



**HAL**  
open science

## Prédiction de l'état du trafic routier basée sur les motifs et les chaînes de Markov

Feda Almuhsen, Nicolas Durand, Leonardo Brenner, Quafafou Mohamed

### ► To cite this version:

Feda Almuhsen, Nicolas Durand, Leonardo Brenner, Quafafou Mohamed. Prédiction de l'état du trafic routier basée sur les motifs et les chaînes de Markov. 25èmes Rencontres de la Société Francophone de Classification (SFC 2018), Sep 2018, Paris, France. hal-01858562

**HAL Id: hal-01858562**

**<https://hal.science/hal-01858562>**

Submitted on 12 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Prédiction de l'état du trafic routier basée sur les motifs et les chaînes de Markov

Feda AlMuhisen, Nicolas Durand, Leonardo Brenner, Mohamed Quafafou

Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France  
{prénom.nom}@univ-amu.fr

**Résumé.** Cet article propose une nouvelle méthode de prédiction de l'état du trafic routier sur de courtes fenêtres temporelles. Cette méthode bénéficie des méthodes de partitionnement spatial, d'extraction de motifs et de modélisation Markovienne. À partir de trajectoires, nous extrayons des régions fréquentes où des véhicules passent de manière récurrente en utilisant les motifs fermés fréquents et nous détectons l'état du trafic en se basant sur l'évolution des motifs dans le temps. Nous prédisons ensuite l'état suivant du trafic pour les régions fréquentes en se basant sur les chaînes de Markov. Les expérimentations effectuées sur des données réelles montrent que la méthode proposée est plus efficace qu'une méthode de base.

## 1 Introduction

L'un des principaux enjeux de la ville intelligente (Smart City) est la surveillance et la prévision du trafic. Dans ce contexte, il est nécessaire de fournir une vision globale de l'état actuel ou futur du trafic dans la ville (Chen et al., 2015). Par exemple, les prévisions de trafic à court terme peuvent être utilisées directement par les experts pour prendre des mesures pertinentes contre les tendances actuelles ou naissantes du trafic. Le développement de tels systèmes nécessite une analyse des données historiques des trajectoires. Ce type d'analyse est également utilisée dans d'autres applications telles que l'amélioration de la navigation de route et la planification des zones urbaines. Le but de cet article est de prédire l'état du trafic dans un avenir proche pour une région spécifique. Nous introduisons une nouvelle méthode appelée TS2PM (Traffic States Prediction Patterns and Markov Chains), basée sur le partitionnement de l'espace, l'extraction de motifs et la modélisation par des chaînes de Markov (Stewart, 1994). Pour cela, la ville est divisée en régions en utilisant une valeur de granularité spatiale, puis l'état du trafic est détecté pour chaque région à travers des fenêtres temporelles (AlMuhisen et al., 2018). En donnant l'état du trafic pour un intervalle de temps, la méthode prédit l'état du trafic dans la fenêtre de temps suivante pour chaque région en utilisant des chaînes de Markov. Enfin, les prévisions de l'état du trafic sont visualisées sur une carte géolocalisée. Les expérimentations sur des données réelles montrent que la méthode proposée est plus précise qu'une méthode de référence qui n'utilise pas des motifs.

## 2 Méthode proposée

Dans cette section, nous présentons notre méthode, TS2PM, pour la prévision de l'état du trafic. La méthode modélise des régions et des trajectoires de façon discrète et détecte l'évolution de l'état de trafic au cours du temps en utilisant des modèles stochastiques. Elle se compose principalement des étapes suivantes :

