Chapter 2
Social Network Analysis: Concepts and Definitions

Abstract The aim of this chapter is to present the main definitions and concepts associated with social network analysis. These definitions and concepts will help to understand the fundaments of graph theory and the following micro, meso- and macro-measurements.

Keywords Social network analysis · Concepts · Graph theory · Digraphs

2.1 Definitions and Concepts

The Social Network Analysis (SNA) is based on Graph Theory (Barnes and Harary 1983), a mathematical study of sets of vertices connected by edges. The techniques model pairwise relations between the vertices.

To better understand the network perspective, consider the social network of Team of football players shown in Fig. 2.1. It is an example of a sociogram, also called a network graph, which is a common way of visualizing networks. Like all networks, it consists of two primary building blocks: vertices (also called nodes or points or agents) and edges (also called lines or ties or arcs or connections). The vertices are represented by number of players of a Team, and the edges are represented by the lines that point from one vertex to another. Other similar context, for example, is presented in (Hansen et al. 2011) to Twitter's users.

The kind of relations considered between vertices in graph theory, in terms of mathematics, are described by graphs that represent networks and those graphs are called directed graphs, undirected graphs, and weighted graphs (digraphs) (Pavlopoulos et al. 2011). Thus, in the following presented some elementary concepts on Graphs Theory used in SNA.

F.M. Clemente et al., *Social Network Analysis Applied to Team Sports Analysis*,
SpringerBriefs in Applied Sciences and Technology,
DOI 10.1007/978-3-319-25855-3_2

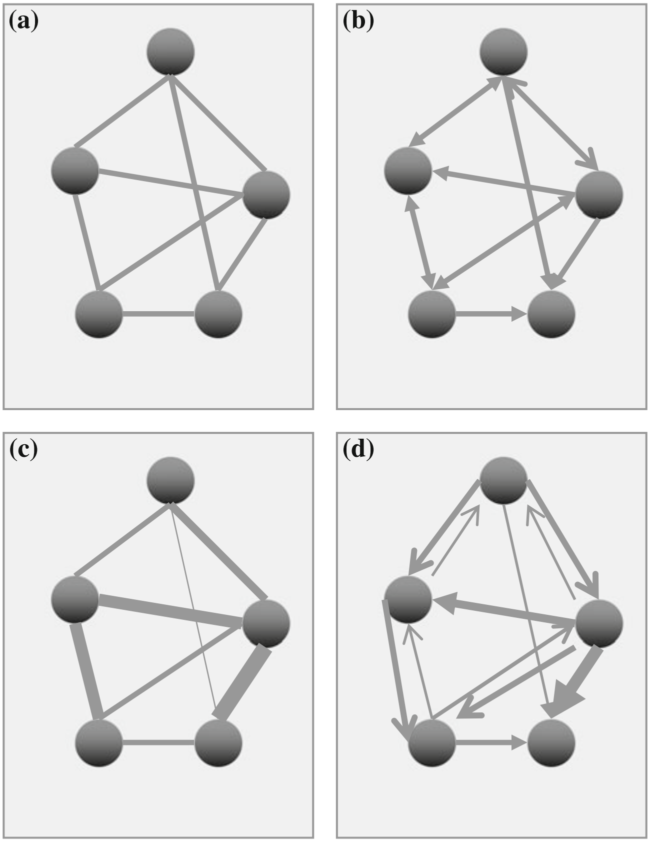

Fig. 2.1 Examples of unweighted graph (**a**), unweighted digraph (**b**), weighted graph (**c**), and weighted digraph (**d**)

Definition 2.1 (Gross and Yellen 2004) A graph $G = (V, E)$ consists of two sets V and E, where V is a set of vertices (or nodes), E is a set of edges and each edge has a set of one or two vertices associated to it, which are called its endpoints (or neighbors) and an edge is said to joint its endpoints.

Remark 2.1 Given two vertices u and v. Then the edge, the single connection between vertices u and v, is represented by pair (u, v).

Definition 2.2 (Gross and Yellen 2004) If vertex u is an endpoint of edge l, then u is said to be incident on l, and l is incident on u.

