HAI916I

Module : IA pour génie logiciel

Rapport de TP2 sur Analyse formelle de concepts

Réalisé par :
M lle MEKHNACHE Toudherth – GL M2

Promotion 2023/2024

Table des matières

Chapitre 1

Analyse formelle des concepts

1.1 Introduction

L'analyse formelle des concepts (AFC) est une méthode puissante pour découvrir des relations cachées entre les entités et les attributs des données. Ce rapport décrit comment nous avons utilisé l'AFC pour créer un contexte formel à partir de données sur les logiciels antivirus et leurs caractéristiques. Nous avons construit un treillis de concepts pour explorer les liens entre les entités et les attributs, montrant ainsi comment cette méthodologie peut révéler des informations essentielles dans des ensembles de données complexes.

1.2 Création du Contexte Formel

1.2.1 Nature des Données

Dans le cadre de cette étude, nous avons créé un jeu de données sur les logiciels antivirus sous Windows et leurs caractéristiques. Dans un contexte où la sécurité informatique est d'une importance cruciale, l'analyse des logiciels antivirus offre une perspective essentielle pour la protection des systèmes et des données contre les menaces en ligne. Nous avons choisi de travailler avec des logiciels antivirus en raison de leur pertinence dans le monde numérique actuel, où la sécurité des systèmes est une préoccupation majeure.

1.2.2 Format du Contexte Formel

Pour créer notre contexte formel, nous avons suivi la structure typique de l'analyse formelle des concepts, qui consiste en un ensemble d'entités (les logiciels antivirus) et un ensemble de caractéristiques (les attributs ou fonctionnalités). Chaque logiciel antivirus est considéré comme une entité, tandis que les caractéristiques sont les attributs qui décrivent leur comportement.

Voici un extrait de notre contexte formel :

TABLE 1.1 – Caractéristiques des Logiciels Antivirus

Logiciel	On demand scan	On access scan	Boot time scans	Heuristics	Cloud AV	Fire wall	IDS	IPS	Sand box	Email Security	Anti Spam
Kaspersky	1	1	1	1	0	0	0	0	0	1	0...
Avast	1	1	1	1	1	1	0	1	1	0	1...
AVG	1	1	1	1	1	0	0	0	0	1	0...
Avira	1	1	1	1	1	0	0	0	0	1	1...
Bitdefender	1	1	1	1	1	0	0	0	0	0	1...
ZoneAlarm	1	1	0	1	0	1	0	0	0	0	0...
Immunet	1	1	0	0	1	0	0	0	0	1	0...
Clam	1	0	1	1	0	0	0	0	0	0	0...
G-DATA	1	1	1	1	0	0	0	0	0	0	0...
McAfee	1	1	1	1	0	0	0	0	0	1	0...
Windows-Defender	1	1	0	0	0	1	0	0	1	1	0...
Panda-Free	1	1	1	1	1	1	0	0	0	0	0...
Vba32-AntiVirus	1	1	0	0	0	0	0	0	0	0	0...

Chaque extention (logiciel antivirus) est associée à un ensemble de intentions (attributs) représentées par des valeurs binaires (1 pour la présence, 0 pour l'absence). Ce format nous permet d'appliquer l'analyse formelle des concepts.

Les différentes extensions et intentions de notre modèle de jeu de données sont les suivantes :

1.2.3 Extensions (Logiciels Antivirus)

Dans notre jeu de données, les entités utilisées sont les suivantes :

— Kaspersky
— Avast
— AVG
— Avira
— Bitdefender
— ZoneAlarm
— Immunet
— Clam
— G-DATA
— McAfee
— Windows-Defender
— Panda-Free
— Vba32-AntiVirus

1.2.4 Intentions (Caractéristiques)

Dans notre jeu de données sur les logiciels antivirus, les caractéristiques utilisées sont les suivantes :

- On-demand scan
- On-access scan
- Boot-time scans
- Heuristics
- CloudAV
- Firewall
- IDS (Intrusion Detection System)
- IPS (Intrusion Prevention System)
- Sandbox
- Email Security
- AntiSpam
- Web protection
- Macro protection
- Live Update
- Support
- Settings Import/Export

1.2.5 Objectif de l'Analyse

L'objectif de notre analyse est de découvrir des relations et des implications entre les logiciels antivirus et leurs caractéristiques. En utilisant l'analyse formelle des concepts, nous cherchons à identifier des schémas et des structures cachées dans nos données, ce qui peut fournir des informations utiles pour la prise de décision en matière de sécurité informatique et la compréhension des logiciels antivirus.

Nous allons également construire le treillis de concepts à partir de notre contexte formel pour explorer les relations d'implication entre les concepts, ce qui nous aidera à dégager des connaissances plus profondes.

