Optimisation par colonies de fourmis pour la recherche d'attaques

Introduction à la recherche en laboratoire

Aurélien Pepin

Ensimag — LIG

24/05/2018

Présentation

Contexte du travail

Recherche de scénarios malicieux

- Thème général : sécurité des systèmes (d'information).
  - Contrer les attaques d'initié ;
  - Identifier les scénarios d'utilisation malicieux.

- Outil 1 : modélisation d'un système d'information.
- Outil 2 : techniques de **model checking**.
  - Système $\equiv$ ensemble d'états d'exécution ;
  - Un état est caractérisé par des variables et des relations.
  - Identifier les états qui sont le résultat d'un scénario malicieux.

Contexte du travail

Recherche de scénarios malicieux

- Thème général : sécurité des systèmes (d'information).
  - Contrer les attaques d'initié ;
  - Identifier les scénarios d'utilisation malicieux.
- Outil 1 : modélisation d'un système d'information.
- Outil 2 : techniques de **model checking**.
  - Système $\equiv$ ensemble d'états d'exécution ;
  - Un état est caractérisé par des variables et des relations.
  - Identifier les états qui sont le résultat d'un scénario malicieux.
- Problème : explosion combinatoire de l'espace d'états

Cadre du sujet

Outils et environnement

- *Modélisation du système* : méthode B.
- *Model checker* : ProB (via l'API en Java).
- *Exemple de système étudié* : SI d'une banque.
  - Politique RBAC (Role-based access control) ;
  - Scénario : usurpation de compte par le banquier.

Cadre du sujet

Colonies de fourmis

- Ensemble d'algorithmes **heuristiques** inspirés des fourmis.
- Idée : reproduire le comportement collectif des fourmis.
  - Par exemple, la stigmergie pour la recherche de nourriture.

Cadre du sujet

Colonies de fourmis

- Ensemble d'algorithmes **heuristiques** inspirés des fourmis.
- Idée : reproduire le comportement collectif des fourmis.
  - Par exemple, la stigmergie pour la recherche de nourriture.

- Algorithme proposé : API (*Pachycondyla Apicalis*).

Cadre du sujet

Colonies de fourmis

- Algorithme proposé : API (*Pachycondyla Apicalis*).