Definition 2.3 (Gross and Yellen 2004) A vertex u is adjacent to vertex v if they are joined by any edge.

Definition 2.4 (Pavlopoulos et al. 2011) A complete graph is a graph in which every pair of vertices is adjacent.

Definition 2.5 (Gross and Yellen 2004) Two adjacent vertices may be called neighbors.

Example 2.1 Let $G = (V, E)$ be a unweighted graph with $n = 5$ vertices (Fig. 2.2).

Definition 2.6 (Pavlopoulos et al. 2011) An undirected graph is connected if one can get from any vertex to any other vertex by following a sequence of edges.

Definition 2.7 (Gross and Yellen 2004) A walk in a graph G is an sequence of vertices and edges $W = v_0, e_1, v_1, e_2, \ldots, e_n, v_n$ shut that for $j = 1, \ldots, n$, the vertices v_{j-1} and v_j are the endpoints of the edge e_j.

Definition 2.8 (Gross and Yellen 2004) A graph that has no loops and includes no more than one edge between a pair of vertices is called a simple graph.

Definition 2.9 (Wasserman and Faust 1994) A graph G_S is a subgraph of G if the set of the vertices of G_S is a subset of vertices of G, and the set of edges in G_S is a subset of the edges in the graph G.

Definition 2.10 (Wasserman and Faust 1994) A subgraph G_S, is generated by a set of vertices if G_S has vertex set and edge set where the set of edges includes all edges of graph G that are between pairs of vertices of the G_S.

Remark 2.2 (Wasserman and Faust 1994) In Definition 2.10 is define vertex-generated subgraphs. In SNA context is only a subset of the g members of network.

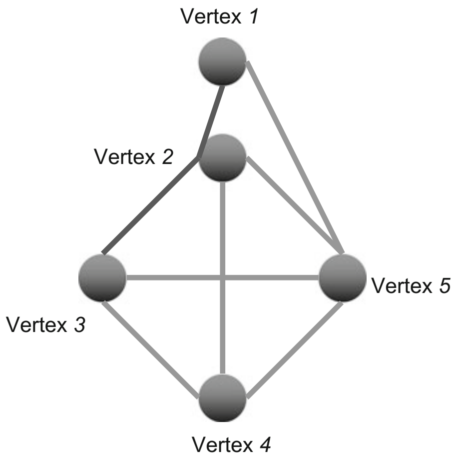

Fig. 2.2 Unweight graph G with five vertices. The set of all vertices is $V = \{n_1, n_2, n_3, n_4, n_5\}$ where $n_i := Vertex\,i$. The edge $e_1 = (n_1, n_5)$ *is the single connection between vertices* n_1 *and* n_5. The set all edges is $\mathrm{E} = \{e_1, e_2, e_3, e_4, e_5\} = \{(n_1, n_5), (n_1, n_2), (n_2, n_3), (n_2, n_5), (n_2, n_4), (n_3, n_5)$. The vertex n_1 *is neighbor of vertex* n_5 *and the set of all neighbors of* n_5 *is* $N(n_5) = \{n_1, n_2, n_3, n_4\}$

Definition 2.11 (Wasserman and Faust 1994) A dyad, representing a pair of agents and the possible edge between them is a (vertex-generated) subgraph consisting of a pair of vertices and the possible edge between the vertices.

Remark 2.3 (Wasserman and Faust 1994) In a graph an unordered pair of vertices can be in only one of two states: either two vertices are adjacent or they are not adjacent.

Definition 2.12 (Wasserman and Glaskiewicz 1994) A triad is a subgraph consisting of the tree vertices and the possible edges among them.

Remark 2.4 (Wasserman and Faust 1994) In a graph, a triad may be in one of four possible sates, depending on whether, zero, one, two or three edges are presented among of tree vertices in the triad.

Definition 2.13 (Wasserman and Faust 1994) A triad involving the vertices u, v, z is transitive if $(u, v), (v, z) \in E$ then $(u, z) \in E$.

Theorem 2.1 (Wasserman and Faust 1994) *A graph is transitive if every triad it contains is transitive.*

Definition 2.14 (Gross and Yellen 2004) The length of walk is the number of edges (counting repetitions).