1.3 Construction du Treillis de Concepts

1.3.1 Utilisation de FCA4J

Apres avoir préparer nos données sous forme d'un tableau dans un fichier au format CSV telque les lignes représentent les logiciels antivirus (extensions) et les colonnes représentent les caractéristiques (intentions), ce format nous a permis de représenter nos entités et attributs de manière structurée..

Nous avons importé nos données dans FCA4J en utilisant le fichier au format CSV.

1.3.2 Construction du Contexte Formel

Avec FCA4J, nous avons défini notre contexte formel en spécifiant les entités (produits) et les attributs (caractéristiques). Voici un extrait de notre contexte formel :

- **Les extentions (entités)**

- Kaspersky
- Avast
- AVG
- Avira
- Bitdefender
- ZoneAlarm
- Immunet
- Clam
- G-DATA
- McAfee
- Windows-Defender
- Panda-Free
- Vba32-AntiVirus

- **Les intentions (caractéristiques)**
  - On-demand scan
  - On-access scan
  - Boot-time scans
  - Heuristics
  - CloudAV
  - Firewall
  - IDS (Intrusion Detection System)
  - IPS (Intrusion Prevention System)
  - Sandbox
  - Email Security
  - AntiSpam
  - Web protection
  - Macro protection
  - Live Update
  - Support
  - Settings Import/Export

1.3.2.1 Construction du Treillis de Concepts :

En utilisant les données structurées dans notre contexte formel, FCA4J a généré le treillis de concepts. Ce treillis représente les relations d'implication entre les entités et les attributs. Voici un exemple simplifié :

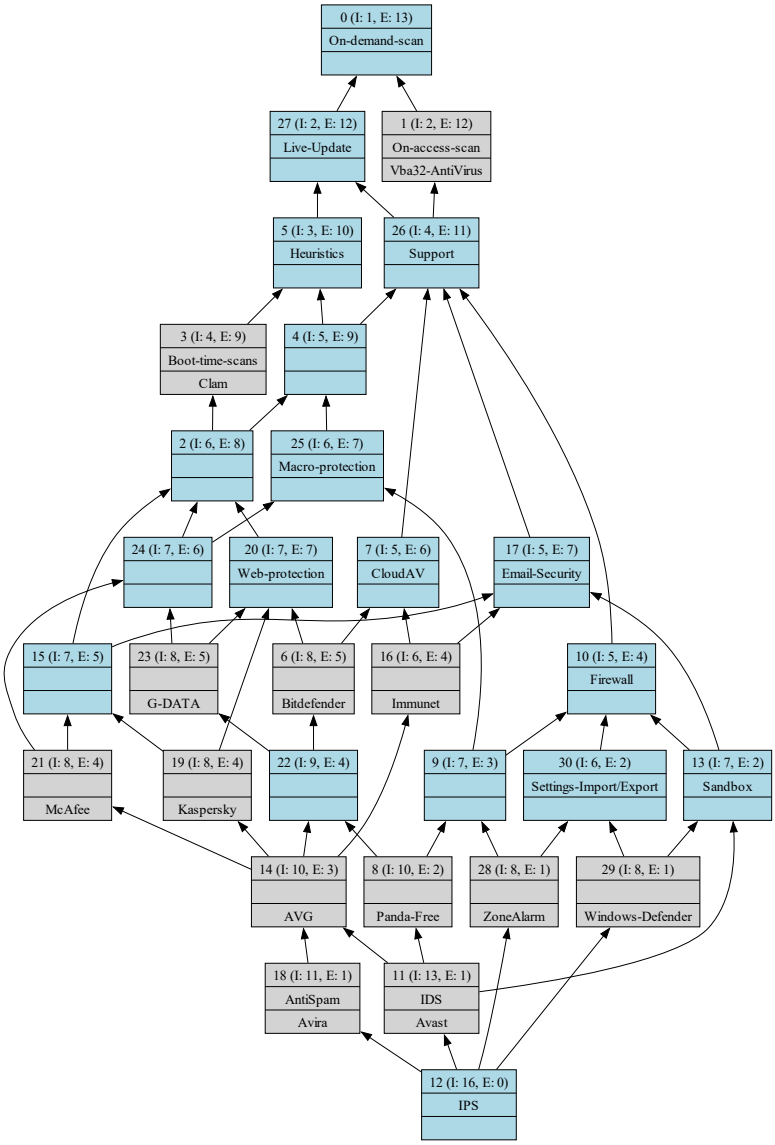


FIGURE 1.1 – Treillis des concepts partiel

1.4 Interprétation des résultats

Les concepts en analyse formelle des concepts sont les paires (ensemble d'entités, ensemble de caractéristiques) qui décrivent les relations entre les entités et les caractéristiques. Ils représentent des sous-ensembles des ensembles d'entités et de caractéristiques, formant ainsi une hiérarchie de concepts.

