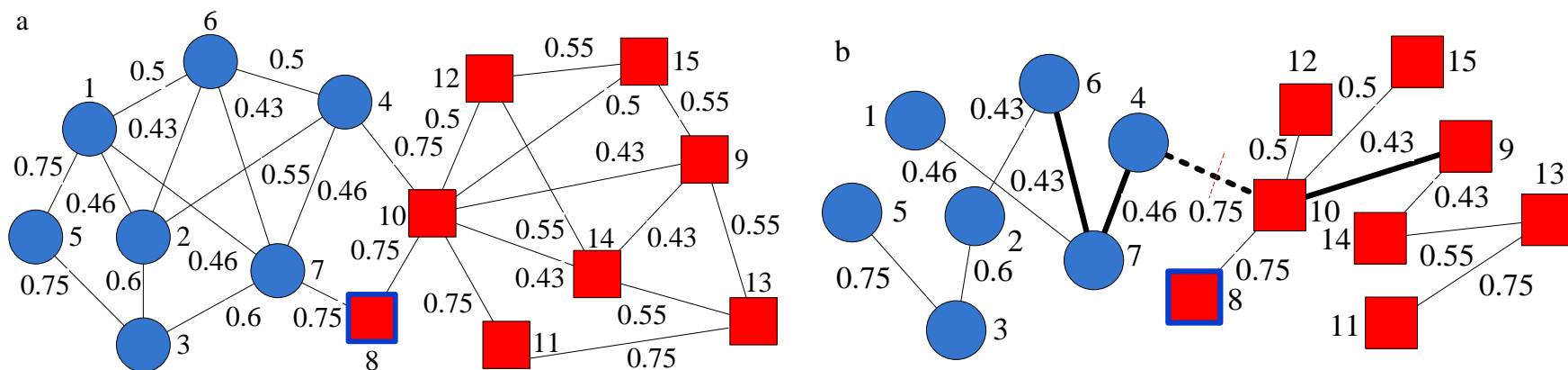


图1. 一个网络与它的最小生成树





## 2. 产生距离矩阵

- 距离矩阵和相似性矩阵的区别；
- 距离矩阵的产生：

以公式（1）计算节点 $v_i$ 和节点 $v_j$ 之间的距离

$$d_{ij} = \frac{1}{n_a 1/1 + n_b 1/2 + n_c 1/3}. \quad (1)$$

其中， $n_a, n_b, n_c$ 分别为一阶路径、二阶路径、三阶路径的个数；

- 去掉叶子节点产生距离矩阵；
- 用Kruskal Algorithm算法或者Prim算法得到最小生成树



## 3. 第2轮最小生成树（2nd-MST）

- 什么是第2轮最小生成树。

也是最小生成树，但不能有和第1轮最小生成树重复的边。

- 第2轮最小生成树的产生：

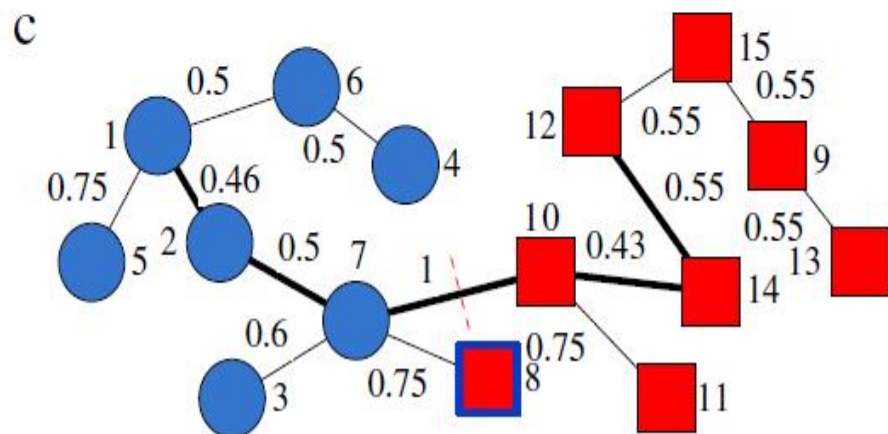
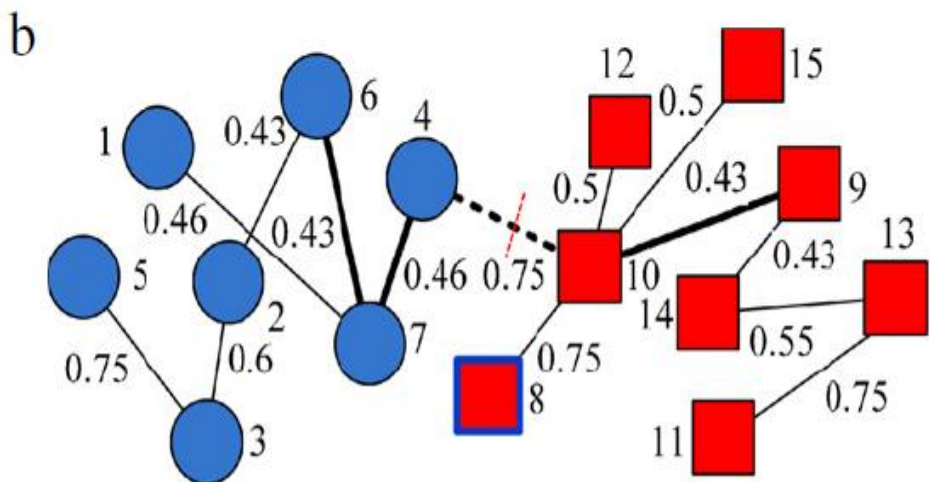
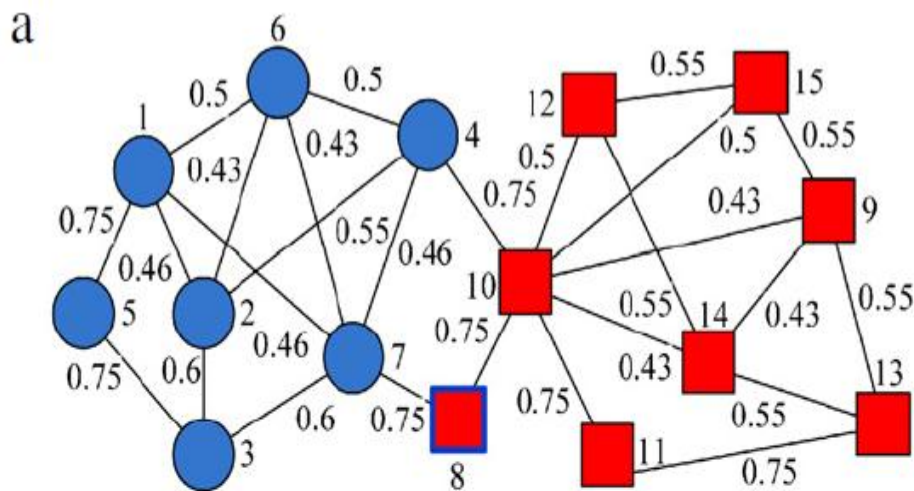
网络中删除1st-MST中存在的边，重新产生最小生成树；

等价于在距离矩阵中将1st-MST存在点边赋值无穷大。

- 2nd-MST和 1st-MST具有相同的特性。



## 3. 第2轮最小生成树 (2nd-MST)





## 4. 由两个最小生成树怎么分割网络

**第一步：** 去掉社区节点数少于某个阈值 $\varepsilon$ 的划分，比如3-5。

**第二步：** 去掉划分边的距离值小于路径距离平均值的划分。

$$S_{\varepsilon}^{T_1} = \left\{ \left( s_{11}^{T_1}, s_{12}^{T_1} \right), \left( s_{21}^{T_1}, s_{22}^{T_1} \right), \dots, \left( s_{L_1 1}^{T_1}, s_{L_1 2}^{T_1} \right) \right\},$$

$$S_{\varepsilon}^{T_2} = \left\{ \left( s_{11}^{T_2}, s_{12}^{T_2} \right), \left( s_{21}^{T_2}, s_{22}^{T_2} \right), \dots, \left( s_{M_1 1}^{T_2}, s_{M_1 2}^{T_2} \right) \right\},$$

