

# Prédiction de séries temporelles : application à la structuration de flux audiovisuels

Jean-Philippe Poli<sup>1,2</sup>

Jean Carrive<sup>1</sup>

<sup>1</sup> Equipe Description des Contenus Audiovisuels

Direction de la Recherche et de l'Expérimentation

Institut National de l'Audiovisuel, 4 avenue de l'Europe, 94366 Bry-sur-Marne cedex - France

{jppoli,jcarrive}@ina.fr

<sup>2</sup> LSIS - UMR CNRS 6168

Universite Paul Cézanne (Aix-Marseille 3)

Domaine Universitaire de Saint-Jérôme

Avenue Escadrille Normandie-Niemen, 13397 MARSEILLE CEDEX 20 - France

{jean-philippe.poli}@lsis.org

## Résumé

*Les séries temporelles sont abondamment employées dans des domaines très variés. De nombreux chercheurs se sont intéressés à caractériser et à prédire des séries temporelles régulières, c'est-à-dire dont les valeurs sont mesurées à intervalles temporels réguliers. Nous proposons de considérer la grille de programme d'une chaîne de télévision comme une série temporelle dont les dates sont les heures de début d'un programme et les valeurs en sont le genre. Une prédiction d'une telle série temporelle nous permet à la fois de connaître la succession des genres d'émissions mais aussi d'estimer l'heure de début d'un programme afin de rechercher localement dans le signal la rupture avec l'émission précédente. Nous présentons dans cet article un moyen de prédire une telle série temporelle compte tenu d'une difficulté jamais rencontrée dans la littérature : la prédiction à la fois de la date et de la valeur de la prochaine observation de la série temporelle.*

## Mots clefs

Série temporelle, prédiction, grille de programmes, arbre de régression, modèle de Markov, structuration automatique de flux audiovisuel.

## 1 Introduction

La plupart des chaînes de télévision diffusent à présent leurs programmes en continu. La conservation de leurs diffusions, présentant autant d'intérêts sociologiques qu'éducatifs ou même ludiques, n'est plus un problème sur le plan technique grâce aux coûts abordables de leur captation et de leur sauvegarde. La difficulté réelle, à laquelle est confronté l'Institut National de l'Audiovisuel, réside

dans la mise à disposition de ces archives dont la vitesse de croissance ne fait qu'augmenter avec la multiplication des chaînes et des moyens de diffusion. En effet, les flux télévisuels nécessitent d'être décrits afin de permettre une recherche ultérieure. Afin d'accomplir une telle tâche, il est souhaitable de décomposer ces flux en plus petites unités. L'unité de la programmation d'une chaîne de télévision étant l'émission, nous nous proposons de décomposer automatiquement un flux télévisuel en émissions afin que chacune d'elles soit documentée.

Si les chercheurs en indexation automatique se penchent activement sur des problèmes tels que la reconnaissance des types d'émissions, la structuration et le résumé automatiques de vidéos[1], ils ne se sont que peu intéressés, à notre connaissance, à la structuration de flux. Pourtant, les grilles prévisionnelles publiées chaque semaines dans les magazines ou par voie électronique (EPG) représentent une source importante d'information sur le flux et les programmes qui le composent. En effet, les programmes mentionnés dans ces grilles prévisionnelles sont typés, résumés et la plupart du temps annotés. Idéalement, la structuration d'un flux consisterait à l'aligner sur la grille prévisionnelle. Malheureusement, une grille prévisionnelle est incomplète puisque que les émissions courtes (météo, trafic routier, magazines courts sur la vie quotidienne...) n'y figurent pas. Notre approche consiste à augmenter statistiquement la complétude des grilles de programmes prévisionnelles. Cependant, dans ce papier, nous nous intéressons à un modèle de prédiction du contenu d'un flux, indépendamment de sa grille prévisionnelle. Pour cela, nous considérons la grille de programmes d'une chaîne comme une série temporelle dont chaque temps serait la date de diffusion (i.e. le jour

de la semaine et l'heure de diffusion) et chaque valeur serait un type de programmes. Compte tenu du fait que les programmes ont des durées différentes, l'espacement temporel entre deux valeurs ne sera pas régulier. La difficulté de l'approche consiste donc à la prédiction d'une série temporelle irrégulière pour laquelle il faut prédire à la fois les dates et les valeurs.