Dans notre analyse des logiciels antivirus, nous avons généré un treillis de concepts pour explorer les relations entre les logiciels et leurs caractéristiques. Voici quelques conclusions que nous pouvons tirer de cette analyse

— **Concept 1 :**

```
({Avira} , {On-demand-scan, On-access-scan, Boot-time-scans,
   Heuristics, CloudAV, Email-Security,AntiSpam, Web-
   protection, Macro-protection, Live-Update, Support})
```

— **Concept 2 :**

```
({ Kaspersky, Avast, AVG, Avira} , {On-demand-scan, On-access
   -scan, Boot-time-scans, Heuristics, Support, Settings-
   Import/Export})
```

— **Concept 3 :**

```
({ Avast, AVG, Avira } , {CloudAV, Firewall, Email-Security,
   AntiSpam, Web-protection, Macro-protection})
```

— **Concept 4 :**

```
({ Kaspersky } , {  Sandbox })
```

— **Concept 5 :**

```
({ ZoneAlarm } , {  On-access-scan, Firewall, Support,  Email
    Security, AntiSpam, Web protection, Macro protection,
   Live Update })
```

Ces concepts nous permettent de conclure ce qui suit :

Dans le Concept 2 et le Concept 5, Kaspersky, Avast, AVG et Avira partagent de nombreuses caractéristiques, tandis que ZoneAlarm se distingue par ses caractéristiques de pare-feu.

Dans le Concept 3 et le Concept 4, les extensions des deux concepts partagent également de nombreuses caractéristiques, tandis que Kaspersky se distingue par l'intention de Sandbox.

— **Concept 6 :**

```
({ Avast, AVG, Avira , Immunet} , { On-demand-scan, On-access
   -scan, CloudAV, Email-Security, Live-Update, Support})
```

— **Concept 7 :**

```
({ Avast, AVG, Avira , Panda-Free} , { On-demand-scan, On-
   access-scan, Boot-time-scans, Heuristics,CloudAV, Web-
   protection, Macro-protection, Live-Update, Support})
```

— **Concept 8 :**

```
({ Bitdefender } , {   On-demand-scan, On-access-scan, Boot-
   time-scans, Heuristics,CloudAV, Live-Update, Support, Web-
   protection })
```

Ce Concept 8 représente le logiciel antivirus Bitdefender, qui se distingue des autres logiciels par son ensemble spécifique de caractéristiques, notamment l'analyse à la demande, l'analyse lors de l'accès, l'analyse au démarrage, la détection heuristique, le support et la protection web.

— **Concept 9 :**

```
({ Avast, AVG, Avira ,Panda-Free, Kaspersky, Bitdefender, G-
   DATA, McAfee, Clam } , {  On-demand-scan, On-access-scan,
   Boot-time-scans, Heuristics})
```

