Guze Sambor
Gdynia Maritime University, Poland

The graph theory approach to analyze critical infrastructures of transportation systems

Keywords

critical infrastructures, domination set, domination number, minimal spanning tree

Abstract

The main aim of the paper is to use algorithms and parameters of graph theory as tool to analyze the transpiration systems. To realize this goal the well-known information about graph theory algorithms and parameters will be introduced and described. The possible application of graph theory algorithms and parameters to analyze the critical infrastructures of exemplary transportation system will be shown.

1. Introduction

Nowadays people are dependent on critical systems such as transportation, electricity, water supply, sewage, ICT. It is important thing to keep these complex systems in good condition. Thus risk and reliability analyses are needed to understand the impact of threats and hazards. However, these problems become more and more complex, because of existing strong interdependencies both within and between infrastructure systems.

2. Critical infrastructures

Generally, according to the official definition, the critical infrastructure is a term used to describe assets that are essential for the functioning of a society and economy. The following facilities are related to this subject [1]:

- electricity and heating generation, transmission and distribution;
- gas and oil production, transport and distribution;
- telecommunication;
- water supply;
- agriculture, food production and distribution;
- public health (hospitals, ambulances);
- transportation systems (fuel supply, railway network, airports, harbours, inland shipping);
- financial services (banking);
- security services (police, military).

There are several regional critical-infrastructure protection programmes. The main aims of all of them are:

- to indentify important assets,
- to analyze a risk based on major threat scenarios and the vulnerability of each asset,
- to indentify, select and make prioritisation of counter-measures and procedures.

These goals are common for all facilities presented above. In the paper, we take into account only the transportation systems. However, the presented approach can be applied to any of listed facilities.

3. Review of graph theory topics

Bellow chapter is showing general definitions, parameters and algorithms of the domination in graph theory and other related topics. It will be used in further part of the paper to analyze critical infrastructures of exemplary transportation systems.

3.1. Domination in graphs

Let $G=(V,E)$ be a connected simple graph where V - set of n vertices , E - set of m edges. The set of neighbors of vertex v in G is denoted by $N_G(v)$.

In the following definitions the different type of domination sets and numbers will be introduced, i.e. general, connected and independent [4], [5].

Definition 1

A set $D \subseteq V(G)$ is the dominating set of G if for any $v \in V$ either $v \in D$ or $N_G(v) \cap D \neq \varnothing$. The domination number $\gamma(G)$ of a graph G is the minimum cardinality of a dominating set of G.

Definition 2
A set $D_C \subseteq V(G)$ is a connected dominating set of G if every vertex of $V \setminus D_C$ is adjacent to a vertex in D_C and the subgraph induced by D_C is connected. The minimum cardinality of a connected dominating set of G is the connected domination number $\gamma_c(G)$.

Definition 3
A set $D_I \subseteq V(G)$ is an independent dominating set of G if no two vertices of D_I are connected by any edge of G. The minimum cardinality of an independent dominating set of G is the independent domination number $\gamma_I(G)$.
The concept related to domination in graphs is also the bondage number. The definition is as follows [3].

Definition 4
The bondage number $b(G)$ of graph G is the cardinality of a smallest set E of edges for which $\gamma(G-E) > \gamma(G)$.

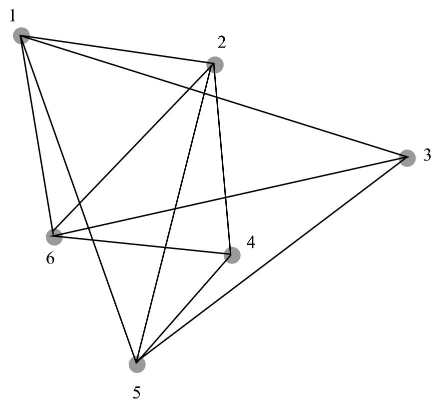

Figure 1. The exemplary undirected graph G.

For the exemplary graph G presented in Figure 1 the particular domination sets and numbers (Definitions 1 – 3) are as follows:

$D = \{1,2,5\}$, $\gamma(G) = 2$;
$D_C = \{2,4,6\}$, $\gamma_c(G) = 2$;
$D_I = \{2,3\}$, $\gamma_I(G) = 2$.