Nous nous proposons d'estimer un intervalle de durées pour chaque émission, à l'aide d'un arbre de décision, afin de prédire la date de la prochaine valeur, et d'utiliser un modèle Markovien contextualisé afin de prédire les valeurs. Nous exposerons à la fin les résultats que nous avons obtenus.

## 2 Séries temporelles

**Définition 1** (*Série temporelle*) On appelle série temporelle la suite d'observations  $(o_t, t \in \mathbb{T})$  d'une variable  $O$  à différentes dates  $t$ .

**Remarque 1** Dans cet article, nous qualifions de régulières les séries temporelles dont les différentes dates sont régulièrement espacées, et d'irrégulières celles qui n'ont pas un taux d'échantillonnage constant.

Les séries temporelles régulières sont extrêmement étudiées en économétrie[2], en astronomie[3] et dans bien d'autres domaines[4]. Beaucoup d'outils permettent de caractériser leurs évolutions et de prédire des valeurs en fonction de l'historique de la série temporelle. Ces différentes communautés ne se sont que très peu intéressées aux séries temporelles irrégulières. Pourtant, la prise en compte de telles séries permettrait d'étudier des données dont la collecte est événementielle. [5] propose des modèles statistiques pour étudier des séries temporelles irrégulières à valeurs continues en se préoccupant de leur caractère stationnaire. Puisque l'hypothèse de travail de l'auteur est la possibilité de décomposer la série en une fonction déterministe et un bruit aléatoire, ces modèles ne peuvent pas être utilisés dans notre travail. En effet, nos séries temporelles vont représenter des grilles de programmes de chaînes de télévision. Les valeurs seront donc symboliques puisqu'elles représenteront les différents types d'émissions. La difficulté de la prédiction de telles séries réside dans le fait qu'il faut prédire d'une part la date de la prochaine observation, et d'autre part sa valeur.

Compte-tenu de la difficulté de la prédiction de la date d'un programme, nous faisons l'hypothèse que la date d'initialisation du système  $t_0$  coïncide avec le début d'une émission. Les autres dates  $t_{i+1}$  de la série temporelle seront calculées à partir de  $t_i$  en y ajoutant la durée de l'émission. Nous présentons dans les sections suivantes la modélisation d'une série temporelle irrégulière sans perdre de vue que notre but est l'application à la structuration des flux.

## 3 Prédiction des valeurs par un modèle de Markov

Dans cette section, nous nous intéressons à la modélisation d'une série temporelle par un modèle de Markov intervenant dans la prédiction des valeurs.

### 3.1 Modèles semi-Markovien contextuels

Lorsqu'il s'agit de prédire les valeurs d'une série temporelle, il est nécessaire de choisir un modèle statistique qui permettra de calculer une valeur en fonction de l'historique des observations. Certains auteurs se sont intéressés aux réseaux de neurones[6], aux SVM (machines à vecteurs supports)[7] ou encore aux modèles de Markov[8]. Notre choix s'est porté sur ces derniers qui sont très utilisés pour représenter des processus stochastiques séquentiels, malgré leur inadaptation à la prise en compte de durées. De plus, nous allons utiliser ce modèle non pas pour estimer la probabilité d'une séquence d'observations mais plutôt pour obtenir de lui les séquences les plus probables que nous considérons comme des prédictions. Nous avons dû introduire une extension des modèles de Markov afin de limiter les possibilités de transition d'un état à l'autre et éviter d'avoir trop de séquences d'observations possibles mais improbables. Pour cela, nous avons contextualisé les transitions entre états.