Definition 2.15 (Gross and Yellen 2004) A walk is closed if the initial vertex is also the final vertex; otherwise, it is open.

Definition 2.16 (Gross and Yellen 2004) A trail in a graph is a walk shut that no edge occurs than more once.

Definition 2.17 (Gross and Yellen 2004) A path in a graph is a trial shut that no internal vertex is repeated.

Definition 2.18 (Gross and Yellen 2004) A cycle is a closed path of length at least 1.

Definition 2.19 (Wasserman and Glaskiewicz 1994) The geodesic distance, $d(n_i, n_j)$, between two vertices, n_i and n_j is the length of shortest path between them and in cases that no path was generated it is possible to set $d(n_i, n_j) = \infty$ assuming that the vertices are so far between each other so they are not connected.

Remark 2.5 (Wasserman and Faust 1994) In unweighted graph the distance between n_i and n_j is equal to the distance between n_j and n_i; $d(n_i, n_j) = d(n_j, n_i)$.

Definition 2.20 (Gross and Yellen 2004) Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, where G_1 and G_2 has the same number of vertices. G_1 and G_2 are isomorphic if for all $n_i, n_j \in V_1$ and $v_i, v_j \in V_2$ there exists a one-to-one mapping $f : V_1 \to V_2$, $f(n_i) = v_i$ and $f(n_j) = v_j$, such that $(n_i, n_j) \in E_1$ if only if $(v_i, v_j) \in E_2$.

Definition 2.21 (Pavlopoulos et al. 2011) Bipartite graph is an undirected graph $G = (V, E)$ in which V can be partitioned into two sets V_1 and V_2 such that $(u, v) \in E$ implies either $u \in V_1$ and $v \in V_2$ or $v \in V_1$ and $u \in V_2$.

The Definitions 2.1–2.21 represent the class of graphs called of undirected and unweighted graph.

The other broad class are the unweighted directed graphs that we show the following in Definition 2.22–2.41.

Definition 2.22 (Gross and Yellen 2004) A directed graph (or digraph) is a graph each of whose edges is directed shut that a direct edge (or arc) is an edge that linked in the initial vertex to the terminal vertex.

Remark 2.6 (Wasserman and Faust 1994) The difference between an arc (in a digraph) and a edge (in a graph) is that an arc is an ordered pair of vertices (to reflect the direction of the edge between of two vertices) whereas a edge is unordered pair of vertices.

Definition 2.23 (Wasserman and Glaskiewicz 1994) Considering that a given vertex is the first (sender) or second (receiver) in the ordered pair defining the arc. A vertex u is adjacent to vertex v if the arc between u and v exist in digraph and a vertex v is adjacent to vertex u if the arc between v and u exist in digraph.

Remark 2.7 (Wasserman and Faust 1994) When a digraph is presented as a diagram the vertices as represented as points and the arcs represented as directed arrows. The arc (u, v) is represented by any arrow from the point representing u to the point representing v.

Example 2.2 Let $G = (V, E)$ be a unweighted digraph with $n = 5$ vertices (Fig. 2.3).

Definition 2.24 (Gross and Yellen 2004) A digraph that has no loops and includes no more than one arc that linked in the initial node to the terminal node is called a simple digraph.

Definition 2.25 (Wasserman and Faust 1994) A dyad in digraphs is subgraph that consisting of the two vertices and the possible arcs between of them.

Definition 2.26 (Wasserman and Faust 1994) A triad in digraphs is subgraph that consisting of the tree vertices and the possible arcs between of them.

Theorem 2.2 (Wasserman and Faust 1994) A digraph is transitive if every triad it contains is transitive.

Definition 2.27 (Gross and Yellen 2004) In a digraph G, W is directed walk, if the edge e_j is directed from v_{j-1} to v_j, where v_{j-1} and v_j are vertices of the G.

Definition 2.28 (Wasserman and Faust 1994) The length of directed walk is the number of instances arcs in it (an arc is counted each time it occurs in the walk).

