



Projet Data Mining

Pierre TURPIN, Jean-Marie COMETS

25 mars 2014

Table des matières

1	Caractérisation du flux vidéo	3
2	Prédiction de l'audience d'un flux	4
3	Classement des "meilleurs" joueurs	4

Au vu des nombreux problèmes de "scaling" que nous pouvions rencontrer avec le jeu de données prévu, l'intégralité de ce rapport repose sur l'analyse d'un échantillon fixé des données, soit 5 000 streams.

1 Caractérisation du flux vidéo

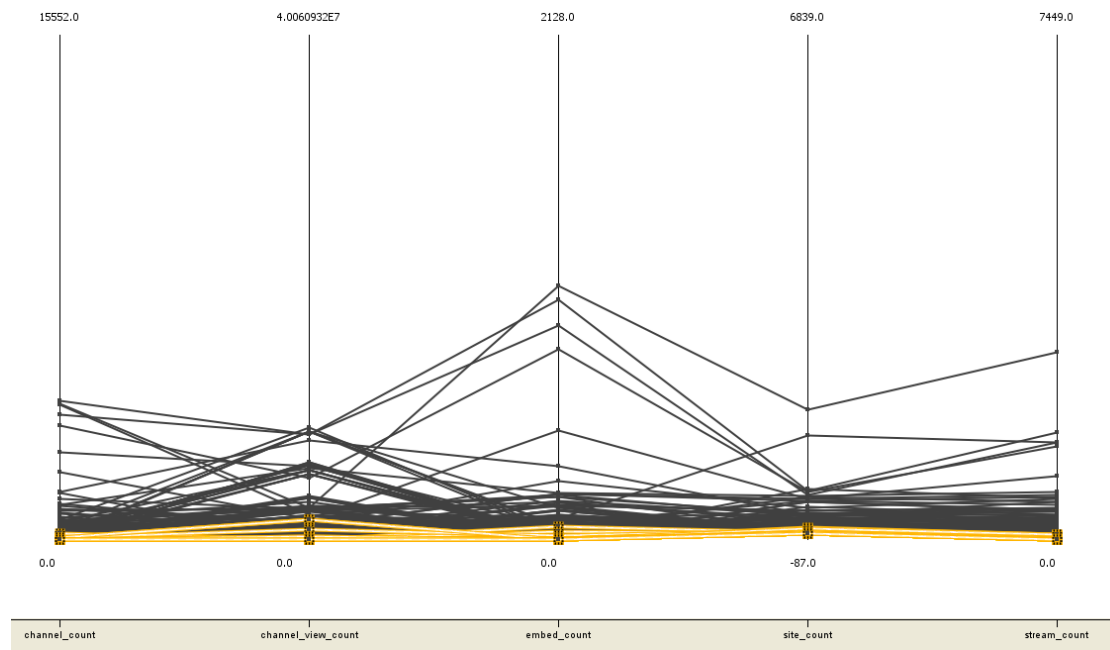


FIGURE 1 –

Dans le jeu de données utilisé, la plupart des jeux étaient en doublons car ils n'étaient pas tous orthographiés de la même façon (majuscules/minuscules, espaces, ...). Nous n'avons pas pu, à cause de la taille des données, corriger toutes les entrées afin d'unifier l'écriture des jeux. Les résultats ne sont donc pas complètement exacts mais donnent tout de même une bonne approximation de la réalité.

En tout il y a 160 jeux différents en comptant les doublons. Nous avons établi la part de vue des différents jeux en groupant sur le champ *meta-name* et en sommant les différents compteurs de vue *site-count*, *channel-view-count*, *channel-count*, *embedded-count*. Cette dernière somme représente une heuristique de popularité (*popularity*). Comme une grande quantité de jeu était très minoritaire selon notre heuristique, nous avons regroupé ces derniers (en seuillant la popularité) dans une seule catégorie *misc*.

La figure 5 représente la popularité de chaque jeu. L'ensemble des jeux *misc* forment 19% de popularité tandis que 10 autres jeux prennent les 80% restant. Il y a donc une très grande disparité dans les jeux vidéo et une petite minorité de 10 jeux écrasent totalement 150 autres jeux.