**Définition 2** (*Contexte*) Un contexte  $\theta$  est un ensemble de variables  $x_1, \dots, x_n$  à valeurs respectivement dans les domaines discrets ou continus  $\{D_1, \dots, D_n\}$ . Une instance  $\theta_i$  de ce contexte correspond à l'affectation d'une valeur à chacune de ses variables :

$$\forall i \in \{1, \dots, n\}, x_i = v_i \text{ avec } v_i \in D_i.$$

Dans cet article, nous appellerons *contexte* l'instance d'un contexte sans perte de généralité.

**Exemple 1** Dans notre cas, un contexte  $\theta$  représentera le contexte de diffusion, avec une variable *Heure* qui représentera l'heure de diffusion par un entier entre 0 et 86399, et une variable *Jour* qui représentera le jour de la semaine par un entier entre 0 et 6.

$$\theta = \{\text{Heure}, \text{Jour}\} \\ \text{et } D_{\text{Heure}} = \{0, \dots, 86399\}, D_{\text{Jour}} = \{0, \dots, 7\}.$$

**Définition 3** (*Fonction d'évolution*) Soit  $\Theta$  l'ensemble de toutes les instances possibles d'un contexte  $\theta$ . On appelle fonction d'évolution la fonction  $F$  définie par :

$$F : \Theta \times D_{p_1} \times \dots \times D_{p_m} \rightarrow \Theta \\ \theta_i, p_1, \dots, p_n \rightarrow \theta_{i+1}$$

où  $D_{p_i}$  est le domaine de valeurs du paramètre extérieur  $p_i$ .

**Exemple 2** En reprenant l'exemple 1, nous pouvons définir  $F$  comme étant la fonction qui à  $\theta_i$ , qui représente le début de l'émission en cours, ajoute sa durée pour obtenir  $\theta_{i+1}$ . Soit  $\mathbb{D}$  l'ensemble des durées possibles, et  $L$  une durée particulière. Alors dans notre exemple,  $F$  peut être définie par :

$$F : \Theta \times \mathbb{D} \rightarrow \Theta$$

$$\left\{ \begin{array}{l} \text{Heure} = h \\ \text{Jour} = j \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \text{Heure} = (L + H) \bmod 86400 \\ \text{Jour} = (j + \lfloor \frac{L+H}{86400} \rfloor) \bmod 7 \end{array} \right\}$$

**Définition 4** (Modèles semi-Markoviens contextuels) Un modèle semi-Markovien contextuel est totalement défini par le  $n$ -uplet  $\langle S, \Sigma, \Theta, F, \pi_\theta, A_\theta, B_\theta \rangle$ , où :

- $S$  est un espace d'états de cardinalité  $n$  et  $s_i$  représente le  $i^e$  état de la séquence d'états,
- $\Sigma$  est un alphabet de cardinalité  $m$  et  $\epsilon_j$  représente le  $j^e$  symbole de l'observation,
- $\Theta$  est l'ensemble des instances du contexte  $\theta$ ,
- $F$ , est une fonction d'évolution des instances du contexte  $\theta$ ,
- $\pi_\theta$  est un vecteur stochastique paramétré dont la  $i^e$  coordonnée correspond à la probabilité que la séquence d'état débute par l'état  $i$  :

$$\forall \theta \in \Theta, \sum_{i=1}^n \pi_i(\theta) = 1.$$

$\pi_i$  est une fonction de  $\theta$  qui représente la distribution initiale dans le contexte  $\theta$  :

$$\forall i \in \{1, \dots, n\}, \pi_i(\theta_1) = P(s_1 = i | \theta_1),$$

- $A$  est une matrice stochastique  $n \times n$  où  $a_{ij}$  représente la probabilité que l'état  $i$  soit suivi par l'état  $j$  dans la séquence d'état et où chaque  $a_{ij}$  est une fonction de  $\theta$  :

