Codes séparateurs et surveillance de trafic

Thomas Bellitto

Jeudi 5 novembre 2015

Doctorant au LaBRI
université de Bordeaux
Équipe LaBRI «Combinatoire et Algorithmique», thème «Graphes et Optimisation»
Équipe-projet INRIA «RealOpt»

1. Notions préliminaires
   - Théorie des langages
   - Codes séparateurs
   - Séparateurs et méthode de résolution
2. Surveillance de trafic
   - Le problème
   - Séparation sur un langage
3. Résolution
   - Cas fini
   - Un cas infini : l'identification totale
   - Restriction aux chemins sans demi-tour

Alphabet, mot, langage

Définitions de base

- Alphabet : ensemble fini non vide
- Mot : n-uplet d'éléments d'un alphabet (lettres)
- Langage : ensemble de mots sur un alphabet donné

Expression régulière

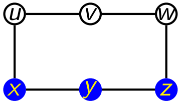
Langage rationnel

- $\varnothing$ est rationnel
- pour tout mot a, $\{a\}$ est rationnel
- pour tout langage rationnels L_1 et L_2, $L_1 + L_2 = \{u \in A^* \mid u \in L_1 \vee u \in L_2\}$ est rationnel
- pour tout langage rationnels L_1 et L_2, $L_1 L_2 = \{w \in A^* \mid \exists u \in L_1, \exists v \in L_2, w = uv\}$ est rationnel
- pour tout langage rationnel L, $L^* = \sum_{k \in \mathbb{N}} L^k$ est rationnel

Sur un graphe

Code séparateur sur un graphe : ensemble de sommets $\mathcal{C} \subset V$ tel que chaque sommet soit caractérisé par ses voisins (lui inclus) dans le code.

$$\forall v, v' \in V, \quad \underbrace{N[v] \cap \mathcal{C}}_{\text{signature du sommet } v} = N[v'] \cap \mathcal{C} \Rightarrow v = v'$$



$$N[u] \cap \mathcal{C} = \{x\} \qquad N[x] \cap \mathcal{C} = \{x, y\}$$
$$N[v] \cap \mathcal{C} = \{\} \qquad N[y] \cap \mathcal{C} = \{x, y, z\}$$
$$N[w] \cap \mathcal{C} = \{z\} \qquad N[z] \cap \mathcal{C} = \{y, z\}$$

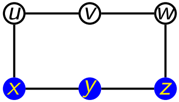
Problème : trouver un code séparateur de taille minimale

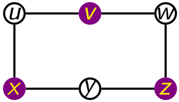
Codes identifiants

Ensemble dominant : ensemble tel que tout sommet du graphe ait un voisin dans le code

$$\forall v \in V, N[v] \cap \mathcal{C} \neq \varnothing$$

Code identifiant : code séparateur + ensemble dominant





Tests couvrants

Problème de test couvrant

Ensemble d'individus $\mathcal{I}$, ensemble d'attributs $\mathcal{A}$.
Recherche du plus petit ensemble $\mathcal{C} \subset \mathcal{A}$ tel que chaque individu de $\mathcal{I}$ soit caractérisé par les attributs de $\mathcal{C}$ qu'il possède.

Généralisation du problème précédent.

Très nombreuses applications : détection de motif, routage ou détection de faute dans des réseaux, bio-informatique (analyse de grandes molécules), médecine (identification de bactéries)...

Ensembles séparateurs

Ensembles séparateurs

L'ensemble séparateur de deux individus i et i' est l'ensemble des attributs qui les distinguent (différence symétrique de leurs attributs).

Un code $\mathcal{C}$ sépare i et i' ssi $\exists x \in \mathrm{Sep}(i, i'), x \in \mathcal{C}$

$$\begin{cases} \forall i \neq i' \in \mathcal{I}, \displaystyle\sum_{a \in \mathrm{Sep}(i,i')} x_a \geqslant 1 \\ \text{minimiser } \displaystyle\sum_{a \in \mathcal{A}} x_a \end{cases}$$

Présentation du problème

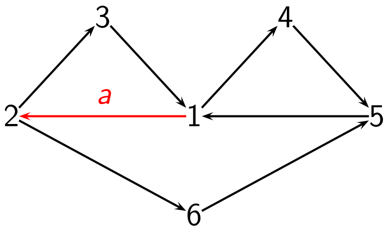
Réseau représenté par un graphe orienté. Possibilité de placer des capteurs sur les arcs.
Signature d'un chemin : séquence ordonnée des capteurs activés.

Surveillance de trafic

Un marcheur se déplace dans le réseau en suivant un chemin dans un ensemble donné de possibilités.
Problème : trouver le plus petit ensemble de capteurs tel que tous les chemins aient une signature différente

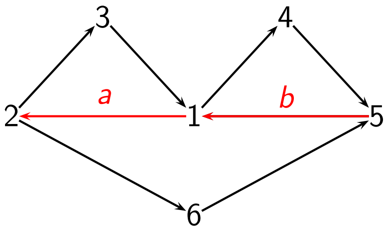
Principales difficultés

- L'ensemble des capteurs activés ne suffit pas
  Ex : $(1, 2, 3, 1)$ et $(1, 2, 3, 1, 2, 3, 1)$



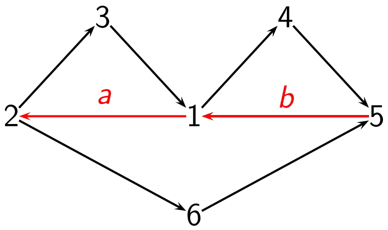
Principales difficultés

- L'ensemble des capteurs activés ne suffit pas
  Ex : $(1, 2, 3, 1)$ et $(1, 2, 3, 1, 2, 3, 1)$
- Le nombre de fois qu'ils ont été activés ne suffit pas non plus
  Ex : $(1, 2, 3, 1, 4, 5, 1)$ and $(1, 4, 5, 1, 2, 3, 1)$



