Complex Networks

Master of Science in Electrical Engineering

Erivelton Geraldo Nepomuceno

Department of Electrical Engineering
Federal University of São João del-Rei

April 9, 2019

1.4 Trees

- Usually a tree is defined as a connected graph containing no cycles. We then say that a tree is a connected acyclic graph.
- Forest: a graph whose connected components are all trees.
- Cayley tree: an infinite tree in which each node is connected to z neighbours.

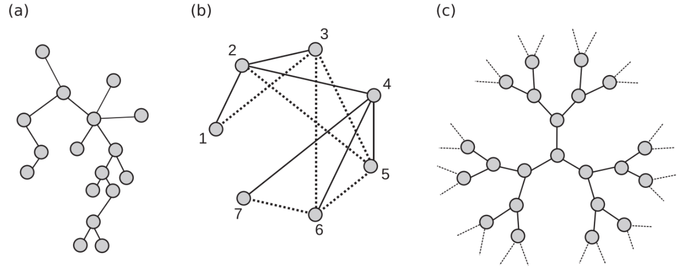


Figure 13: A tree with $N = 17$ nodes (a). A spanning tree (solid lines) of a graph with $N = 7$ nodes and 12 links (solid and dashed lines) (b).Three levels of a Cayley tree with $z = 3$ (c).

Example 3 (Trees in the real world)

- Any printed material and textbook that is divided into sections and subsections is organised as a tree.
- Large companies are organised as trees, with a president at the top, a vice-president for each division, and so on.
- Each drive is the root of an independent tree of files and directories (folders).
- In nature, plants and rivers have a tree-like structure.

Definition 13 (Trees)

A tree can be alternatively defined as: (A) a connected acyclic graph; (B) a connected graph with $K = N - 1$ links; (C) an acyclic graph with $K = N - 1$ links.

1.5 Graph Theory and the Bridges of Königsberg

- The theorem proposed by Leonhard Euler in 1736 as a solution to the Königsberg bridge problem: an example of graph theory abstraction.
- The problem to solve is whether or not it is possible to find an optimum route that traverses each of the bridges exactly once, and eventually returns to the starting point.

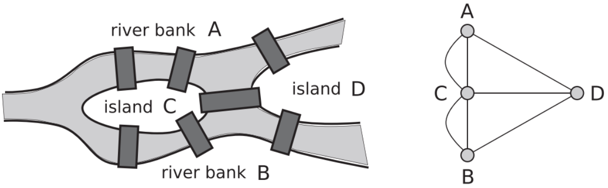


Figure 14: The city of Königsberg at the time of Leonhard Euler (left).The river is coloured light grey, while the seven bridges are dark grey. The associated multigraph, in which the nodes corresponds to river banks and islands, and the links represents bridges (right).

- A brute force approach to this problem consists in starting from a side, making an exhaustive list of possible routes, and then checking one by one all the routes.
  - In a complete graph $N = 4$, it means $N(N-1)^{N(N-1)/2} = 2916$.
- Euler came up with an elegant idea:
  - First, he introduced the idea of the graph.
  - Is it possible to find a trail containing all the graph edges?
- Explanation of the following theorem. [Video-04].

Theorem 1 (Euler theorem)

A connected graph is Eulerian iff each vertex has even degree. It has a Eulerian trail from vertex i to vertex j, $i \neq j$, iff i and j are the only vertices of odd degree.

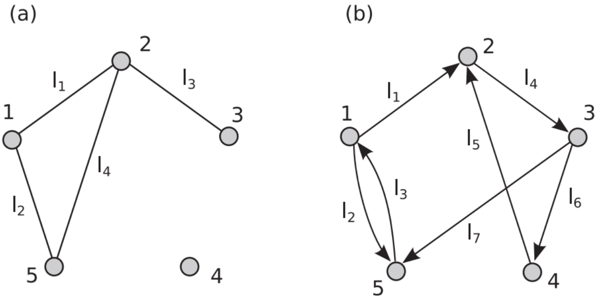
1.6 How to Represent a Graph

- Drawing a graph is a certainly a good way to represent it.
- However, when the number of nodes and links in the graph is large, the picture we get may be useless.
- An alternative representation of a graph can be obtained by using a matrix.