The above definitions refer to the general concept of undirected graphs. There are topics related to weight functions. Thus, the vertex-weighted graph is defined ([4], [5]).

Definition 5
A vertex-weighted graph (G, w_v) is defined as a graph G together with a positive weight-function on its vertex set $w_v:\ V(G) \to R > 0$.
According to above definition we get next definitions.

Definition 6
The weighted domination number $\gamma_w(G)$ of (G, w_v) is the minimum weight $w(D) = \sum_{v \in D} w(v)$ of a set $D \subseteq V(G)$ such that every vertex $x \in V(D) - D$ has a neighbor in D.

Definition 7
The weighted connected domination number $\gamma_{Cw}(G)$ of (G, w_v) is the minimum weight $w(D_C) = \sum_{v \in D_C} w(v)$ of a set $D_C \subseteq V(G)$ such that every vertex of $V \setminus D_C$ is adjacent to a vertex in D_C and the subgraph induced by D_C is connected.

Definition 8
The weighted independent domination number $\gamma_{Iw}(G)$ of (G, w_v) is the minimum weight $w(D_I) = \sum_{v \in D_I} w(v)$ of a set $D_I \subseteq V(G)$ such that if no two vertices of D_I are connected by any edge of (G, w).

Let us assume, that labels on graph (Figure 1) represents the weight-function. The domination numbers defined in Definitions 6 – 8, for this graph are equals to:

$\gamma_w(G) = 3$,
$\gamma_{Cw}(G) = 3$
$\gamma_{Iw}(G) = 5$.

In general, finding a minimum dominating set is NP-hard problem, but efficient approximation algorithms do exist. It can be done under assumption that any dominating set problem can be formulated as a set covering problem. Thus, the greedy algorithm for finding domination set is an analog of one that has been presented in [7]. This algorithms is formulated as follows [7]:

Algorithm 1:
1. Let $V = \{1, ..., n\}$, and define $D = \phi$.
2. Greedy add a new node to D in each iteration, until D forms a dominating set.

3. A node j, is said to be covered if $j \in D$ or if any neighbor of j is in D. A node that is not covered is said to be uncovered.
4. In each iteration, put into D the least indexed node that covers the maximum number of uncovered nodes.
5. Stop when all the nodes are covered.

For graph in Figure 1, Greedy will return $D = \{1,2\}$.

In case of the minimum connected domination set, the Greedy algorithm is also used. However, to define them some preliminaries are necessary[8].
We consider graph G and subset C of vertices in G. We can divided all vertices into three classes:

- belong to C are called black v_b;
- not belong to C but adjacent to C are called gray v_g;
- not in C and not adjacent to C are called white v_w.

Under assumption that C is connected dominating set if and only if there is no white vertex and the subgraph induced by black vertices is connected. The sum of the number of white vertices and the number of connected components of the subgraph induced by black vertices (black components) equals 1. The greedy algorithm with potential function equal to the number of white vertices plus number of black components is as follows [8].

Algorithm 2:
Set $w := 1$;
while $w = 1$ ***do***
If there exists a white or gray vertex such that coloring it in black and its adjacent white vertices in gray would reduce the value of potential function
then choose such a vertex to make the value of potential function reduce in a maximum amount
else set $w := 0$;

3.2. Spanning trees

The issues discussed in Subsection 3.1. and 3.2. was related to the analysis of nodes of the transportation system. However, the edges between the nodes are also important. As an appropriate tool for the transportation system analysis in terms of its infrastructure connecting each node the spanning tree is proposed.

Definition 9
The spanning tree *T* of a connected, undirected graph *G* is a tree composed of all the vertices and some (or perhaps all) of the edges of *G*.

Informally, a spanning tree of *G* is a selection of edges of *G* that form a tree *spanning* every vertex. It means, that every vertex lies in the tree, but no cycles (or loops) are formed.

According to Definition 5, the edge-weighted graph is introduced.

