An aproximmation to the Minimum Branch Vertex Spanning Tree problem

Ramón Daniel Regueiro Espiño, Avraham Rosenberg & Abdoulaye Sakho

École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise

Introduction

- Spanning trees : a way to modelize flow structure.
- Aim : to obtain enough good approximated solutions of the Minimum Branch Vertex Spanning Tree (MBVST)

Our project phases :

- **Operational Research** part :
  1. Construction of the ILP
  2. Nodes Preprocessing
- **Data Science** part :
  1. Estimation of the number of branch vertex and the time to solve the ILP
  2. Estimation of branch vertex nodes
- Final deliverable : algorithm given a spanning tree

The ILP problem

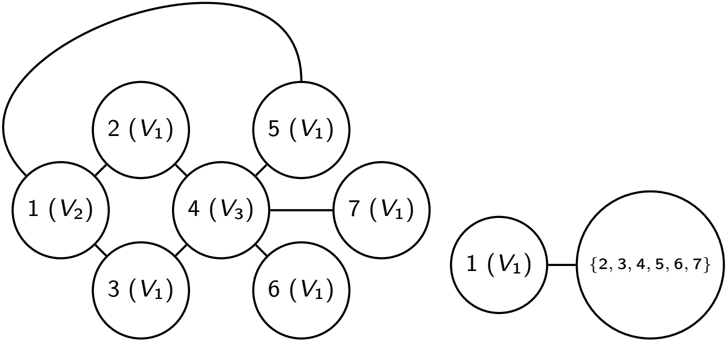
$$
\begin{aligned}
\min \ & \sum_{v \in V} y_v \\
\text{subject to : } & \sum_{e \in E} x_e = n - 1, \\
& \sum_{(s,v) \in A^+(s)} f_{sv} - \sum_{(v,s) \in A^-(s)} f_{vs} = n - 1, \\
& \sum_{(v,u) \in A^+(v)} f_{vu} - \sum_{(u,v) \in A^-(v)} f_{uv} = -1 \ \forall v \in V \setminus \{s\}, \\
& f_{uv} \leq (n-1) x_e \ \forall e \in \{u, v\} \in E, \\
& f_{vu} \leq (n-1) x_e \ \forall e \in \{u, v\} \in E, \\
& \sum_{e \in A(v)} x_e - 2 \leq (n-1) y_v \ v \in V, \\
& x_e \in \{0, 1\} \ \forall e \in E, \quad y_v \in \{0, 1\} \forall v \in V, \quad f_{u,v}, f_{vu} \geq 0 \ \forall e = \{u, v\} \in E
\end{aligned}
$$

Nodes Preprocessing I

Following [Merabet et al., 2018], we can clasify our nodes in three types :

- V_1 : nodes with no more than k edges.
- V_3 : nodes that deleting them and its edges will result in at least $k+1$ connected components.
- V_2 : $V \setminus (V_1 \cup V_3)$, these are our candidate nodes.

Nodes Preprocessing II



(a) Before

(b) After

Figure – Tranformation using the method for a graph G.

Effect of graph reduction

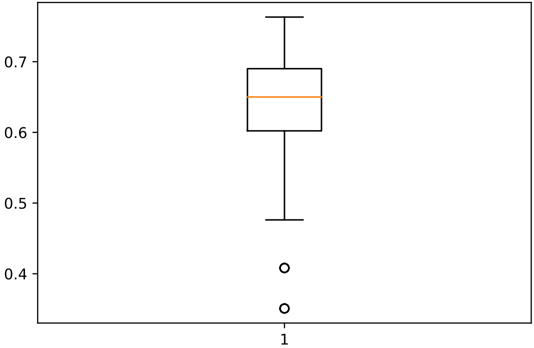


Figure – Relation between the number of nodes after the reduction of the graph and the initial number of node of the graph.

The first data frame

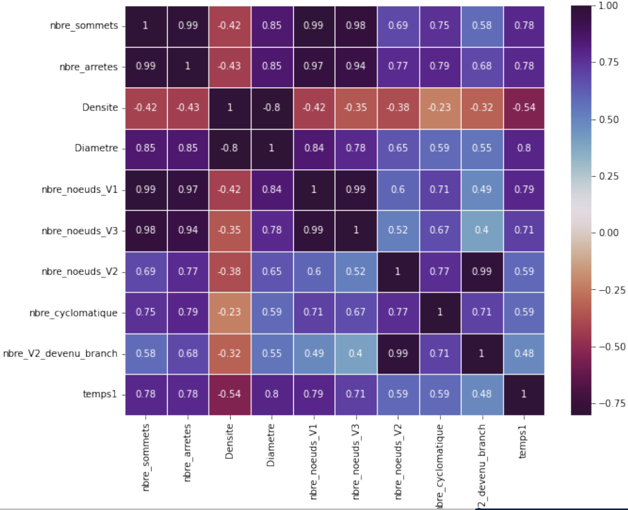
In a first step we constructed a data frame where each row/observation represents a given graph. For each graph we have retrieved several pieces of information, here they are :

- nbre_sommets
- nbre_arretes
- Densite
- Diametre
- nbre_noeuds_V1
- nbre_noeuds_V3
- nbre_noeuds_V2
- nbre_cyclomatique
- nbre_V2_devenu_branch
- temps1

This dataframe was used to build two types of Machine Learning models, which we will see later in the presentation

The first data frame

Study of correlations :



Estimation of the time to solve the ILP

Using the same dataframe, this model predicts, for a given graph, the time it will take the LP to find the solution graph of the problem.