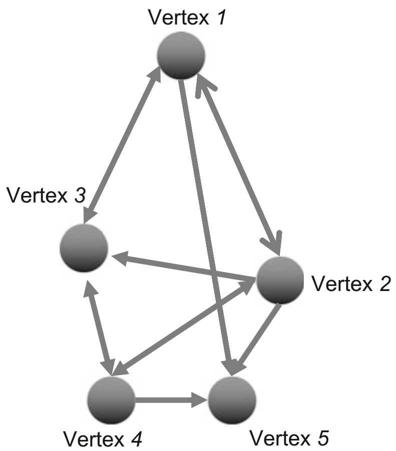

Fig. 2.3 Unweight digraph with five vertices. The set of all vertices is $V = \{n_1, n_2, n_3, n_4, n_5\}$ where $n_i := Vertex\, i$. The arc $e_1 = (n_2, n_5)$ *is the single connection of* n_2 *for* n_5. The set of all arcs is $\mathrm{E} = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}, e_{11}, e_{12}\} = \{(n_2, n_5), (n_2, n_3), (n_2, n_4), (n_2, n_1), (n_1, n_2), (n_4, n_2), (n_4, n_3), (n_4, n_5), (n_3, n_4), (n_3, n_1), (n_1, n_3), (n_1, n_5)\}$. The vertex n_5 *is outgoing neighbor of vertex* n_1 *and the set of all outgoing neighbors of* n_1 *is* $N^{out}(n_1) = \{n_2, n_3, n_5\}$. n_2 *is incoming neighbor of* n_1 *and he set of all incoming neighbors of* n_1 *is* $N^{in}(n_1) = \{n_2, n_3\}$

Definition 2.29 (Wasserman and Glaskiewicz 1994) A directed trail in a digraph is a directed walk in which no arc is included more than one.

Definition 2.30 (Bondy and Murty 2008) A directed path (or simply a path) in digraph is a directed walk in which no vertex and no arc is included more than once.

Definition 2.31 (Wasserman and Faust 1994) A length of a path is the number of arcs in it.

Definition 2.32 (Wasserman and Glaskiewicz 1994) A semiwalk joining vertices v_i and v_j is a sequence of vertices and arcs which successive pairs of vertices are incident with an arc from the first to the second, or by an are from the second to the first.

Remark 2.8 In a semiwalk the direction of the arcs is irrelevant.

Definition 2.33 (Wasserman and Faust 1994) The length of a semiwalk is the number of instances of arcs in it.

Definition 2.34 (Wasserman and Faust 1994) A semipath joining vertices v_i and v_j is a sequence of distinct vertices were all successive pairs of vertices are connected by an arc from the first to the second, or by an arc from the second to the first for all successive pairs of vertices.

Remark 2.9 In a semipath the direction of the arcs is irrelevant.

Definition 2.35 (Wasserman and Faust 1994) The length of a semipath is the number of instances of arcs in it.

Definition 2.36 (Wasserman and Faust 1994) A cycle in digraph is a closed direct walk of at least tree vertices in which all vertices except the first and the last are distinct.

Definition 2.37 (Wasserman and Faust 1994) A semicycle in digraph is a closed direct semiwalk of at least tree vertices in which all vertices except the first and the last are distinct.

Remark 2.10 (Wasserman and Faust 1994) In a semicycle the arcs may go in either direction, whereas in a cycle the arcs must "point" in same direction.

Definition 2.38 (Wasserman and Glaskiewicz 1994) Given a pair of vertices n_i, n_j of digraph G with n vertices, $i, j = 1, \ldots, n$ and $i \neq j$. A pair of vertices n_i and n_j, is:

(i) Weakly connected if they are joined by a semipath;
(ii) Unilaterally connected if they are joined by a path from n_i, n_j, or a path from n_i to n_j;
(iii) Strongly connected if there is a path from n_i to n_j and a path from n_j to n_i; the path from n_i to n_j may contain different vertices and arcs than the path from n_j to n_i;
(iv) Recursively connected if they are strongly connected and the path from n_i to n_j uses the same vertices and arcs as the path from n_j to n_i in reverse order.