Definition 10
An *edge-weighted graph* (G, w_e) is defined as a graph graph G together with a positive weight-function on its edge set w_e : $E(G) \to R > 0$.

Furthermore, for edge-weighted graphs it is possible to find minimum spanning tree, which is defined as follows.

Definition 11
A *minimum spanning tree (MST)* of an edge-weighted graph is a spanning tree whose weight (the sum of the weights of its edges) is no greater than the weight of any other spanning tree.
It can be done according to two well-known algorithms: Kruskal's and Prim's. They can be shown as follows [2]:

Algortihm 3 (Kruskal's)
1. Find the cheapest edge in the graph (if there is more than one, pick one at random). Mark it with any given colour, say red.
2. Find the cheapest unmarked (uncoloured) edge in the graph that doesn't close a coloured or red circuit. Mark this edge red.
3. Repeat *Step 2* until you reach out to every vertex of the graph (or you have *n-1* coloured edges).

The red edges form the desired minimum spanning tree.

Algortihm 4 (Prim's)
1. Pick any vertex as a starting vertex - v_{start} . Mark it with any given colour (red).
2. Find the nearest neighbor of v_{start} (call it P_1). Mark both P_1 and the edge $v_{start}P_1$ red. Cheapest unmarked (uncoloured) edge in the graph that doesn't close a coloured circuit. Mark this edge with same colour of Step 2.
3. Find the nearest uncoloured neighbor to the red subgraph (i.e., the closest vertex to any red vertex). Mark it and the edge connecting the vertex to the red subgraph in red.
4. Repeat Step 3 until all vertices are marked red.

The red subgraph is a minimum spanning tree.

4. Application of graph theory algorithms to analyze transportation systems

The knowledge presented in Section 3 can be applied to analyze transportation system, particularly its infrastructure. For instance it can be used to choice of routes and nodes classified as critical infrastructure [6]. Using these tools in this way, it is possible to improve the safety and reliability considered systems. Unfortunately, it is possible only during designing and selection of critical infrastructures. In these stages using the number of failures parameter should be sufficient for safety and reliability analysis. In the other hand, the maintenance cost parameter allows us to cost-optimal design of the critical infrastructure. Therefore, the possibility of improving the safety and reliability of the system corresponds to select the optimal critical infrastructures of the transportation systems according to mentioned parameters. To show possibility of application discussed methods, the academic example is shown bellow.

Example

Let us consider the PKP EIC map of connections (*Figure 2*). We choose twelve nodes and describe them the number and the maintenance costs. The edges between the nodes are described by the number of failures. Under these assumption, we get the exemplary schema presented in *Figure 3*.

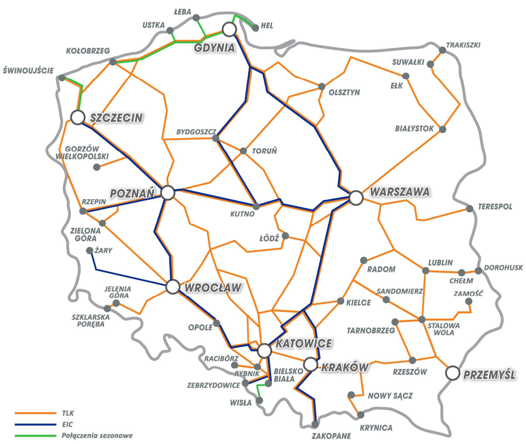

Figure 2. The map of PKP EIC connections, http://intercity.pl

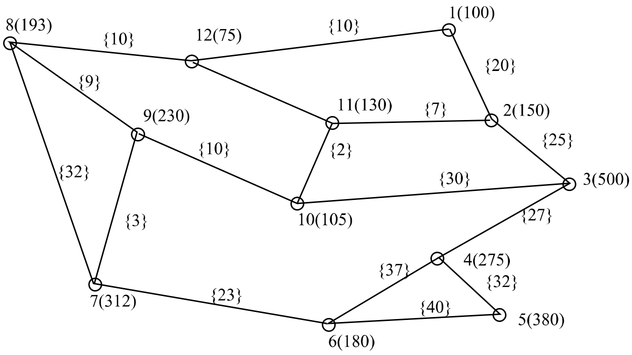

Figure 3. The exemplary scheme of railway infrastructure

Our main goal is to choose the minimal spanning tree and minimal independent domination set, i.e. finding an independent domination number.