$$\forall \theta \in \Theta, \forall i \in \{1, \dots, n\}, \sum_{j=1}^n a_{ij}(\theta) = 1.$$

$$\forall k, t \in \mathbb{N}, \forall i, j \in \{1, \dots, n\},$$

$$a_{ij}(\theta_k) = P(s_{t+1} = j | s_t, \theta_k),$$

- $B$  est une matrice stochastique  $n \times m$  où  $b_{ik}$  représente la probabilité d'observer le symbole  $k$  lorsque l'on est sur l'état  $i$  :

$$\forall \theta \in \Theta, \forall i \in \{1, \dots, n\}, \sum_{k=1}^m b_{ik}(\theta) = 1.$$

$$\forall k, t \in \mathbb{N}, \forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, m\}$$

$$b_{ij}(\theta_k) = P(\epsilon_{t+1} = j | s_t, \theta_k).$$

Afin de conserver la facilité d'expression d'une probabilité d'avoir une certaine séquence d'observations, il est nécessaire d'exprimer l'hypothèse de Markov contextuellement. Soit  $T$  la longueur d'une séquence d'observations :

$$P(s_t | s_{t-1}, \theta_1, \dots, \theta_t, \dots, \theta_T) = P(s_t | s_{t-1}, \theta_t)$$

$$P(s_t, \epsilon_t | s_1, \dots, s_{t-1}, \epsilon_1, \dots, \epsilon_{t-1}, \theta_1, \dots, \theta_t) = P(s_t, \epsilon_t | s_t, \theta_t)$$

Autrement dit, les calculs des probabilités de transition et d'émission ne se font que par rapport au contexte courant et à l'état sur lequel on se trouve. Aucun historique des états ou des contextes n'intervient dans les calculs.

Avec ce modèle, nous allons représenter une grille de programmes en associant chaque état à un type d'émission. Mais au lieu de considérer un ensemble de symboles, nous allons utiliser un ensemble continu de valeurs entières afin de représenter les durées des programmes. Nous verrons dans la section suivante comment obtenir les probabilités des durées des émissions.

Il faut à présent entraîner le modèle afin d'en estimer les paramètres.

### 3.2 Estimation des paramètres du modèle

Intuitivement, il est aisé de voir que la prise en compte d'un contexte et de durées dans le modèle de Markov va augmenter le coût de l'apprentissage. Dans notre application à la structuration des flux et avec le contexte et la fonction d'évolution que nous avons choisis, l'estimation des paramètres du modèle semi-Markovien contextuel va donc dépendre fortement de la prédiction des dates de la série temporelle. En effet la fonction d'évolution  $F$  nécessite de connaître la durée entre deux évènements, c'est-à-dire dans notre cas la durée de l'émission. Nous proposons dans la section suivante une méthode de prédiction des durées.

En supposant le problème des durées résolu, l'estimation des paramètres de notre modèle revient à l'estimation des paramètres d'une chaîne de Markov, en calculant pour un contexte donné les fréquences d'occurrences pour chaque type, puis chaque transition.

L'algorithme parcourt l'ensemble des exemples contenus dans la base. Pour chaque exemple, l'heure de début est estimée en fonction de l'heure de début réelle et de la durée du programme précédent. Et pour toutes les durées possibles de l'exemple en cours, on calcule les probabilités de transition.

Pour donner un ordre de grandeur du coût de l'apprentissage, nous pouvons chiffrer à 604800 le nombre de contextes possibles dans  $\Theta$  et à 38 le nombre de genres d'émissions que nous considérons.

A ce stade, les valeurs de la matrice  $B$  du modèle semi-Markovien contextuel n'ont pas besoin d'être évaluées. L'outil de régression des durées que nous allons utiliser devra être en mesure de nous fournir la probabilité de ces durées.