Definition 2.39 (Gross and Yellen 2004) A directed graph is:

(i) Weakly connected if all pairs of vertices are Weakly connected;
(ii) Unilaterally connected if all pairs of vertices are unilaterally connected;
(iii) Strongly connected if all pairs of vertices are strongly connected;
(iv) Recursively connected if all pairs of nodes are recursively connected.

Definition 2.40 (Rubinov and Sporns 2010; Wasserman and Faust 1994) The geodesic distance, $d(n_i, n_j)$, between two vertices, n_i and n_j is the length of shortest path between them and in cases that no path was generated it is possible to set $d(n_i, n_j) = \infty$ assuming that the vertices are so far between each other so they are not connected.

Remark 2.11 In unweighted digraph the distance between n_i and n_j is not equal to the distance between n_j and n_i; $d(n_i, n_j) \neq d(n_j, n_i)$ (Wasserman and Faust 1994).

Definition 2.41 (Chen 1997) Two digraphs are said to be isomorphic if verify simultaneously the following conditions:

(i) Their associated undirected graphs are isomorphic;
(ii) The directions of their corresponding edges are preserved for some correspondences of (i).

Definition 2.42 (Wasserman and Glaskiewicz 1994) In the case of a any graph (digraph) with all connections measured between any two vertices is called weighted graph (digraph).

Remark 2.12 (Wasserman and Faust 1994) The value of the weight of edge between any two vertices, v_i, v_j is represented by w_{ij}. In a valued graph the edge between vertex u and vertex v is identical to the edge between vertex v and vertex u and thus there is only a single value for each unordered pair of vertices. In the case weighted digraph, a valued digraph the arc between vertex u and vertex v no is identical to the arc between vertex v and vertex u and thus there is two values, one for each possible arc for the ordered pair of vertices.

Example 2.3 Given one weighted graph G_1 and one weighted digraph G_2, both with n = 5 vertices (Fig. 2.4).

Definition 2.43 (Wasserman and Faust 1994) A dyad in weighted graphs has a edge between vertices with a specific strength and in weighted digraphs has arcs between vertices with a specific strength, respectively.

Definition 2.44 (Wasserman and Faust 1994) A triad in weighted graphs has a edges between vertices with a specific strength, respectively, and in weighted digraphs has arcs between vertices with a specific strength, respectively.

Definition 2.45 (Umeyama 1988) Two weighted undirected graphs, G_1, G_2 are said to be isomorphic if verify simultaneously the following conditions:

(i) Their associated unweighted and undirected graphs are isomorphic;

(ii) $\sum_{i=1}^{n}\sum_{j=1}^{n}\left(w_{ij}^1 - w_{ij}^2\right)^2 = 0$, where w_{ij}^1 is the weight of the edge between any two vertices, n_i, n_j and w_{ij}^2 is the weight of the edge between any two vertices, v_i, v_j with $f(n_i) = v_i$ and $f\left(n_j\right) = v_j$, such that $\left(n_i, n_j\right) \in E_1$ and $\left(v_i, v_j\right) \in E_2$.

Definition 2.46 (Bunimovich and Webb 2014) Two weighted digraphs, G_1, G_2 are said to be isomorphic if for all $n_i, n_j \in V_1$ and $v_i, v_j \in V_2$ there exists a one-to-one mapping $f : V_1 \rightarrow V_2, f(n_i) = v_i$ and $f\left(n_j\right) = v_j$, such that w_{ij}^1 is the weight of the arc from n_i to n_j with $\left(n_i, n_j\right) \in E_1$ if only if w_{ij}^2 is the weight of the arc from v_i to, v_j with $\left(v_i, v_j\right) \in E_2$ and $w_{ij}^1 = w_{ij}^2$.

The information of an unweighted or weighted graph or digraph may also be expressed in a variety of ways in matrix form presented in the Definition 16.

Definition 2.47 (Wasserman and Faust 1994) The $n \times n$ matrix $A = \left[a_{ij}\right]$ is the adjacency matrix of a unweighted graph or unweighted digraph with n vertices such that $a_{ij} = 1$ if $\left(n_i, n_j\right) \in E$ or $a_{ij} = 0$ if $\left(n_i, n_j\right) \notin E$, $i, j = 1, \ldots, n$. In the case of weighted graphs or weighted digraphs $a_{ij} = w_{ij}$ if $\left(n_i, n_j\right) \in E$ or $a_{ij} = 0$ otherwise. Thus, the matrix A is called weighted adjacency matrix.

