Exercises with AMPL

Antonio Mucherino

Laboratorie d'Informatique, École Polytechnique

Course on Operations Research (ISC610A)

Semester I - 2008/09

TD4 - November 27^{th} 2008

The *Proogle* project

The Proogle project has as aim to correlate past and current projects of an important firm, in order to help predicting the developments of current projects.

We studied the Proogle project and we found that three interesting problems needed to be solved:

1. clustering of past and current projects;
2. location of the densest subgraph of a given graph G having edges of the same color, in order to create interfaces in the software architecture;
3. clustering of the modules of the software architecture, in order to separate the needed work on different teams in an efficient way.

In the following, we will focus on the second problem.

Recalling some definitions

Do you remember what a graph is?

Definition

A *graph* is an ordered pair $G = (V, E)$ comprising a set V of vertices or nodes together with a set E of edges or links, which are 2-element subsets of V.

- **Undirected graph**: a graph in which edges have no orientation.
- **Directed graph** or **Digraph**: a graph $G = (V, A)$, where A is a set of *ordered* pairs of vertices, even called arcs or directed edges.
- **Weighted graph**: a graph in which numbers (weights) are assigned to each edge. It can be *directed* and *undirected*. It is denoted by $G = (V, E, w)$ or $G = (V, A, w)$, where w represents the weights.

Recalling some definitions

Do you remember what arc coloring is?

Let $G = (V, A)$ be a directed graph. A function

$$\mu : A \longrightarrow N,$$

that associates an integer number to each arc of G, is called arc coloring of G.

Since we can associate a color to each integer number, the function μ actually associates a color to each arc of G.

Subgraphs of the graph G can be located by considering all its arcs having the same color:

$$H = (U, F) \quad : \quad U \subseteq V,\, F \subseteq A, \quad \forall e, f \in F \;\; \mu(e) = \mu(f).$$

The densest subgraph problem

Given

- a digraph $G = (V, A)$,
- an arc coloring μ of G,
- a color k,

find the densest subgraph H of G in which the arcs have the same color k.

How can we solve this problem?

- we need to formulate a mathematical model,
- we can solve it by AMPL/CPLEX.

The mathematical model

Parameters

- V, set of vertices of G
- A, set of arcs of G
- μ, arc coloring of graph G
- k, a prefixed arc color

Variables

- x_v, binary, indicates if the vertex v is contained into the densest uniformly colored subgraph (U, F):

$$x_v = \begin{cases} 1 & \text{if } v \in U \\ 0 & \text{otherwise} \end{cases}$$

The mathematical model

Objective function

- The densest subgraph has the maximum number of arcs and the minimum number of vertices:

$$\max \left(\sum_{(u,v)\in A} x_u x_v - \sum_{v\in V} x_v \right)$$

Constraints

- There cannot be arcs having different colors:

$$\forall (u,v) \in A$$
$$x_u x_v \leq \min\left(\max(0, \mu(u,v) - k + 1), \max(0, k - \mu(u,v) + 1)\right)$$

Product of binary variables

We are going to solve this optimization problem by AMPL.

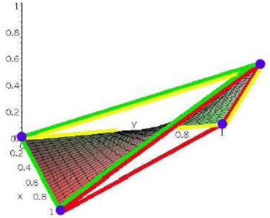
- We are going to use CPLEX for solving this problem.
- Observation: we have a product of binary variables in the objective function.
- Problem: CPLEX does not handle nonlinear terms.
- Solution: use a suitable reformulation for removing the product in the objective function.
- **How can we do that**?

Product of binary variables

From Leo Liberti's lectures:

Product of binary variables

- Consider binary variables x, y and a cost c to be added to the objective function only of $xy = 1$
- $\Rightarrow$ Add term cxy to objective
- Problem becomes mixed-integer (some variables are binary) and nonlinear
- Reformulate "xy" to MILP form (ProdBin reform.):



- replace xy by z
- add $z \leq y$, $z \leq x$
  $z \geq 0$, $z \geq x + y - 1$
- $x, y \in \{0, 1\} \Rightarrow$
  $z = xy$

The mathematical model

A new variable

- y_{uv}, real in $[0,1]$, it represents the product between x_u and x_v.