## 4 Prédiction des dates et estimation des probabilités

Afin que notre représentation des grilles de programmes par notre extension des modèles markoviens soit valide, il est nécessaire pour un type d'émission donné de pouvoir estimer la probabilité d'une durée prédite. Plus généralement, pour pouvoir prédire une série temporelle irrégulière, nous avons besoin de prédire la durée entre deux évènements déclenchant un relevé des données.

### 4.1 Arbre de régression pour la prédiction des durées

Nous avons vu que la difficulté de prédire une série temporelle irrégulière dont la collecte des valeurs repose sur le déclenchement d'un évènement est un problème difficile qui demande la prévision de l'intervalle entre deux évènements. Dans le cas de la représentation des grilles de programmes, l'évènement est le début d'une nouvelle émission et il faut donc être en mesure de prédire la durée d'une émission. Contrairement à ce que nous pensions au début des travaux, même si deux émissions appartiennent à la même collection, leurs durées peuvent être très différentes. Par exemple, le magazine *Tout le monde en parle* sur France 2 peut durer entre une heure et demi et trois heures.

Il existe différentes méthodes de régression mais nous avons opté pour une méthode symbolique tirée des arbres de décision. Ces derniers sont des classifieurs dont l'entraînement supervisé se fait sur des classes symboliques. Quinlan dans [9] propose de remplacer ces classes symboliques par des classes continues. On parle alors d'arbres de régression. Comme dans la plupart des méthodes de régression, on suppose qu'il existe une seule variable de sortie, celle à approcher, et plusieurs variables d'entrées. L'avantage des arbres de décision ou de régression est que l'ensemble des variables d'entrées peut comporter aussi bien des variables continues que des variables symboliques.

Un arbre est construit récursivement par partition de l'ensemble d'apprentissage. Au début, tous les exemples sont dans un même ensemble. Le système va tenter de partitionner cet ensemble en fonction des valeurs des variables d'entrée. En fonction d'un certain critère, un test est choisi et le système réitère le processus sur chacune des partitions obtenues. Dans le cas d'un arbre de décision, lorsque les données d'une partition ont atteint l'entropie désirée, le noeud est transformé en noeud terminal, appelé feuille. Tout nouveau cas présenté à l'arbre et arrivant sur cette feuille sera attribué à la classe dominante dans la partition associée à la feuille. Dans le cas d'un arbre de régression, il s'agira de la moyenne et de l'écart-type des individus présents dans la partition. Parfois même, une fonction est rattachée à la feuille pour régresser linéairement localement la variable d'entrée.

Afin d'adapter cet apprentissage symbolique à nos be-

soins, nous avons remplacé l'écart-type par la valeur minimum et maximum des durées ; cela nous permet d'obtenir une fenêtre temporelle dans laquelle on peut s'attendre à l'évènement suivant. En effet, si notre fenêtre temporelle devait être formée par la moyenne plus ou moins l'écart-type, elle ne couvrirait pas tous les cas de l'ensemble d'apprentissage. De ce fait, le critère proposé par Quinlan dans [9] a été remplacé par :

$$\Delta Error = \max(T) - \min(T) - \sum_i \frac{|T_i|}{|T|} (\max(T_i) - \min(T_i))$$

où  $T$  représente l'ensemble à partitionner et  $T_i$  une de ses partitions.

Les tests que nous avons effectués montrent que la variable d'entrée la plus importante est le genre de programme. Nous forçons donc l'apprentissage à commencer par cet attribut. Les autres tests effectués sur les données concernent l'heure de début, le jour de la semaine, ainsi que le genre du programme précédent. Nous nous sommes en effet aperçus que ce dernier était un bon moyen de réduire les problèmes causés par les retards de diffusion. Par exemple, considérons la suite constituée d'un magazine court, d'une météo et d'un magazine long. Si le premier magazine commence trop tôt, il faut éviter que le deuxième soit considéré comme un magazine court au lieu d'un long.

