

# Rapport hebdomadaire de stage

---

Alexis Blanchet

26 avril 2018

## 1 INTRODUCTION

Cette semaine a été divisée en trois axes principaux : tout d'abord refaire les mécanismes de détection d'anomalies afin d'améliorer encore la détection des potentielles coupures/fin de programme. Ensuite traiter les fichiers .xml fournis par mediametrie contenant l'ensemble des données relatives aux programmes TV. Et enfin poursuivre la recherche des pratiques commerciales liées à la pub et plus particulièrement la juridiction et la jurisprudence la concernant. On s'appuiera pour cela sur le CSA et ses différents jugements.

## 2 LÉGISLATION

Le décret du 27 mars 1992 définit la publicité comme « toute forme de message télévisé diffusé contre rémunération ou autre contrepartie en vue soit de promouvoir la fourniture de biens ou services, y compris ceux qui sont présentés sous leur appellation générique, dans le cadre d'une activité commerciale, industrielle, artisanale ou de profession libérale, soit d'assurer la promotion commerciale d'une entreprise publique ou privée ». Une période d'au moins vingt minutes doit s'écouler entre deux interruptions successives à l'intérieur d'une même émission, et ce qu'il s'agisse d'une œuvre cinématographique, d'une œuvre audiovisuelle (constitutive, ou non, d'une série, d'un feuilleton ou d'un documentaire) ou d'un programme destiné à la jeunesse.

Les œuvres cinématographiques et audiovisuelles ne peuvent pas faire l'objet de plus de deux interruptions publicitaires. Et concernant les œuvres cinématographiques, elles doivent se limiter à une durée de six minutes au total.

Toutefois, la diffusion d'une œuvre audiovisuelle ou cinématographique par France Télévisions, et celle d'une œuvre cinématographique par les services de télévision de cinéma ne

peut faire l'objet d'aucune interruption publicitaire. Il en va différemment lorsqu'une émission est composée de parties autonomes, et pour une émission sportive ou retransmettant un événement ou un spectacle comprenant des intervalles. Dans ces cas, les messages publicitaires peuvent être insérés entre ces parties autonomes ou dans ces intervalles sans limitation du nombre d'interruptions. Sur les chaînes privées, le temps d'antenne consacré à la publicité est encadré différemment selon leur mode diffusion :

1. sur les chaînes diffusées par voie hertzienne terrestre (c'est-à-dire la TNT), il est limité à neuf minutes par heure en moyenne quotidienne sur l'ensemble des périodes de programmation au cours desquelles cette diffusion est autorisée, et à douze minutes pour une heure d'horloge donnée. Afin de favoriser leur essor, les nouvelles chaînes de la TNT bénéficient de règles allégées pendant un délai de sept ans à compter de la date du début des émissions, le temps consacré à la publicité étant seulement limité par le plafond de douze minutes par heure d'horloge donnée. A l'issue de ce délai, elles devront également respecter la durée de neuf minutes par heure en moyenne quotidienne.
2. Sur les chaînes distribuées par câble, par ADSL ou diffusées par satellite, la durée consacrée à la publicité est fixée par voie conventionnelle avec le CSA. Elle ne peut excéder douze minutes pour une heure d'horloge donnée

Les règles encadrant la durée des messages publicitaires sont plus strictes sur les chaînes publiques : cette durée ne peut dépasser six minutes par heure en moyenne quotidienne, ni huit minutes pour une heure d'horloge donnée.

De plus, depuis le 5 janvier 2009, les chaînes de France Télévisions (France 3 Régions exceptées) ne doivent plus diffuser de publicité de marques de 20 heures à 6 heures du matin. Cette interdiction ne s'applique qu'à la publicité et ne concerne donc pas les messages d'intérêt général, les publicités génériques (pour faire la promotion de la pomme, des produits laitiers, etc.) ou les parrainages, qui peuvent continuer à être diffusés

le CSA peut de plus autoriser le dépassement du nombre d'écrans publicitaire à l'occasion d'événements particuliers. Par exemple : "Le CSA a autorisé la société France Télévisions à insérer, au sein des émissions consacrées au Téléthon diffusées en soirée le vendredi 4 décembre 2015 sur France 3 et le samedi 5 décembre 2015 sur France 2, deux écrans publicitaires dont les recettes seront intégralement reversées à l'Association française contre les myopathies, dès lors qu'il s'agit d'une interruption exceptionnelle liée à une opération caritative."