New objective function

- We substitute $x_u x_v$ with y_{uv}:

$$\max \left(\sum_{(u,v)\in A} y_{uv} - \sum_{v\in V} x_v \right)$$

École Polytechnique

The mathematical model

A new constraint

- We substitute $x_u x_v$ with y_{uv}:

$$\forall (u,v) \in A$$
$$y_{uv} \leq \min\left(\max(0, \mu(u,v) - k + 1), \max(0, k - \mu(u,v) + 1)\right)$$

Linearization constraints

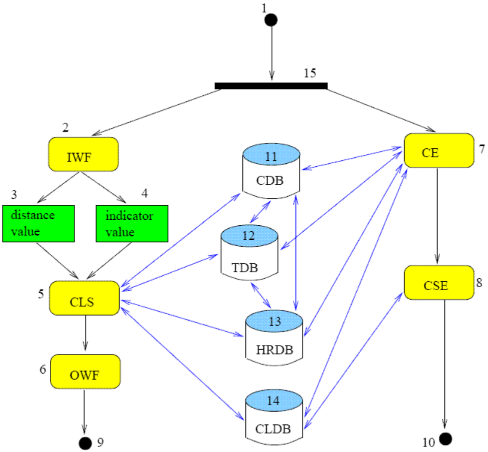
$$\forall (u,v) \in A,\ y_{uv} \leq x_u$$
$$\forall (u,v) \in A,\ y_{uv} \leq x_v$$
$$\forall (u,v) \in A,\ y_{uv} \geq x_u + x_v - 1$$

densestSubgraph.mod

```
# densestSubgraph.mod

param n >= 1, integer;
set V := 1..n;
set E within {V,V};
param c{E}; # edge weights
param I{E}; # edge inclusions

# arc colours
param kmax default 10;
param k <= kmax, >= 0, integer, default 1;
param mu{E} >=0, integer, <= kmax;

# variables
var x{V} binary;
var y{(u,v) in E} >= 0, <= min(max(0, mu[u,v]-k+1), max(0,k-mu[u,v]+1));

# model
maximize densesubgraph : sum{(u,v) in E} I[u,v] * c[u,v] * y[u,v] - sum{v in V} x[v];

# linearization constraints
subject to lin1 {(u,v) in E} : y[u,v] <= x[u];
subject to lin2 {(u,v) in E} : y[u,v] <= x[v];
subject to lin3 {(u,v) in E} : y[u,v] >= x[u] + x[v] - 1;
```

densestSubgraph.dat

We will use the structure of the graph
representing the architecture of the Proogle project.



densestSubgraph.dat

```
# densestSubgraph.dat

param n := 15;
param : E : c l mu :=
1 15 1 1 1
2 15 1 1 1
2 3 1 1 1
2 4 1 1 1
3 5 1 1 1
4 5 1 1 1
5 6 1 1 1
5 11 1 1 2
5 12 1 1 2
5 13 1 1 2
5 14 1 1 2
6 9 1 1 1
7 8 1 1 1
7 11 1 1 2
7 12 1 1 2
7 13 1 1 2
7 14 1 1 2
7 15 1 1 1
8 10 1 1 1
8 14 1 1 2
11 12 1 1 2
11 13 1 1 2
12 13 1 1 2
;
```

densestSubgraph.run

The .run file refers to the other two text files
(the model and the data),
and sets the solver to be used.

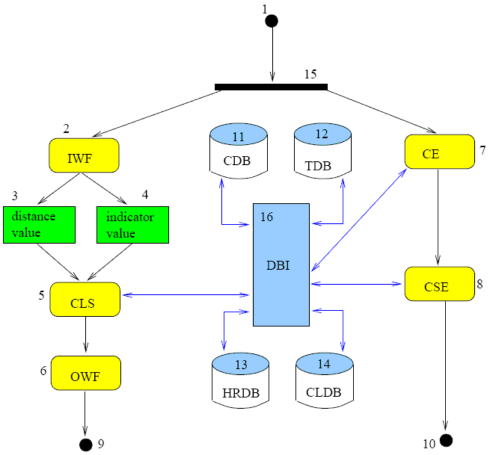
```
# densestSubgraph.run

model densestSubgraph.mod;
data densestSubgraph.dat;
let k := 2; # choose the color
option solver cplex;

solve;

display x;
```

Calling AMPL

```
[antonio.mucherino@ferrari ~]$ cat densestSubgraph.run | ampl
ILOG AMPL 10.100, licensed to "ecolepolytechnique-palaiseau".
AMPL Version 20060626 (Linux 2.6.9-5.ELsmp)
ILOG CPLEX 10.100, licensed to "ecolepolytechnique-palaiseau", options: e m b q use=8
CPLEX 10.1.0: optimal integer solution; objective 5
12 MIP simplex iterations
0 branch-and-bound nodes

x [*] :=
 1   0
 2   0
 3   0
 4   0
 5   1
 6   0
 7   1
 8   0
 9   0
10   0
11   1
12   1
13   1
14   1
15   0
;

[antonio.mucherino@ferrari ~]$
```

The interface

The solution of this optimization problem helps in improving the software architecture of the Proogle project. An interface is added.



The proposed exercises can be downloaded from:

www.antoniomucherino.it