Cadre du sujet

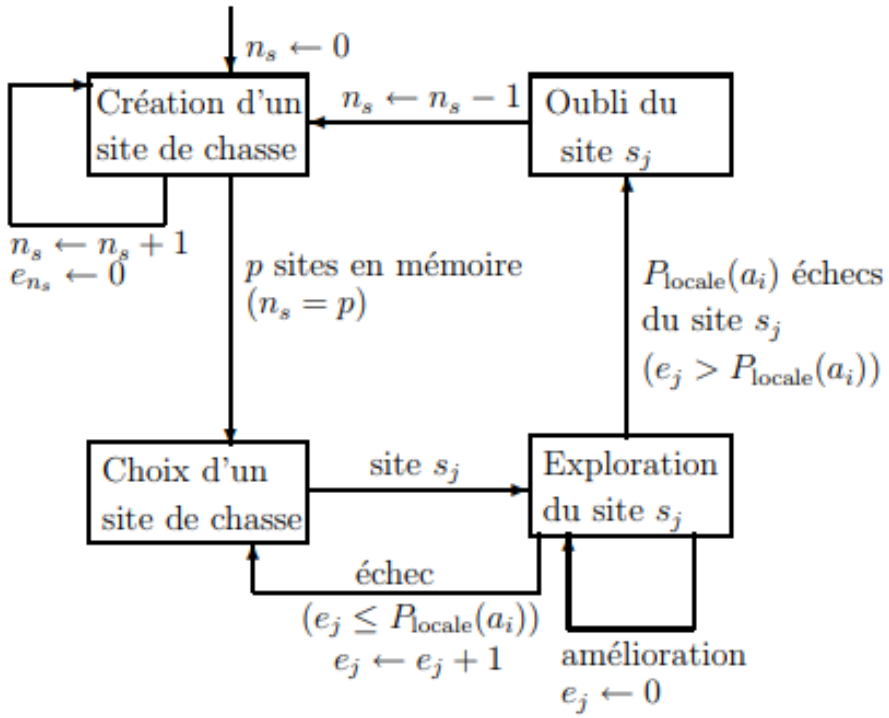
Colonies de fourmis

- Algorithme proposé : API (*Pachycondyla Apicalis*).
  - Au niveau global (colonie, nid) :
    - Envoyer des fourmis en exploration
    - Déplacer régulièrement le nid sur les sites intéressants

Cadre du sujet

Colonies de fourmis

- Algorithme proposé : API (*Pachycondyla Apicalis*).
  - Au niveau global (colonie, nid) :
    - Envoyer des fourmis en exploration
    - Déplacer régulièrement le nid sur les sites intéressants
  - Au niveau local (fourmi) :
    - Chasser de proche en proche, garder les *bons* sites



Réalisations

Paramètres de l'algorithme

- Algorithme générique : paramètres à adapter
  - Relatifs à l'algorithme :
    - Nombre de fourmis qui chassent ;
    - Patience d'une fourmi avant d'abandonner un site ;
    - Condition de déplacement du nid ;
    - etc.
  - Relatifs au domaine d'application (sécurité des SI) :
    - Heuristique d'exploration de l'espace de recherche ;
    - Fonction d'évaluation de ce qu'est un *bon* site ;
    - Condition d'arrêt de l'algorithme ;
    - etc.

Paramètres de l'algorithme

Relatifs au domaine d'application

Exploration de l'espace d'états

- Installation du nid à la racine de l'espace d'états
- Exploration aléatoire des transitions
- Ajout du retour en arrière pour la fourmi
  - Non prévu par ProB
  - Transformer l'espace d'états en arbre pour éviter les boucles

Paramètres de l'algorithme

Relatifs au domaine d'application

Fonction d'évaluation

- Est-ce qu'on se rapproche de l'état dangereux ?
- Prédicat (réponse $\in \{V, F\}$) trop « strict »

Paramètres de l'algorithme

Relatifs au domaine d'application

Fonction d'évaluation

- Est-ce qu'on se rapproche de l'état dangereux ?
- Prédicat (réponse $\in \{V, F\}$) trop « strict »

- Idée : indice de similarité, réponse $\in [0, 1]$
  - 0 : l'état possède toutes les variables recherchées
  - 1 : l'état n'a rien à voir
- Exemple pour une variable de classe :

$$\delta(A^*, A) = 1 - \frac{|A^* \cap A|}{|A^* \cup A|}$$

- Moyenne (pondérée) de la similarité de chaque variable

Paramètres de l'algorithme

Relatifs au domaine d'application

Condition d'arrêt

- Choix 1 : nombre d'itérations de l'algorithme
- Choix 2 : la meilleure solution n'évolue plus
- Choix 3 : nombre d'appels à la fonction d'évaluation
  - Appel à ProB coûteux (mémoïsation)
  - Comparer des exécutions avec différents paramètres

Résultats

Pour 20 exécutions de l'algorithme...

- Le scénario critique est recréé dans 95 % des cas ;
- Le chemin comporte 18 opérations en moyenne (optimal : 11) ;
- Le nombre moyens d'états à visiter est 247 ;
  - Recherche exhaustive : 36 000 états ;
  - Recherche guidée par prédicats : 12 895 états ;

Perspectives

- Supprimer les transitions inutiles
- Passage à l'échelle de l'algorithme ?

Merci pour votre attention

Bibliographie I

Nicolas Monmarché.
Algorithmes de fourmis artificielles : applications à la classification et à l'optimisation.
Université François Rabelais - Tours, 2000.

Maxime Dadoua.
Recherche d'attaques d'initié en systèmes d'information.
Rapport d'IRL - Ensimag et LIG, 2016.