Definition 14 (Adjacency matrix)

The adjacency matrix A of a graph is a $N \times N$ square matrix whose entries a_{ij} are either ones or zeros according to the following rule:

$$a_{ij} = \begin{cases} 1 & \textit{iff } (i,j) \in \mathcal{L} \\ 0 & \textit{otherwise} \end{cases} \tag{1}$$



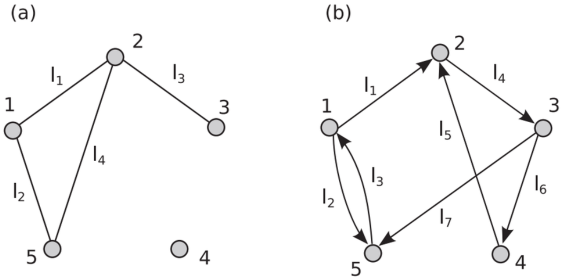
$$A_u = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \quad A_d = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 15: The first graph is undirected and has $K = 4$ links, while the second graph is directed and has $K = 7$ arcs.

- A_u is symmetric and contains $2K$ non-zero entries.
- The number of ones in row i, or equivalently in column i, is equal to the degree of vertex i.
- The adjacency matrix of a directed graph is in general not symmetric. This is the case of A_d. This matrix has K elements different from zero.
- The number of ones in row i is equal to the number of outgoing links k_i^{out}.
- The number of ones in column i is equal to the number of ingoing links k_i^{in}.
- A bipartite graph can be described by an $N \times V$ adjacency matrix A, such that entry $a_{i\alpha}$, with $i = 1, \ldots, N$ and $\alpha = 1, \ldots, V$, is equal to 1 if node i of the first set and node α of the second set are connected, while it is 0 otherwise.
- Example 1.13 - Recommendation systems: collaborative filtering: Potential seminar.

Definition 15 (Incidence matrix)

The incidence matrix B of an undirected graph is an $N \times K$ matrix whose entry b_{ik} is equal to 1 whenever the node i is incident with the link l_k, and is zero otherwise. If the graph is directed, the adopted convention is that the entry b_{ik} of B is equal to 1 if arc k points to node i, it is equal to -1 if the arc leaves node i, and is zero otherwise.



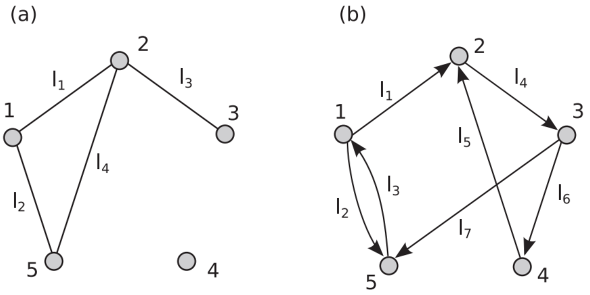
$$B_u = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad B_d = \begin{pmatrix} -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 16: The incidence matrices respectively associated with the undirected and directed graphs.

Definition 16 (Edge list)

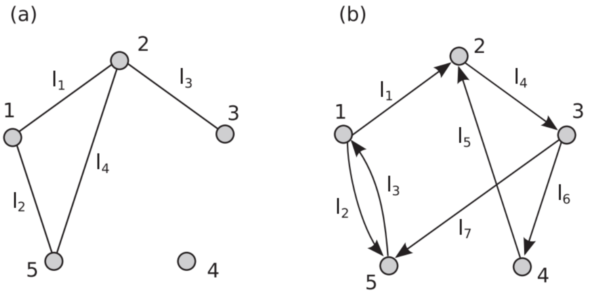
The edge list of a graph, also known as *ij*-form of the adjacency matrix of the graph, consists of two vectors i and j of integer numbers storing the positions, i.e. respectively the row and column indices of the ones of the adjacency matrix A. Each of the two vectors has M components, with $M = 2K$ for undirected graphs, and $M = K$ for directed graphs.

- The *ij*-form of the adjacency matrix is very flexible, because the order in which the links are stored is not relevant.
- This form is often used during the construction of a graph which is obtained by adding a new link at a time.



$$A_u = (\mathbf{i},\mathbf{j}),\ \mathbf{i} = \begin{pmatrix} 1\\1\\2\\2\\2\\3\\5\\5 \end{pmatrix},\ \mathbf{j} = \begin{pmatrix} 2\\5\\1\\3\\5\\2\\1\\2 \end{pmatrix}, \quad A_d = (\mathbf{i},\mathbf{j}),\ \mathbf{i} = \begin{pmatrix} 1\\1\\2\\3\\3\\4\\5 \end{pmatrix},\ \mathbf{j} = \begin{pmatrix} 2\\5\\3\\4\\5\\2\\1 \end{pmatrix}$$

Figure 17: *ij*-form of the adjacency matrices.



$$A_u = \left(\begin{array}{c|lll} \text{node} & \text{in-neighbours} & & \\ 1 & 2 & 5 & \\ 2 & 1 & 3 & 5 \\ 3 & 2 & & \\ 4 & & & \\ 5 & 1 & 2 & \end{array}\right), \quad A_d = \left(\begin{array}{c|ll} \text{node} & \text{out-neighbours} & \\ 1 & 2 & 5 \\ 2 & 3 & \\ 3 & 4 & 5 \\ 4 & 2 & \\ 5 & 1 & \end{array}\right)$$

Figure 18: The lists of neighbours corresponding to the two graphs.