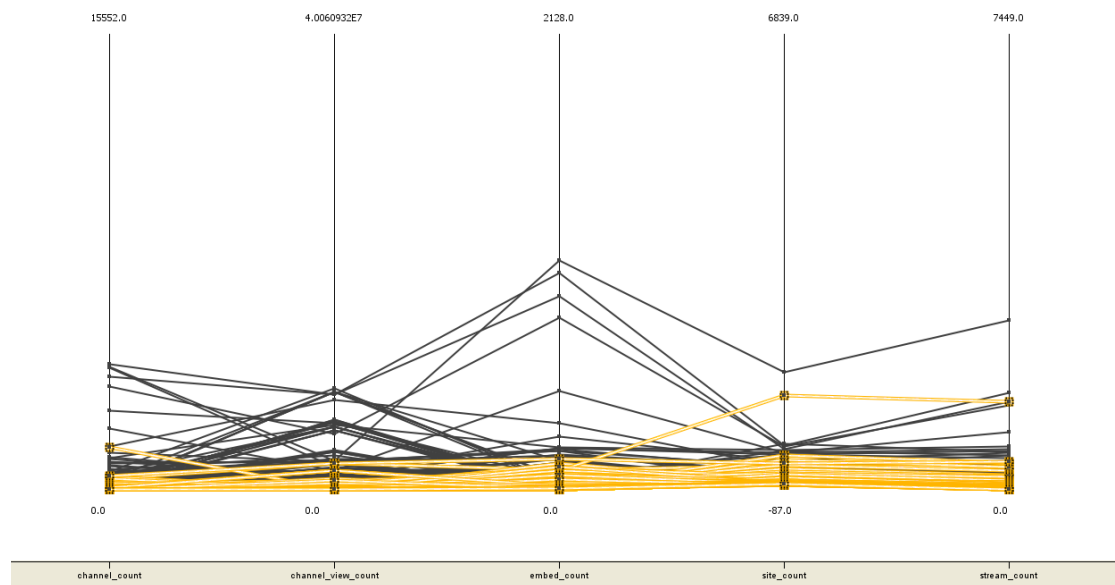


FIGURE 2 –

Le tableau 1 montre alors un classement des jeux les plus populaires sur la plateforme Twitch.