Il arrive parfois que les arbres de décision ou de régression souffrent d'un sur-apprentissage. Dans ce cas là, les classes ou les valeurs de l'ensemble d'apprentissage sont parfaitement prédites, mais le taux d'erreur sur de nouveaux exemples est élevé. La plupart du temps, l'arbre est élagué, ce qui nécessite un ensemble de test. Comme nous n'avons pas suffisamment d'exemples, compte tenu de la diversité des durées, nous avons préféré introduire deux paramètres  $\rho$  et  $\lambda$ .  $\rho$  est la distance minimale entre la durée maximum et la durée minimum en-dessous de laquelle la partition de l'ensemble de données doit s'arrêter. Le paramètre  $\lambda$ , lui, impose une distance minimum entre la durée minimale et la durée maximale. Il est très important de faire la distinction entre les deux paramètres. Les sous-ensembles sont partitionnés tant que l'amplitude des valeurs est supérieur à  $\rho$ . Mais il se peut qu'un test crée un sous-ensemble avec une durée minimale très proche de la durée maximale. Alors on force la distance entre les deux à être égale à  $\lambda$ .

### 4.2 Estimation des probabilités par une gaussienne asymétrique

Nous sommes à présent en mesure de prédire des durées avec un arbre qui nous fournit une durée moyenne, une durée minimale et une durée maximale.

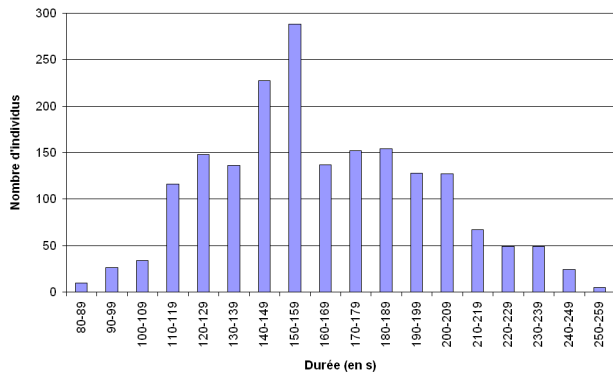


Figure 1 – Distribution des valeurs pour une feuille de l'arbre représentant la météo.

La figure 1 nous montre un exemple de la répartition des durées dans l'ensemble correspondant à la feuille représentant les bulletins météo. Nous faisons donc l'hypothèse que les distributions peuvent être approchées par une gaussienne asymétrique comme celle proposée dans [10]. La probabilité d'une durée  $d \in \mathbb{D}$  peut être donnée par :

$$A(d, \mu, \sigma^2, r) = \frac{2}{\sqrt{2\pi}} \frac{1}{\sigma(r+1)} \begin{cases} e^{-\frac{(d-\mu)^2}{2\sigma^2}} & \text{si } d > \mu \\ e^{-\frac{(d-\mu)^2}{2r^2\sigma^2}} & \text{sinon} \end{cases}$$

Dans cette expression,  $\mu$  représente la valeur moyenne des durées retournée par la feuille de l'arbre de régression,  $\sigma$  est égale à la valeur absolue de la différence entre la valeur moyenne et la valeur maximale est  $r$  est donné par la relation :

$$r = \left| \frac{\mu - \text{valeur minimale}}{\mu - \text{valeur maximale}} \right|$$

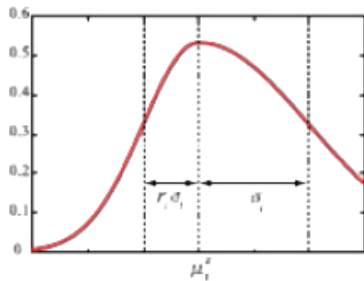


Figure 2 – La distribution gaussienne asymétrique selon [10].

