

Total : 6 à 7 heures

## **Lecture 1 : Le pattern-matching (1 heure)**

Il s'agit de vous familiariser avec le contexte du **pattern-matching** (recherche de motifs en français) : historique, principaux algo, cas d'application.

Wikipedia n'est pas dans ce cas un bon point de départ. Utilisez plutôt des références académiques simples comme des diapos ou des notes de cours (exemples de sources sur chamilo).

Vous allez tomber assez vite sur les algorithmes suivants :

- **Brute force**
- **Knuth-Morris-Pratt (KMP)**
- **Boyer-Moore**
- **Rabin-Karp**
- **Aho-Corasick**

Il y a une littérature et des références (tutos, vidéos,...) très abondantes sur ces algorithmes.

Regardez tout de suite l'algo '**brute force**' pour rentrer dans le sujet et comprendre les enjeux algorithmiques du sujet.

## **Lecture 2 : Rabin-Karp (1,5 heure)**

Regardez en détail l'algo de **Rabin-Karp**. Il n'est pas trop difficile.

Remarques :

- Il y a une seule subtilité dans cet algo à observer, c'est le procédé de calcul des hash-successifs dans la chaîne principale, qui se fait en une seule opération.
- Si l'on change un peu le problème de base, en recherchant cette fois-ci un ensemble de motifs dans la chaîne : avec **Rabin-Karp** on peut à chaque étape non plus tester l'égalité des deux hashes mais tester l'appartenance du hash en cours de la chaîne principale à l'ensemble de hashes des motifs. Pour tester cette appartenance on peut utiliser un **filtre de bloom** (vu dans les cours/TD précédents), qui coutera  $O(1)$  (au prix d'une probabilité de faux négatif). Cette technique est intéressante quand il s'agit de détecter dans un flot de données un ensemble très grand de motifs : par exemple pour des sondes réseau.

## **Lecture 3 : Knuth-Morris-Pratt (1 heure)**

Regardez en détail l'algo de **Knuth-Morris-Pratt**. Attention, il n'est pas évident à comprendre. Prenez votre temps.

### **Exercice 1- *Knuth-Morris-Pratt* (1 heure)**

Il est indispensable de faire au moins un exercice soi-même pour comprendre. Prenez celui des diapos sur chamilo par exemple. Il y a aussi de bons tutos sur Youtube.

### **Lecture 4 (1 heure)**

Regardez rapidement les autres algorithmes : **Boyer-Moore**, **Shift-OR**, ..., pour compléter vos connaissances sur le sujet

### **Lecture 5 (facultative)- Cas d'application à une sonde d'intrusion (1 heure)**

Regardez en détail **Aho-Corasick**.

Puis lisez la référence (sur chamilo) sur l'utilisation de cet algo dans la sonde d'intrusion **SNORT**.

---