— **Concept 10 :**

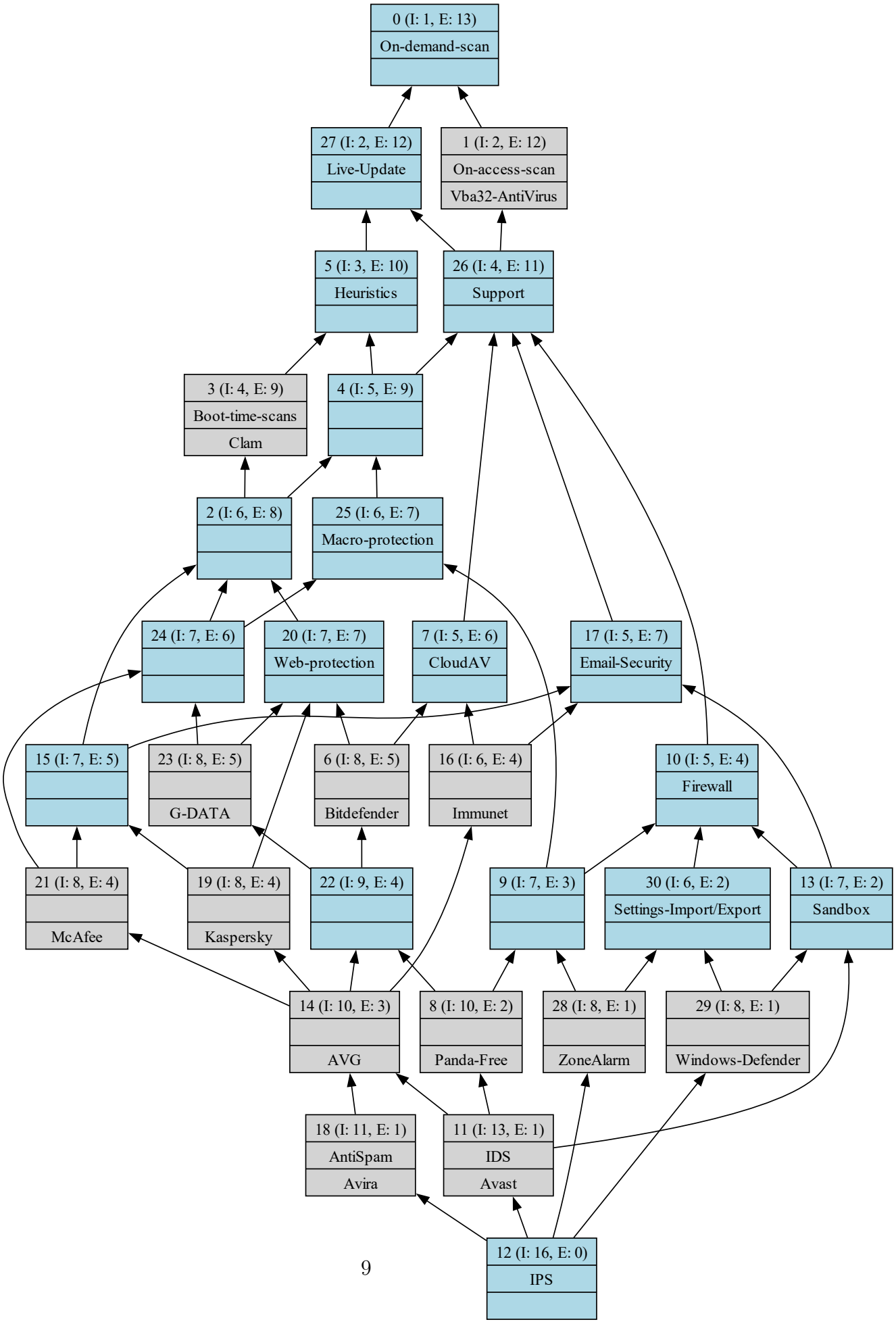
```
({ Avast, AVG, Avira ,Panda-Free, Kaspersky, Bitdefender, G-
   DATA, McAfee, Clam, ZoneAlarm, Immunet, Windows-Defender,
   Vba32-AntiVirus } , {  On-demand-scan, On-access-scan })
```

D'autres concepts (Concept 9 et Concept 10) incluent un ensemble d'entités partageant des caractéristiques similaires, reflétant ainsi des regroupements pertinents dans les données.

Cette interprétation nous permet de mieux comprendre les relations entre les logiciels antivirus et les caractéristiques qui les définissent, ce qui peut être utile pour la comparaison et la sélection de logiciels antivirus en fonction des caractéristiques souhaitées.

1.5 Conclusion

En conclusion, notre analyse formelle des logiciels antivirus à l'aide de la Formal Concept Analysis (FCA) et du logiciel FCA4J a permis de mettre en évidence les relations subtiles entre les entités (logiciels antivirus) et les caractéristiques qui les définissent. Cette approche nous a aidés à découvrir des structures sous-jacentes dans nos données brutes et à regrouper les logiciels antivirus en fonction de leurs caractéristiques communes.


0 (I: 1, E: 13)
On-demand-scan
27 (I: 2, E: 12)
Live-Update
1 (I: 2, E: 12)
On-access-scan
Vba32-AntiVirus
5 (I: 3, E: 10)
Heuristics
26 (I: 4, E: 11)
Support
3 (I: 4, E: 9)
Boot-time-scans
Clam
4 (I: 5, E: 9)
2 (I: 6, E: 8)
25 (I: 6, E: 7)
Macro-protection
24 (I: 7, E: 6)
20 (I: 7, E: 7)
Web-protection
7 (I: 5, E: 6)
CloudAV
17 (I: 5, E: 7)
Email-Security
15 (I: 7, E: 5)
23 (I: 8, E: 5)
G-DATA
6 (I: 8, E: 5)
Bitdefender
16 (I: 6, E: 4)
Immunet
10 (I: 5, E: 4)
Firewall
21 (I: 8, E: 4)
McAfee
19 (I: 8, E: 4)
Kaspersky
22 (I: 9, E: 4)
9 (I: 7, E: 3)
30 (I: 6, E: 2)
Settings-Import/Export
13 (I: 7, E: 2)
Sandbox
14 (I: 10, E: 3)
AVG
8 (I: 10, E: 2)
Panda-Free
28 (I: 8, E: 1)
ZoneAlarm
29 (I: 8, E: 1)
Windows-Defender
18 (I: 11, E: 1)
AntiSpam
Avira
11 (I: 13, E: 1)
IDS
Avast
12 (I: 16, E: 0)
IPS