## 5 Expérimentations

Nous avons procédé à un apprentissage de l'arbre de décision, puis du modèle markovien en utilisant les grilles de programmes de 2003 et 2004 de France 2, dont l'Institut National de l'Audiovisuel est en possession. La base

de données n'a pas été filtrée. Par exemple, nous n'avons pas distingué les vacances scolaires ou l'été de la période normale. Nous n'avons retiré de la base que des problèmes techniques, en laissant volontaire du bruit comme par exemple les jeux Olympiques d'été. Des résultats meilleurs pourraient sans doute être obtenus après filtrage de la base d'apprentissage et de la base de tests.

### 5.1 Test de l'arbre de régression

Nous avons procédé à un test sur l'ensemble des durées de 2005. Il est possible qu'un individu présenté à l'arbre de décision n'aboutisse pas à une feuille. Dans ce cas là, l'individu est considéré comme inconnu et lui sont affectés la durée moyenne, minimale et maximale du dernier noeud qu'il a franchi. Nous considérons comme erreur uniquement les individus dont la durée réelle n'est pas comprise entre les durées minimales et maximales prédites.

De plus, il est souhaitable que les durées soient encadrées le plus finement possible, ce qui peut être obtenu en donnant des valeurs optimales aux paramètres  $\rho$  et  $\lambda$ .

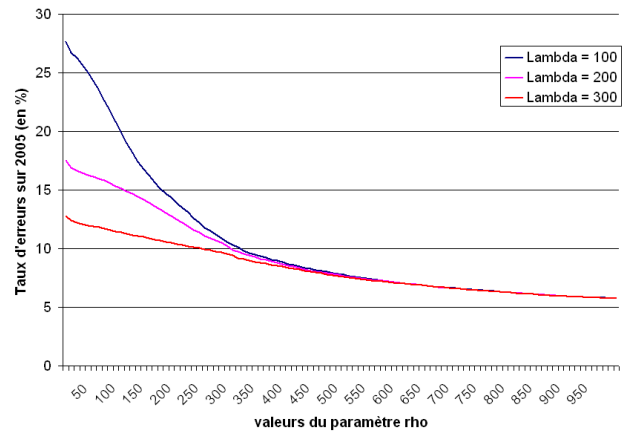


Figure 3 – Impact du paramètre  $\rho$  sur le taux d'erreurs de prédiction pour des arbres entraînés avec différentes valeurs de  $\lambda$ .

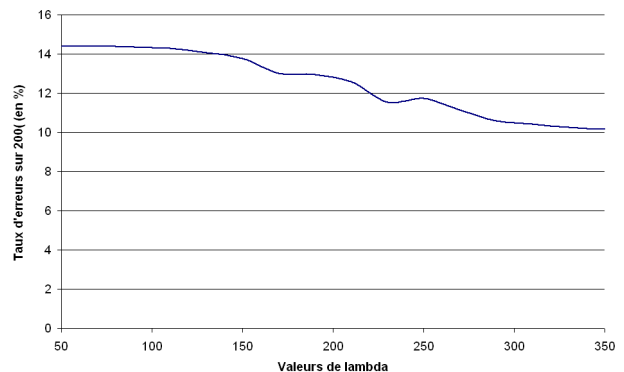


Figure 4 – Impact du paramètre  $\lambda$  sur le taux d'erreur avec  $\rho = 200$ .

Les figures 3 et 4 confirment que pour baisser le taux d'erreurs de prédiction sur 2005, il faut accepter d'encadrer ces durées moins finement. Avec  $\rho = \lambda = 300$ , il est possible d'obtenir 90,5% de bonnes prédictions avec des encadrements des durées d'envergure 200 secondes en moyenne.