$$d_{avg}^{T_1} = \frac{1}{L_1} \sum_{e_l \in T_1} e_l,$$

$$d_{avg}^{T_2} = \frac{1}{M_1} \sum_{e_l \in T_2} e_l.$$



## 4. 例子

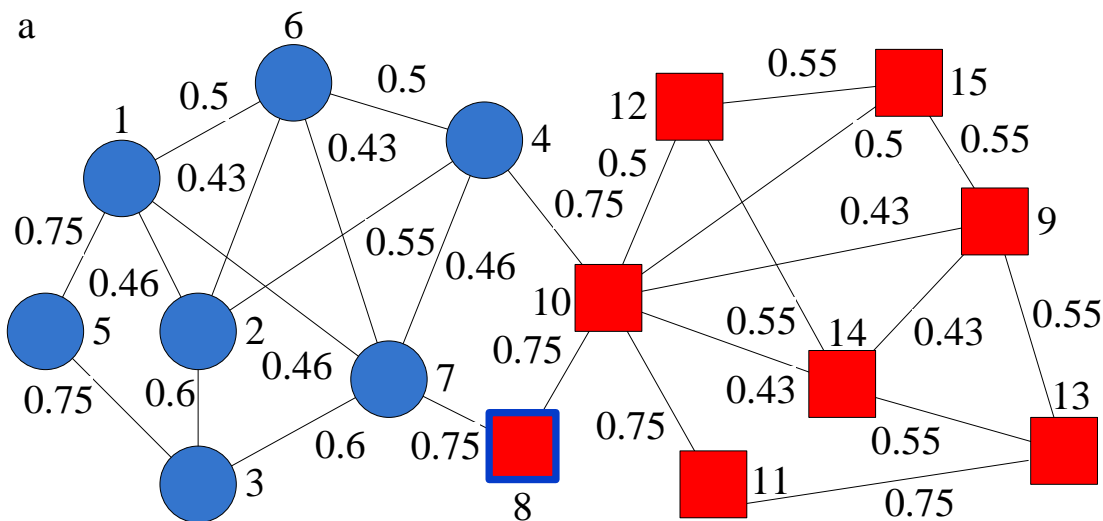
下表是去除社区节点数少于3的分别由1-MST 和2-MST得到的分割

| Separations in the 1 <sup>st</sup> -MST (A) |                   |                                      | Separations in the 2 <sup>nd</sup> -MST (B) |                                      |  |
|---------------------------------------------|-------------------|--------------------------------------|---------------------------------------------|--------------------------------------|--|
| distance<br>↓                               | e <sub>4-10</sub> | 1,2,3,4,5,6,7  8,9,10,11,12,13,14,15 | e <sub>7-10</sub>                           | 9,10,11,12,13,14,15  1,2,3,4,5,6,7,8 |  |
|                                             | e <sub>47</sub>   | 1,2,3,5,6,7  4,8,9,10,11,12,13,14,15 | e <sub>12-14</sub>                          | 9,12,13,15  1,2,3,4,5,6,7,8,10,11,14 |  |
|                                             | e <sub>67</sub>   | 2,3,5,6  1,4,7,8,9,10,11,12,13,14,15 | e <sub>27</sub>                             | 1,2,4,5,6  3,7,8,9,10,11,12,13,14,15 |  |
|                                             | e <sub>9-10</sub> | 9,11,13,14  1,2,3,4,5,6,7,8,10,12,15 | e <sub>12</sub>                             | 1,4,5,6  2,3,7,8,9,10,11,12,13,14,15 |  |
|                                             |                   |                                      | e <sub>10-14</sub>                          | 9,12,13,14,15  1,2,3,4,5,6,7,8,10,11 |  |

去除分割边的距离小于平均值的分割后，分别剩下一个分割，即：  
e<sub>4-10</sub>和e<sub>7-10</sub>。得到v<sub>8</sub>是重叠节点。



## 4. 例子



右图可见v8是重叠节点是合理的。



- (1) 用第2轮最小生成树起到的作用？
- (2) 可以用 第3轮最小生产树吗？



西安电子科技大学  
XIDIAN UNIVERSITY

谢谢