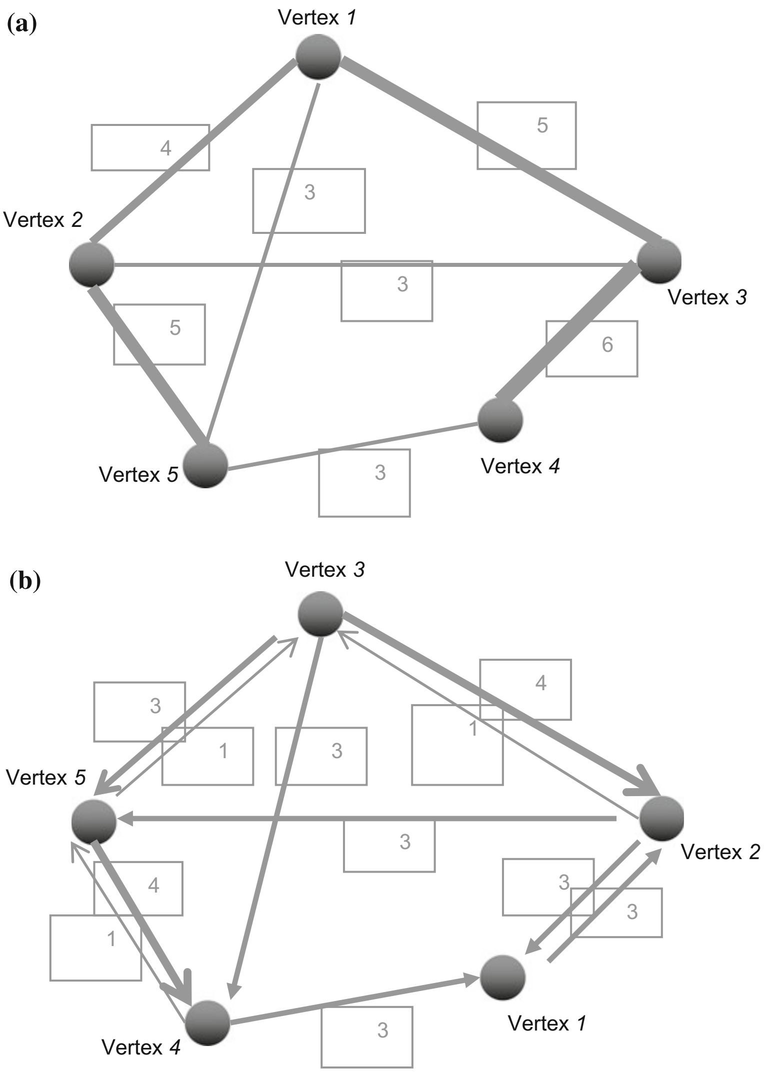

Fig. 2.4 Weighted graph (**a**) and a weighted digraph (**b**). **Weighted graph** $\boldsymbol{G_1}$. The set of all vertices is $V_1 = \{n_1, n_2, n_3, n_4, n_5\}$ where $n_i :=$ *Vertex i*. The edge $e_1 = (n_1, n_5)$ *is the single connection between vertices* n_1 *and* n_5. The set all edges is $E = \{(n_1, n_5), (n_1, n_2), (n_1, n_3), (n_2, n_3), (n_2, n_5), (n_5, n_4), (n_4, n_3)\}$. The vertex n_1 *is neighbor of vertex* n_2 *and the set of all neighbors of* n_2 *is* $N(n_2) = \{n_1, n_3, n_5\}$. *The* $w_{12} = 4$ *is the weight of the edge* $(n_1, n_2) \in E_1$. *We can observe that* $w_{12} = w_{21} = 4$. **Weighted digraph** $\boldsymbol{G_2}$. The set of all vertices is $V_2 = \{n_1, n_2, n_3, n_4, n_5\}$ where $n_i :=$ *Vertex i*. The arc $e_5 = (n_3, n_2)$ *is the single connection of* n_3 *for* n_2 *and the* $w_{32} = 4$ *is the weight of the arc* $(n_3, n_2) \in E_2$. *However,* $w_{23} = 1$ *and, in general,* $w_{ij} \neq w_{ji}$. The set of all arcs is $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}, e_{11}\} = \{(n_1, n_2), (n_2, n_1), (n_2, n_5), (n_2, n_3), (n_3, n_2), (n_3, n_4), (n_3, n_5), (n_5, n_3), (n_5, n_4), (n_4, n_5), (n_4, n_1)\}$. The vertex n_3 *is outgoing neighbor of vertex* n_2 *and the set of all outgoing neighbors of* n_2 *is* $N^{out}(n_2) = \{n_1, n_3, n_5\}$. n_1 *is incoming neighbor of* n_2 *and he set of all incoming neighbors of* n_2 *is* $N^{in}(n_2) = \{n_1, n_3\}$