Principales difficultés

- L'ensemble des capteurs activés ne suffit pas
  Ex : $(1, 2, 3, 1)$ et $(1, 2, 3, 1, 2, 3, 1)$
- Le nombre de fois qu'ils ont été activés ne suffit pas non plus
  Ex : $(1, 2, 3, 1, 4, 5, 1)$ and $(1, 4, 5, 1, 2, 3, 1)$
- Ensemble de chemins à distinguer potentiellement infini



Séparation sur un langage

Projection d'un mot

Projection d'un mot $u \in A^*$ sur un sous-alphabet $A' \subset A$: plus grand sous-mot de $u \subset A'^*$.

Ex : $p_{\{a,b\}}(abacacb) = abaab$

Séparation sur un langage

Recherche du plus petit sous-alphabet $A' \subset A$ tel que la projection sur A' soit injective sur le langage L donné.

Ex : $L = \{aabcc, acabc, baacb, cbaac\}$

Lien avec la surveillance de trafic

Chemin $\leftrightarrow$ mot sur $\underbrace{\text{l'ensemble des arcs du graphe}}_{\text{alphabet}}$

Signature d'un chemin $\leftrightarrow$
projection du mot sur $\underbrace{\text{l'ensemble des arcs munis d'un capteur}}_{\text{sous-alphabet}}$

1. Notions préliminaires
2. Surveillance de trafic
3. Résolution
   - Cas fini
   - Un cas infini : l'identification totale
   - Restriction aux chemins sans demi-tour

Séparateurs forts

Ensemble séparateur fort

Pour deux mots u et $v \in A^*$, $\mathrm{SEP}(u,v) \subset \mathcal{P}(A)$ est tel qu'un sous-alphabet $\mathcal{C}$ sépare u et v ssi $\exists x \in \mathrm{SEP}(u,v), x \subset \mathcal{C}$.

Théorème

Pour tous mots u et v, il existe un ensemble séparateur fort ne contenant que des parties de A de cardinal au plus deux.

Il s'agit de contenir un séparateur fort de chaque paire de mots.

Présentation du problème

- Graphe orienté $G = (V, A)$.
- Ensemble non vide $V_I \subset V$ de points de départ potentiels.
- Ensemble non vide $V_F \subset V$ de destinations potentielles.

Problème : séparer tous les chemins menant d'un sommet de V_I à un sommet de V_F.

Langage atteignable

un langage $L \subset A^*$ est dit atteignable s'il existe un graphe $G = (V, A)$, $V_I \subset V$ et $V_F \subset V$ tels que L soit l'ensemble des chemins menant d'un sommet de V_I à un sommet de V_F.

Théorème de réduction

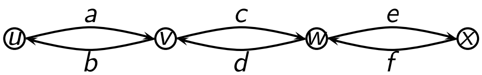
Restriction d'un langage rationnel

- $\overline{\varnothing} = \varnothing$.
- $\forall a \in A^*, \overline{\{a\}} = \{a\}$.
- pour tous langages rationnels L_1 et L_2, $\overline{L_1 + L_2} = \overline{L_1} + \overline{L_2}$.
- pour tous langages rationnels L_1 et L_2, $\overline{L_1 L_2} = \overline{L_1}\ \overline{L_2}$.
- pour tout langage rationnel L, $\overline{L^*} = \varepsilon + \overline{L} + \overline{L}^2$.

Théorème de réduction

Pour tout langage L atteignable sur un alphabet A, $\mathcal{C} \subset A$ sépare L ssi il sépare $\overline{L}$.

Limites du modèle précédent

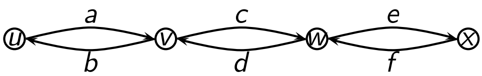


Ensemble des chemins possibles de u à x :

$$(a(c(ef)^*d)^*b)^*(ac(ef)^*d)^*c(ef)^*e$$

- Comportement irréaliste
- Augmente le temps de calcul, diminution donc la taille des instances qu'on peut résoudre
- Solution de mauvaise qualité

Présentation du nouveau problème



Nouveau problème : séparer tous les chemins menant d'un sommet de V_I à un sommet de V_F et qui ne font pas de demi-tour, (*i.e.* qui n'empruntent pas un arc (u, v) si le dernier arc emprunté est l'arc (v, u)).

Ensemble des chemins possibles de u à x :

$$ace$$

Résultats

Langage sans demi-tour

un langage $L \subset A^*$ est dit sans demi-tour s'il existe un graphe $G = (V, A)$, $V_I \subset V$ et $V_F \subset V$ tels que L soit l'ensemble des chemins sans demi-tour menant d'un sommet de V_I à un sommet de V_F.

Les langages sans demi-tour sont rationnels, on peut donc définir leur réduction !

Théorème de réduction

Pour tout langage L sans demi-tour sur un alphabet A, $\mathcal{C} \subset A$ sépare L ssi il sépare $\overline{L}$.

Conclusion

Contributions

- Reformulation du problème de surveillance de trafic dans un réseau comme problème de séparation, introduction d'un nouveau modèle sur les langages
- Mise au point d'outils pour pouvoir résoudre le problème sur des instances infinies intéressantes dans la pratique
- Implémentation de la méthode

Perspectives

Optimisation de l'algorithme.

Merci !