According the Kruskal's Algorithm, the minimal spanning tree is given in *Figure 4* as the set of double edges. In this way, the minimal number of failures is equal to 153 [units].

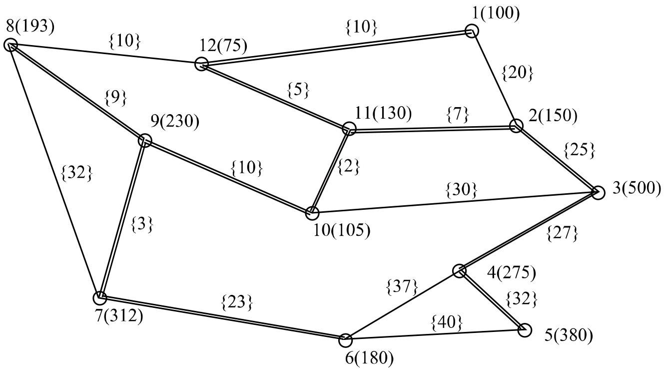

Figure 4. Minimal spanning tree (Kruskal's Algortihm)

Furthermore, according to Algorithm 1, we can find minimal dominating set. It is marked in *Figure 5* with black nodes (vertices). For resulting domination set, the maintenance costs are equaled to 430 [units].

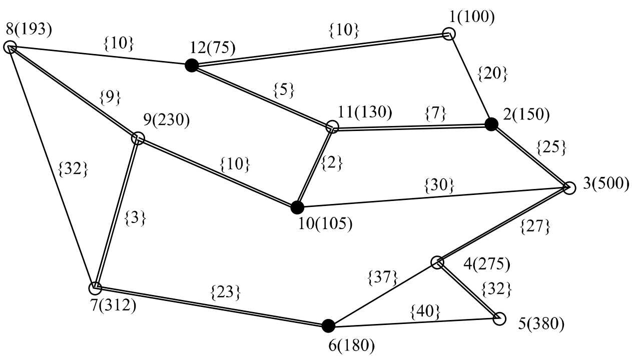

Figure 5. Minimal weighted independent domination set.

5. Conclusion

In the paper, the concept of critical infrastructures has been presented. Furthermore, the selected definitions, parameters and algorithms of graph theory have been introduced and applied to system transportation analysis. As the example of possible application, the critical infrastructure with the lowest maintenance cost of nodes and the smallest number of failure has been determined.

The example is only a presentation of the possibilities of the methods in stationary case.

In future work, these methods will be expand to non-stationary case.

References

[1] COMMISSION OF THE EUROPEAN COMMUNITIES (2006), Communication from the Commission on a European Programme for Critical Infrastructure Protection, Brussels.

[2] Cormen, T.H. & al. (2009). Introduction to Algorithms. *Third Edition. MIT Press*, ISBN 0-262-03384-4. Section 23.2: The algorithms of Kruskal and Prim, pp. 631–638.

[3] Fink, J.F. et. al. (1990). The bondage number of a graph. *Discrete Math*., 86 (1990), pp. 47–57.

[4] Harary, F. (1969). Graph Theory. *Addison-Wesley*, Reading.

[5] Haynes T. W., Hedetniemi, S., Slater, P. (1988). Fundamentals of Domination in Graphs. *CRC Press*.

[6] Kołowrocki, K. (2013). Safety of critical infrastructures. *Journal of Polish Safety and Reliability Association, Summer Safety and Reliability Seminars*. Vol. 4, No. 1, pp. 51-74.

[7] Parekh, A.K. (1991). Analysis of Greedy Heuristic for Finding Small Dominating Sets in Graphs. *Information Processing Letters* Vol. 39, Issue 5, 237 – 240.

[8] Ruan, L. & al., (2004). A greedy approximation for minimum connected dominating sets. *Theoretical Computer Science* Vol. 329, Issues 1–3, 325–330.