2 Prédiction de l'audience d'un flux

3 Classement des "meilleurs" joueurs

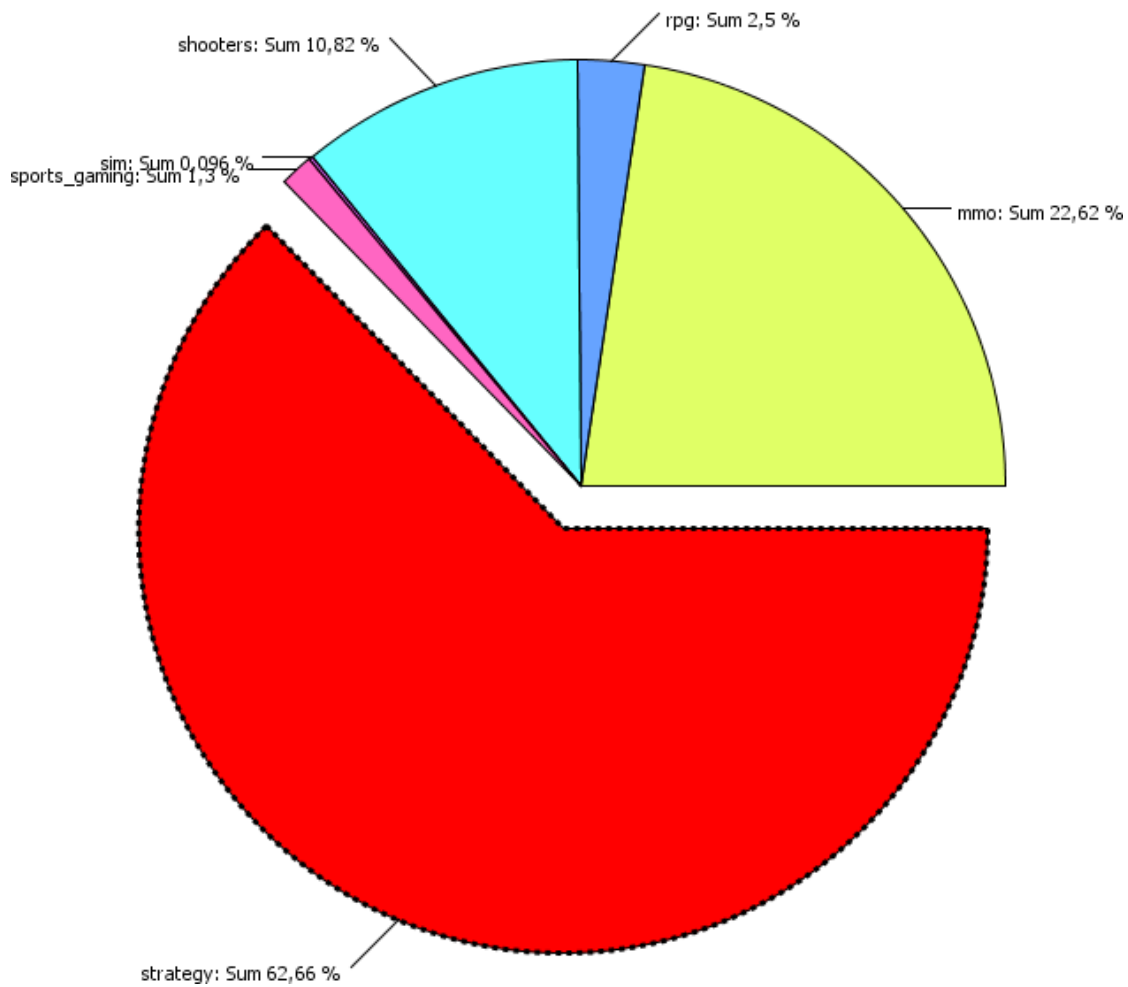


FIGURE 3 –

TABLE 1 – Classement des 10 jeux les plus populaires sur la plateforme Twitch.

Position	Jeux vidéo	Part de popularité (en %)
1	StarCraft II	25.12
2	World of Warcraft : Cataclysm	11.25
3	Super Street Fighter IV	10.18
4	League of Legends	8.98
5	Gears of War	7.46
6	Battlefield 3	5.87
7	Call of Duty : Black Ops	3.58
8	The Ico & Shadow of Colossus Collection	3.26
9	Halo : Reach	3.23
10	Heroes of Newerth	2.21