1. **Détection de l'état du trafic** : cette étape est détaillée dans notre précédent travail (AlMuhisen et al., 2018). La zone étudiée est divisée en régions de taille égale selon une valeur de granularité spatiale, puis chaque trajectoire est transformée et représentée dans cette zone comme un ensemble des régions. Ensuite, cette représentation discrète des trajectoires est utilisée pour construire un treillis de concepts formels fréquents pour encoder la correspondance maximale entre les motifs fréquents découverts et les trajectoires dans une certaine fenêtre temporelle. Chaque élément de ce treillis, appelé concept formel, est composé d'un motif fermé (d'attributs) et d'un ensemble fermé d'objets. Le motif fermé est l'ensemble maximal d'attributs partagé par l'ensemble d'objets. Ce dernier est l'ensemble maximal d'objets contenant le motif. Dans notre contexte, un concept formel contient un ensemble de régions (le motif) et l'ensemble des trajectoires (les objets) qui traversent ces régions. Un concept formel est dit fréquent lorsque le nombre d'objets dépasse un certain seuil fixé. Nous détectons l'état du trafic du motif fermé pour chaque concept formel fréquent en étudiant son évolution à travers une paire de fenêtres temporelle  $(t_i, t_{i+1})$ . L'état du trafic appartient à un ensemble prédéfini de labels :  $\{Latent (LA), Emerging (E), Decreasing (D), Jumping (J), Lost (LO), Nothing (\emptyset)\}$ . L'état *Emerging* signifie que la présence du motif a augmenté dans  $t_{i+1}$  par rapport à  $t_i$ . L'état *Decreasing* signifie que la présence du motif a diminué dans  $t_{i+1}$  par rapport à  $t_i$ . L'état *Latent* signifie que la présence du motif est assez similaire dans  $t_{i+1}$  et  $t_i$ . L'état *Jumping* signifie le motif qui était absent dans  $t_i$  et est apparu dans  $t_{i+1}$ . L'état *Lost* signifie que le motif a disparu dans  $t_{i+1}$ . *Nothing* signifie qu'il n'y a pas de trafic détecté dans la fenêtre temporelle.
2. **Construction du modèle** : pour chaque région, nous construisons une chaîne de Markov (Stewart, 1994) où chaque état de la chaîne représente un état prédéfini du trafic. Les transitions représentent un changement de l'état du trafic d'un intervalle de temps à un autre. À chaque transition, nous associons une probabilité qui est le ratio du nombre de changements d'un état  $i$  à un état  $j$  sur le nombre total de changements à partir de l'état  $i$ . Par exemple, dans la figure 1,  $d \rightarrow la$  indique la probabilité de transition de l'état  $D$  à l'état  $LA$ .
3. **Prédiction** : deux types de prédictions peuvent être obtenus à partir de ce modèle : à long et à court terme. Une analyse stationnaire nous permet de faire une prédiction à long terme sur l'état moyen d'une région. Une analyse transitoire nous permet de prédire l'évolution du modèle à partir d'un état donné et pour une période spécifiée, cela représente une prédiction à court terme. Par exemple, nous pouvons calculer la probabilité que le trafic soit demain dans l'état  $E$  en sachant qu'aujourd'hui il est dans l'état  $D$ .

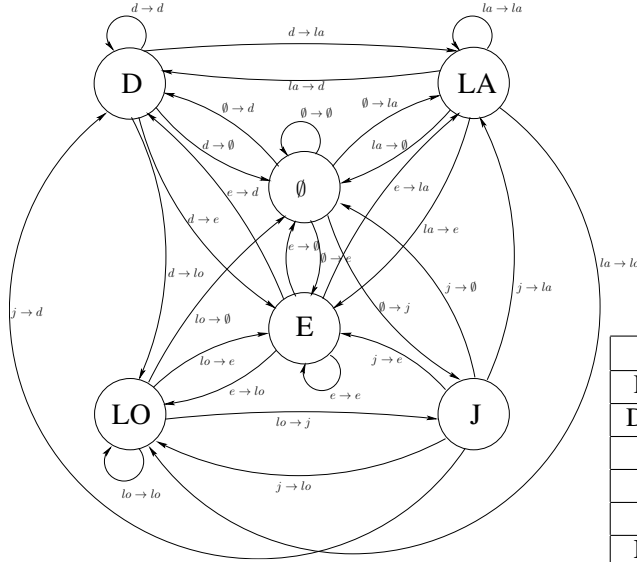


FIG. 1: Chaîne de Markov pour une région

État	Baseline	TS2PM
Emerging	0.696	0.862
Decreasing	0.792	0.930
Latent	0.919	0.980
Jumping	0.890	0.992
Nothing	0.991	0.856
<b>Moyenne</b>	<b>0.858</b>	<b>0.924</b>