### 3 MODIFICATION DES MÉCANISMES DE DÉTECTION D'ANOMALIES

Afin de créer un historique il s'agit de détecter manuellement les anomalies et de les labelliser directement à la main en fonction des programmes TV et pourquoi pas des notes pris directement en regardant la TV. J'ai réécrit les mécanismes d'auto-annotation pour permettre une distinction entre gain anormal de téléspectateurs et perte anormale de téléspectateurs. De plus la normalisation des outliers permet aux données permettant la détection d'anomalies d'être moins sujettes à des pertes très fortes ou à des gains très forts qui écrasent le reste de la distribution des variations. Divers paramètres contribuent avec des poids distincts à la recherche des anomalies. L'attribution des poids n'est pas automatique et est faite à la main.

S'en suit un seuillage sur la probabilité d'être une anomalie et une classification selon les variations locales. De plus l'algorithme ayant pour objectif d'être déployé en Temps réel, aucun futur n'est utilisé dans la classification des événements. Cela engendre parfois un léger retard (1 2 minutes) dans la détection des anomalies (surtout si deux anomalies opposées se suivent).

## 4 ÉTUDE DES DONNÉES DES PROGRAMMES TV

les programmes TV fournis par médiamétrie sont des fichiers .xml contenant tous les programmes de la journée pour chaque chaîne avec un nombre important d'information concernant la diffusion du programme et le programme en lui même. Il existe cependant une disparité important dans les informations disponibles pour chaque programme. En effet le fichier xml autorise l'ajout de données entre balise pour chaque événement. Il y a de plus des informations qui se recoupent entre la diffusion et le programme (mais le programme et encapsulé DANS la diffusion). Ainsi les programmes TV présentent une quantité importante d'information que j'ai pour le moment réduit a l'essentiel afin d'avancer et de pouvoir revenir dessus ensuite quand je saurai quelle information je souhaite récupérer de ces données. les informations sont extraites directement des fichiers .xml et sauvegardées sous forme de DataFrame dans des CSV.

## 5 VISUALISATION DE DONNÉES

comme la semaine dernière, voici quelques graphes permettant de visualisé la détection d'anomalies sur les courbes d'audience. On ne travaillera que sur TF1 pour le moment afin de ne pas se perdre. De plus TF1 a un replay en libre accès et donc facilite grandement la vérification (plus ou moins précise) des événements détectés au cours de notre analyse.

Out[5]:

	PT2	DOLBY51	CLE-GENRE	TITRE	DOLBY	@DATEMODIF	DATEHEURE	CLE-FORMAT	GENRE	STEREO	...	description programme	CS	HD	PT3	INEDIT_EN_CLAIR	@CLEEM	RATIO	DATE	HEURE	minutes	
0	0	0	403	TFou	0	2017-11-08T02:12:33	2017-12-01T06:25:00	12	OrderedDict({'@CLE': '403'}, {'@GENRESIMPLE': ...})	1	...	Magazine jeunesse	NaN	1	0		1	122835453	16:9	2017-12-01	06:25:00	385
1	0	0	951	Météo	0	2017-11-07T12:08:12	2017-12-01T08:25:00	16	OrderedDict({'@CLE': '951'}, {'@GENRESIMPLE': ...})	1	...	Météo	NaN	1	0		0	122842665	16:9	2017-12-01	08:25:00	505
2	0	0	425	Téléshopping	0	2017-11-07T12:08:12	2017-12-01T08:30:00	12	OrderedDict({'@CLE': '425'}, {'@GENRESIMPLE': ...})	1	...	Magazine de télé-achat	NaN	1	0		0	122842666	16:9	2017-12-01	08:30:00	510
3	0	0	951	Météo	0	2017-11-07T12:08:12	2017-12-01T09:15:00	16	OrderedDict({'@CLE': '951'}, {'@GENRESIMPLE': ...})	1	...	Météo	NaN	1	0		0	122842667	16:9	2017-12-01	09:15:00	555
4	0	0	1026	Petits secrets entre voisins	0	2017-11-07T12:08:12	2017-12-01T09:20:00	3	OrderedDict({'@CLE': '1026'}, {'@GENRESIMPLE': ...})	1	...	Série réaliste	NaN	1	0		0	116093509	16:9	2017-12-01	09:20:00	560

5 rows x 39 columns