Definition 17

The compressed row storage, instead, consists of an array **j** of size 2K (K for directed graphs) storing the ordered sequence of the neighbours of all nodes, and a second array, **r**, of size $N + 1$. The k_1 neighbours of node 1 are stored in **j** at positions $1, 2, \ldots, k_1$, the k_2 neighbours of node 2 are stored in **j** at positions $k_1 + 1, k_1 + 2, \ldots, k_1 + k_2$, and so on.

$$A_u = \left(\begin{array}{c|lll} \text{node} & \text{in-neighbours} & & \\ 1 & 2 & 5 & \\ 2 & 1 & 3 & 5 \\ 3 & 2 & & \\ 4 & & & \\ 5 & 1 & 2 & \end{array} \right), \quad A_d = \left(\begin{array}{c|ll} \text{node} & \text{out-neighbours} & \\ 1 & 2 & 5 \\ 2 & 3 & \\ 3 & 4 & 5 \\ 4 & 2 & \\ 5 & 1 & \end{array} \right)$$

Figure 19: Compressed storage. The vectors $\mathbf{r}$ and $\mathbf{j}$ corresponding to matrix A_u are $\mathbf{r} = (1, 3, 6, 7, 7, 9)$ and $\mathbf{j} = (2, 5, 1, 3, 5, 2, 1, 2)$. The vectors $\mathbf{r}$ and $\mathbf{j}$ corresponding to matrix A_d in the compressed row storage are $\mathbf{r} = (1, 3, 4, 6, 7, 8)$ and $\mathbf{j} = (2, 5, 3, 4, 5, 2, 1)$.

Algorithm 11 From CRS to *ij*-format

Input: $\mathbf{r}$

Output: $\mathbf{i}$

```
1: for ℓ = 0 to N − 1 do
2:     for k = r[ℓ] to r[ℓ + 1] − 1 do
3:         i[k] ← ℓ
4:     end for
5: end for
```

Figure 20: Pseudo-code CRS to *ij*-format.

Algorithm 12 From *ij*-to CRS format

Input: $\mathbf{i}, \mathbf{j}$

Output: $\vec{r}, \mathbf{j}_{\text{CRS}}$

$\vec{r} \leftarrow \vec{0}, \vec{p} \leftarrow \vec{0}$ {Set $\vec{r} \in \mathbb{N}^{N+1}$ and $\vec{p} \in \mathbb{N}^{N}$ to zero}
Count the number of elements for each row
for $k = 0$ **to** $M - 1$ **do**
  $r[i[k] + 1] \leftarrow r[i[k] + 1] + 1$
end for
Compute vector $\mathbf{r}$
$r[0] \leftarrow 0$
for $l = 0$ **to** $N - 1$ **do**
  $r[l + 1] \leftarrow r[l] + r[l + 1]$
end for
Compute vector $\mathbf{j}_{\text{CRS}}$
for $k = 0$ to $M - 1$ **do**
  $\ell \leftarrow i[k]$
  $j_{\text{CRS}}[r[\ell] + p[\ell]] \leftarrow j[k]$
  $p[\ell] \leftarrow p[\ell] + 1$
end for

Figure 21: Pseudo-code *ij*-format to CRS.

Exercises - Classrom

1. Change A_u and A_d from slide 7 to *ij*-format. Tip: use `ind2sub`.
2. Change **i** and **j** from slide 12 to adjacency matrices A_u and A_d. Tip: use `sub2ind`.
3. Use the algorithm in slide 16 to **i** and **j** from slide 12 to the CRS format.
4. Implement the algorithm in slide 17 and change **i** and **j** from slide 12 to CRS format.

Exercises - Delivery

1. Use the Dataset 1 and apply the following routines using the experience in class. Deadline: 24 April 2019.
2. Converte from *ij*-format to Adjacency matrix.
3. Elaborate an algorithm to converte from adjacency matrix to indicende matrix.
4. Adjacency matrix to *ij*-format.
5. *ij*-format to the list of neighbours.
6. CRS to *ij*-format.
7. *ij*-format to CRS format.