## 5.2 Expressivité du modèle de Markov

En utilisant un arbre de régression avec les paramètres  $\rho = \lambda = 300$  pour l'entraînement du modèle semi-Markovien contextuel, nous l'avons mis à l'épreuve sur le mois de Mars 2005. L'année 2005 est en effet une mauvaise année pour tester un modèle statistique car elle a connu plein d'évènements (mort du pape, référendum pour la constitution) qui ont bouleversé particulièrement les chaînes nationales, dont fait partie France 2. Il était nécessaire de tester le modèle sans tenir compte de la prédiction des durées. Nous avons pour cela présenté une à une les émissions du mois de Mars en mettant à jour le contexte avec la durée réelle de l'émission. 83% des journées sont *observables* par le modèle. Les journées qui ont été comptées comme des erreurs contenaient toujours une émission déplacée dans la journée ou un communiqué de la chaîne (seuls les évènements exceptionnels ont été retirés de la base de tests). A chaque transition se fait d'un état vers au plus 6 états contre 38 états au plus avec un modèle non contextuel.

## 6 Conclusion

Nous avons présenté dans cet article un moyen de prédire des séries temporelles irrégulières dans le but de pouvoir prédire des grilles de programmes de chaînes de télévision. Nous avons souligné la difficulté de la tâche qui nécessite non seulement la prédiction des valeurs, mais aussi des dates.

Les résultats obtenus avec des méthodes statistiques et des méthodes de régression symbolique simples sont fortement encourageant même si tout apprentissage a ses limites.

Nous poursuivons nos travaux en essayant de prendre en compte cette fois-ci les grilles prévisionnelles. Nous essayons de ramener notre problème à une prédiction de séries temporelles irrégulières compte tenu d'une série temporelle *tuteur* qui est une observation partielle et imprécise de celle-ci.

## Références

[1] Cees G.M. Snoek et Marcel Worring. Multimodal video indexing : A review of the state-of-the-art. *Multimedia Tools and Applications*, 25(1) :5–35, 2005.

[2] C. Lee Giles, Steve Lawrence, et Ah Chung Tsoi. Noisy time series prediction using a recurrent neural network and grammatical inference. *Machine Learning*, 44(1/2) :161–183, July/August 2001.

[3] Eric Perlman et Akshay Java. Predictive Mining of Time Series Data in Astronomy. *Astronomical Data Analysis Software and Systems XII ASP Conference Series*, 295 :431–434, January 2003.

[4] Bo Xu et Ouri Wolfson. Time-series prediction with applications to traffic and moving objects databases. Dans *MobiDe '03 : Proceedings of the 3rd ACM international workshop on Data engineering for wireless and mobile access*, pages 56–60, New York, NY, USA, 2003. ACM Press.

[5] Emre Erdogan, Sheng Ma, Alina Beygelzimer, et Irina Rish. Statistical models for unequally spaced time series. Dans *Proceedings of SIAM Data Mining 2005*, 2005.

[6] C. Lee Giles, Steve Lawrence, et Ah Chung Tsoi. Noisy time series prediction using a recurrent neural network and grammatical inference. *Machine Learning*, 44(1/2) :161–183, July/August 2001.

[7] S. Mukherjee, E. Osuna, et F. Girosi. Nonlinear prediction of chaotic time series using support vector machines. Dans J. Principe, L. Giles, N. Morgan, et E. Wilson, éditeurs, *IEEE Workshop on Neural Networks for Signal Processing VII*, page 511. IEEE Press, 1997.

[8] Gerhard Dangelmayr, Sabino Gadaleta, Douglas Hundley, et Michael J. Kirby. Time series prediction by estimating markov probabilities through topology preserving maps. Dans Bruno Bosacchi, David B. Fogel, et James C. Bezdek, éditeurs, *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation II*, volume 3812, pages 86–93. SPIE, 1999.

[9] J. R. Quinlan. Learning with continuous classes. Dans *Proceedings of 5th Australian Joint Conference on Artificial Intelligence*, pages 343–348, 1992.

[10] T. Kato, S. Omachi, et H. Aso. Asymmetric gaussian and its application to pattern recognition. Dans *Lecture Notes in Computer Science (Joint IAPR International Workshops SSPR 2002 and SPR 2002)*, volume 2396, pages 405–413, 2002.