Remark 2.13 (Wasserman and Faust 1994) One $n \times n$ matrix is also called a matrix of order n. On the other hand how the elements a_{ij} of an (weighted) adjacency matrix A are real numbers so we can write $A \in \mathcal{R}^n$.

Example 2.4 In this example will go to show the corresponding (weighted) adjacency matrix associated with the (weighted) graphs and digraphs presented in the Examples 2.1, 2.2 and 2.3.

The adjacency matrix A associated with unweighted graph G presented in Example 2.1 is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \in \mathcal{R}^{5\times 5}.$$

The adjacency matrix A associated with unweighted digraph G presented in Example 2.2 is

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathcal{R}^{5\times 5}.$$

The weighted adjacency matrix A associated with weighted graph G_1 presented in Example 2.3(a) is

$$A = \begin{bmatrix} 0 & 4 & 5 & 0 & 3 \\ 4 & 0 & 3 & 0 & 5 \\ 5 & 3 & 0 & 6 & 0 \\ 0 & 0 & 6 & 0 & 3 \\ 3 & 5 & 0 & 3 & 0 \end{bmatrix} \in \mathcal{R}^{5\times 5}.$$

The weighted adjacency matrix A associated with weighted digraph G_2 presented in Example 2.3(b) is

$$A = \begin{bmatrix} 0 & 3 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 3 \\ 0 & 4 & 0 & 3 & 3 \\ 3 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 4 & 0 \end{bmatrix} \in \mathcal{R}^{5\times 5}.$$

References

Barnes, J. A., & Harary, F. (1983). Graph theory in network analysis. *Social Networks, 5*, 235–244.

Bondy, J. A., & Murty, U. S. R. (2008). *Graph Theory*. USA: Springer.

Bunimovich, L., & Webb, B. (2014). *Isospectral Transformations: A New Approach to Analysing Multidimensional Systems and Networks*. New York, USA: Springer.

Chen, W. K. (1997). *Graph Theory and Its Engineering Applications*. Singapore: World Scientific Publishing Co., Pte. Ltd.

Gross, L., & Yellen, J. (2004). *Handbook of Graph Theory*. New York, USA: CRC Press.

Hansen, D., Shneiderman, B., & Smith, M. A. (2011). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Burlington, USA: Morgan Kaufmann—Elsevier.

Pavlopoulos, G. A., et al. (2011). Using graph theory to analyze biological networks. *BioData Mining, 4*(1), 10.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage, 52*(3), 1059–1069.

Umeyama, S. (1988). An eigendecomposition approach to weighted graph matching problems. *IEE Transactions on Pattern Analysis and Machine Intelligence, 10*(5), 1988.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York, USA: Cambridge University Press.

Wasserman, S., & Glaskiewicz, J. (1994). *Advances in Social Network Analysis: Research in the Social and Behavioral*. California, USA: SAGE Publications.

http://www.springer.com/978-3-319-25854-6

Social Network Analysis Applied to Team Sports Analysis
Clemente, F.M.; Martins, F.M.L.; Mendes, R.S.
2016, XII, 93 p. 10 illus. in color., Softcover
ISBN: 978-3-319-25854-6