

Compte-rendu Tp data mining

Demongeot Nicolas

Ougueur Karim

Traitement des données :

On a trois phases pour la mise en forme des données sous forme de data set compatible avec spmf :

-La première ou l'on génère un fichier "filename.txt" qui contient le nom de toutes les cartes différentes contenu dans le jeu de donnée fournis ce traitement est effectué par un premier Jar,

-Deuxième partie on génère un fichier "carteNumber.txt" qui va générer la première partie du fichier à passer à spmf (utile seulement si on se sert de l'interface Windows).

-Troisième phase : à partir du fichier "absolute_+.txt" et "filename.txt" on créer le fichier "fileLCM_players.txt" qui contient tous les item set. Les items set de taille sont supprimés.

Les fichiers "CarteNumber.txt" et "fileLCM_players.txt" son ensuite concaténer.

Nous avons créé un petit script Bash qui enchaîne ses les étapes précédentes et exécute spmf. Il y a deux possibilités d'exécutions suivant les données que l'on souhaite observer, premier paramètre la valeur de minsup, second paramètre (optionnel) le nom d'une carte.

- Lorsque l'on ne passe pas de paramètres on effectuera un appel de spmf sur l'intégralité des decks qui ont été identifié dans le dataset d'origine.
- Si l'on passe le nom d'une carte en paramètre on aura le résultat de l'analyse spmf sur le sous-ensemble de deck qui contient la carte passée en paramètre (en soustrayant à chaque fois le pouvoir héroïque à l'item set). Cela va nous permettre par exemple d'analyser les decks de telle ou telle classe en passant en paramètre le nom de son pouvoir héroïque. Mais aussi de connaître les sous ensemble de carte qui sont les plus souvent associées à celle passée en paramètre.

Post-Traitement :

- Nous effectuons un dernier traitement sur la sortie de spmf qui enlève tous les item sets de taille un.

Analyse des résultats :

-Un item set fréquent fermer ici vas nous permettre de potentiellement faire émerger des connaissances, par exemple on va pouvoir regarder si des associations de cartes correspondent à une certaine logique, par exemple des associations de cartes pouvant former des combos.

-Il y a aussi des Itemset fréquent fermer qui ne nous apprennent rien : les item set de taille 1 correspondant aux pouvoirs héroïques, ou les itemset contenant la pièce.

-En utilisant la commande `script_to_use.sh 40 Fireblast` on arrive à discerner les deux duos de cartes les plus utiliser pour le deck mage Frostbolt/WyrmrestAgent et Frosbolt/ Fireball

On arrive aussi à isoler les deux duos de cartes les plus utiliser pour le deck chasseur MadScientist/ EaglehornBow et MadScientist/ KillCommand.

On remarque que MadScientist et eagleHornbow sont des cartes qui combo bien l'une permet de poser une créature sur le board et donne un secret tandis qu'eaglehorn est une arme qui gagne de la durabilité quand un secret allié est révélé.