TAB. 1: TS2PM vs. Baseline

### 3 Expérimentations

Les données utilisées pour les expérimentations, *T-drive*, sont des données réelles collectées par Microsoft Research Asia (Yuan et al., 2011). Ces données correspondent à des trajectoires de taxis sur une semaine, du 2 février au 8 février 2008, à Pékin. Nous avons choisi aléatoirement 222 taxis. Le protocole expérimental peut être résumé comme suit. Tout d'abord, nous avons commencé par prétraiter l'ensemble des trajectoires. Les trajectoires ont été segmentées en fonction de la granularité temporelle fixée à 12 heures. La zone étudiée a été divisée en régions en fixant la granularité spatiale à 60 mètres. Pour chaque paire de fenêtres temporelles, l'état du trafic de chaque région a été détecté. À partir de ces données générées sur l'état du trafic, nous avons construit le modèle markovien pour chaque région. Nous avons appliqué la méthode proposée pour la prédiction en séquence pendant l'intervalle de temps de données d'une semaine comme suit : à partir d'un jour<sub>*i*</sub> l'état du trafic du jour<sub>*j*</sub> a été prédit,  $i = 1 \dots 6$  et  $j = 2 \dots 7$ . La précision moyenne (avg. accuracy) des prévisions a été calculée pour chaque état du trafic. Nous avons comparé notre méthode (TS2PM) avec une méthode de base (Baseline) qui n'utilise pas de motifs pour la détection des états du trafic mais utilise notre modèle markovien. Cette méthode de base détecte les états du trafic de chaque région directement à partir de la fréquentation de celle-ci, i.e., le nombre de trajectoires la traversant. Comme nous pouvons le voir dans le tableau 1, notre méthode surpasse la méthode de base. Les figures 2 et 3 présentent respectivement la carte géolocalisée des états réels et celle des prévisions de l'état du trafic pour une zone de Pékin. Nous pouvons observer qu'une seule région a été mal prédite (région encerclée). De plus, la majorité de l'état du trafic détecté dans cette zone était *Emerging*. Cette zone a dû faire face à une augmentation du trafic.

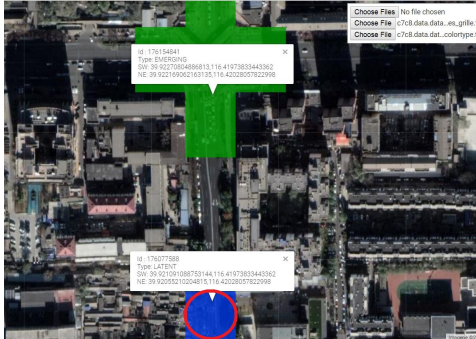


FIG. 2: État réel du trafic

FIG. 3: État prédit du trafic

## 4 Conclusion

Nous avons proposé une nouvelle méthode pour prédire l'état du trafic dans une ville. La méthode est basée sur l'extraction des motifs et les modèles de Markov. Elle se compose des étapes suivantes : détection de l'état du trafic, construction du modèle, prédiction et visualisation. Les expérimentations sur des données réelles montrent que l'approche proposée est plus précise qu'une méthode de base qui n'utilisent pas de motifs. À l'avenir, nous allons travailler sur une évolution du modèle présenté en considérant le temps et les mesures d'accélération.

## Références

- AlMuhsen, F., N. Durand, et M. Quafafou (2018). Detecting behavior types of moving object trajectories. *International Journal of Data Science and Analytics* 5(2-3), 169–187.
- Chen, W., F. Guo, et F.-Y. Wang (2015). A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems* 16(6), 2970–2984.
- Stewart, W. J. (1994). *Introduction to the numerical solution of Markov chains*. Princeton University Press.
- Yuan, N. J., Y. Zheng, X. Xie, et G. Sun (2011). Driving with knowledge from the physical world. In *KDD'11*, San Diego, CA, pp. 316–324.

## Summary

This paper proposes a new method for predicting traffic state within short time windows. This method takes advantage from space-partitioning, pattern extraction and Markov modeling. From trajectories, frequent regions are extracted where vehicles repeatedly pass through by using the frequent closed patterns and the traffic state is detected based on the evolution of these patterns over time. The next state of traffic for the frequent regions is then predicted based on the Markov models. Experiments on real-world data show that the proposed method is more accurate than a baseline method.