- `Models`
- `Results`
- `Utility for Linear Programming`

Estimation of the number of branch vertex

Using the previous dataframe, this model predicts, for a given graph, the number of branching nodes that the associated solution graph will have.

- `Models`
- `Results`
- `Quantiles Regression`
- `Utility for Linear Programming`

Estimation of the branch vertex nodes

A model that for a given node of a given graph, predicts whether that node will be a branching node in the associated solution graph or not. The dataframe for this kind of models :

- `nbre_voisins_V1`
- `nbre_voisins_V2`
- `nbre_voisins_V3`
- `nbre_comp_connexes`
- `nbre_base_cycle`
- `Is_V2`

Estimation of the branch vertex nodes

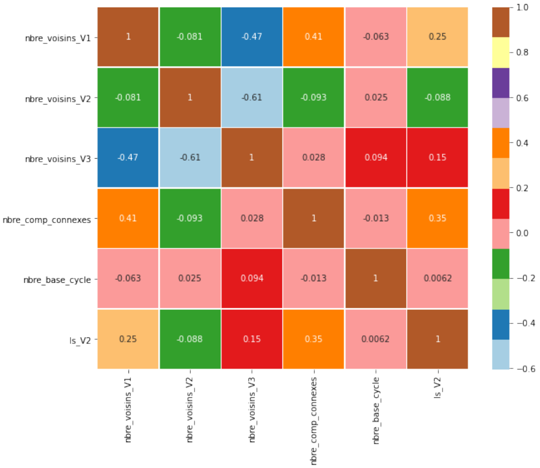
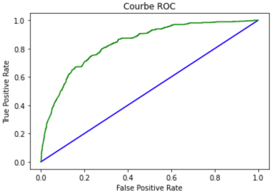


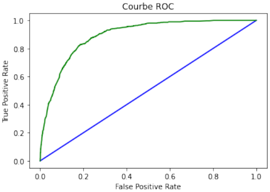
Figure – Correlations of the second dataframe

Estimation of the branch vertex nodes

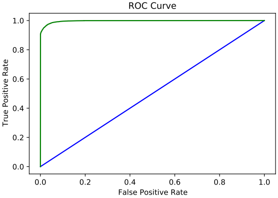
Models and results :



(a) Logistic Regression



(b) Random Forest



(c) Neural Network

Figure – ROC Curves for the considered models.

Final algorithm

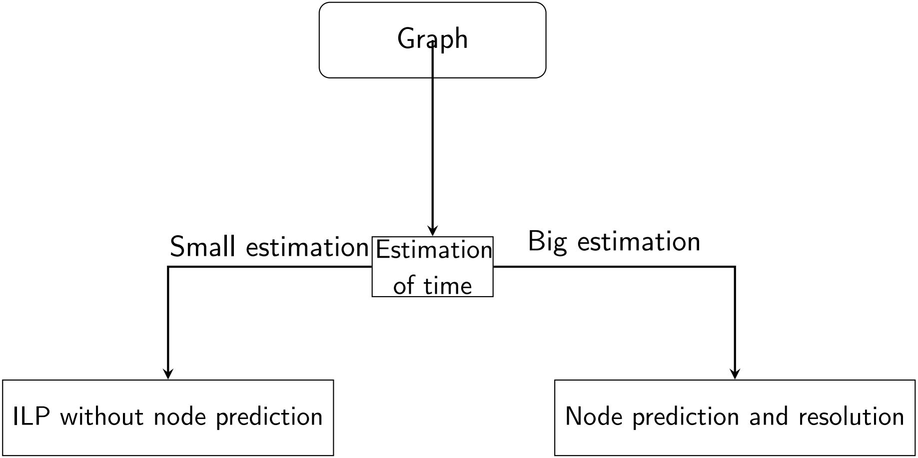


Figure – Flow chart of the final algorithm

Further improvements

- To improve node estimation.
- To measure the trade-off between the ILP time resolution and the probability of being considered branch or no-branch.
- To apply other problem reduction tools.

Références principales

Merabet, M., Desai, J., and Molnár, M. (2018).
A Generalization of the Minimum Branch Vertices Spanning Tree Problem.
In *ROADEF 2018 - 19ème congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision*, Lorient, France.