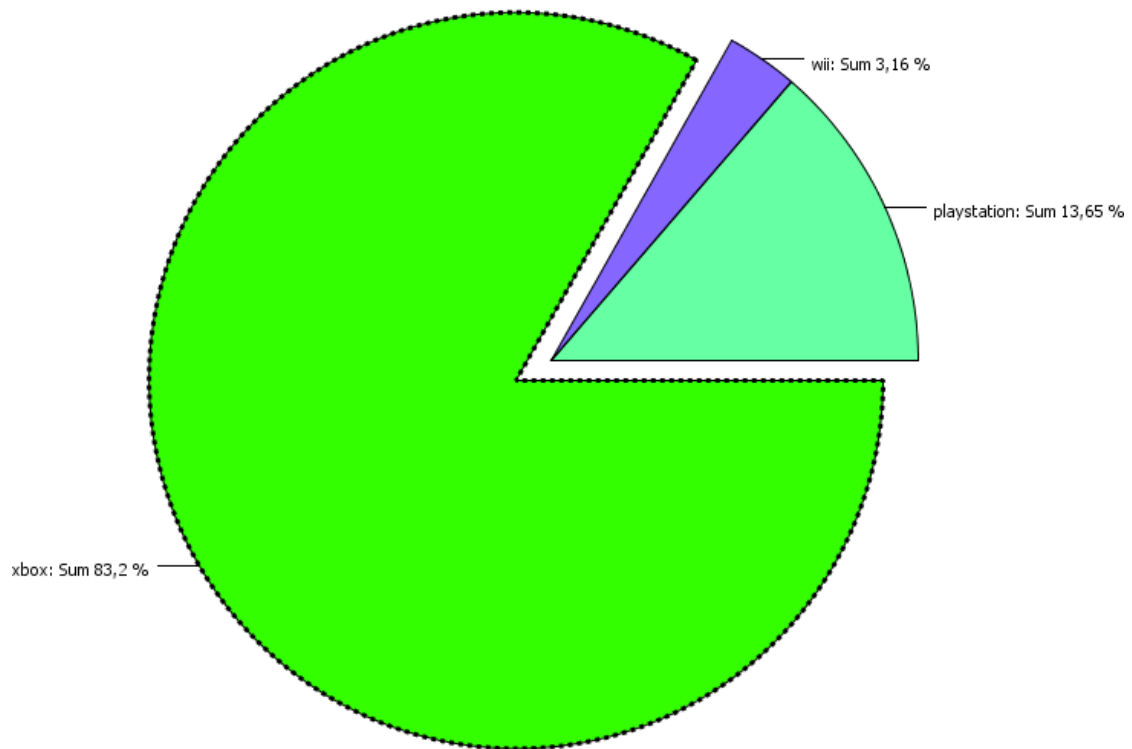


FIGURE 4 –

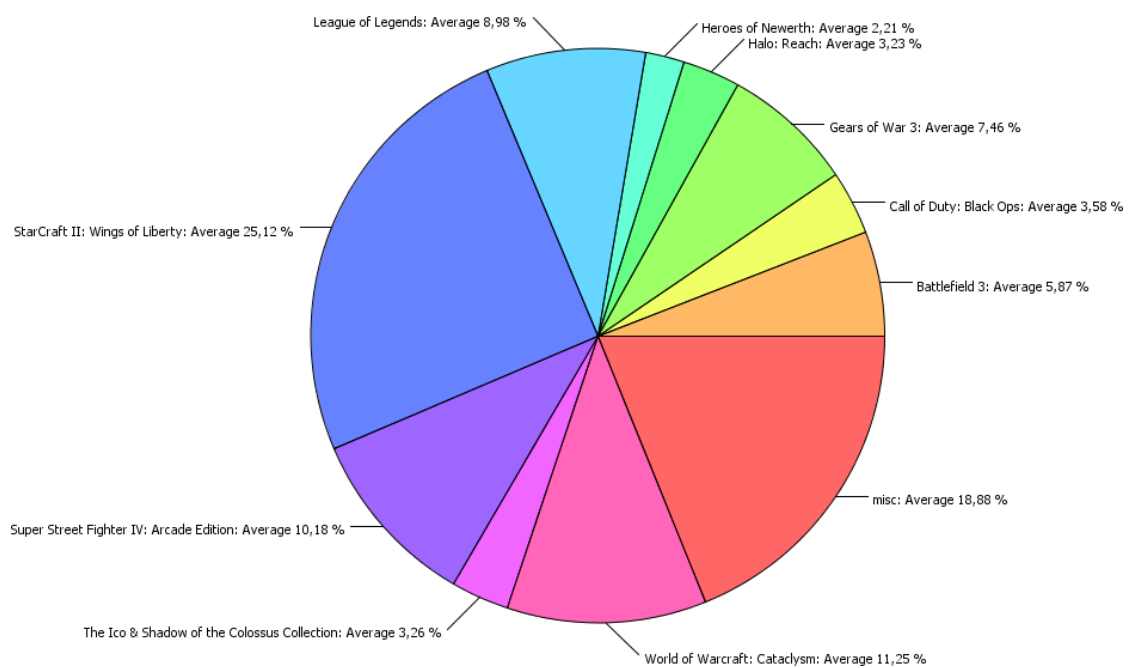


FIGURE 5 – Part de vue en moyenne des jeux vidéo

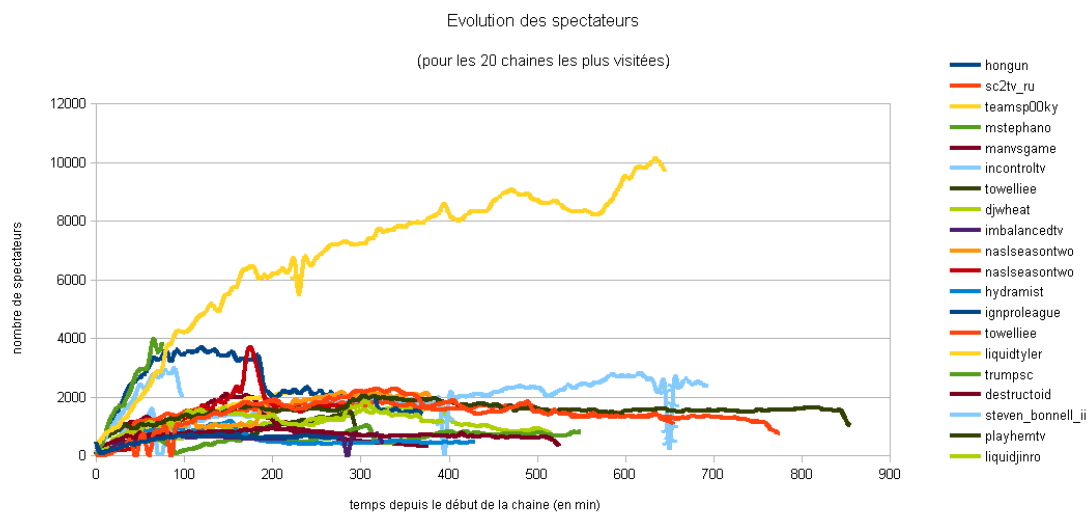


FIGURE 6 –