

# Mise en Cohérence des Objectifs de TIPE (MCOT) Modélisation et Prévision Spatio-Temporelle du Risque de Cambriolage en Ville

Maxime MUNIER, MP2I, 2021-2022

## Positionnement thématique

*Statistique, Probabilité, Informatique*

## Mots-clefs

**Mots-clefs** – Cambriolage – Séisme – Prédiction – Algorithme – Processus  
**Keywords** – Heist – Earthquake – Prediction – Algorithm – Process

## Bibliographie commentée

La prédiction policière [1] est un domaine de recherche dont l'objectif principal est de développer des machines à prédire les crimes, en tirant profit des algorithmes du machine learning et de l'accessibilité croissante à une diversité de données. Depuis 2011, l'enthousiasme autour de « Predpol » [2], le logiciel de police prédictive américain, électrise la terre entière. Son algorithme secret, toujours comparé aux précogs de « Minority Report » ou à la Machine de « Person of Interest » [3], semble tenir plus de la magie que de la science. La société, elle, affiche partout des résultats là où sa technologie est déployée, à savoir si l'on en croit le site de Predpol, une petite dizaine de villes américaines pour l'instant, Los Angeles et Atlanta étant les plus importantes, on y observe une baisse de la criminalité de 10 à 30%.

Depuis quelques années, Andrea Bertozzi [4], anime un groupe de chercheurs de l'Université de Californie à Los Angeles rassemblant mathématiciens, anthropologues et policiers. Leur but est de comprendre et prévenir l'activité criminelle dans

une ville comme Los Angeles. Andrea Bertozzi a vite compris que les systèmes criminels pouvaient être modélisés à l'aide de ses outils favoris : la densité criminelle d'un lieu ayant tendance à alimenter son potentiel criminel, et réciproquement, chaque donnée évolue en fonction de l'autre. Chez les malfrats, on a tendance à se rapprocher peu à peu des zones les plus attractives en matière de criminalité... Ce qui finit par faire augmenter le nombre de crimes commis dans ces zones, et donc par faire augmenter leur attractivité. Pour tenter de comprendre où et quand pourraient se dérouler les prochains crimes, la mathématicienne et son équipe ont pris en compte un certain nombre de paramètres, à commencer par ce qu'on pourrait appeler la « densité criminelle » (c'est-à-dire le nombre de criminels par mètre carré) et le « potentiel criminel » (y a-t-il des crimes potentiellement « intéressants » à commettre?) d'un certain nombre de lieux. Pour la mathématicienne, tout cela est lié à un concept connu sous le nom de « théorie de la vitre brisée » [5], ainsi nommé parce que se basant sur le principe suivant : si une vitre brisée n'est pas immédiatement réparée, il est probable que d'autres vitres du même bâtiment seront bientôt brisées à leur tour, le lieu étant considéré comme abandonné et délabré. Un principe qui fonctionne aussi avec une maison cambriolée qui, parce qu'elle est identifiée comme insuffisamment protégée, a plus de chances de l'être à nouveau. D'un point de vue purement mathématique, l'équipe de recherche en est venue à la conclusion suivante : de par leur structure, les équations obtenues rappellent celles qui modélisent l'évolution de certaines bactéries. Bertozzi voit du même œil la répartition et l'évolution de la criminalité : un crime non signalé, non élucidé ou non puni peut servir d'origine à la constitution progressive d'un « point chaud », une zone géographique dans laquelle les crimes seront de plus en plus nombreux. La dynamique des tremblements de terre et de leurs répliques semble également très proche de la façon dont un crime peut être suivi par d'autres, au même endroit ou dans les environs.

C'est ce lien avec la sismologie en particulier qui semble relativement intéressant [6]. En effet, on pourrait croire que les tremblements de terre sont aléatoires et distribués selon le modèle de Poisson de sorte que chaque tremblement de terre soit indépendant des autres. Mais en réalité, s'il y a un tremblement de terre, il va très probablement y avoir des contrecoups, c'est-à-dire une série de tremblements de terre, au même endroit, en succession rapide. Les scientifiques et les mathématiciens ont développé le « Processus de Hawkes » [7] qui prend en compte le fait que des événements ne sont pas complètement indépendants les uns des autres. Ce qui est intéressant, c'est que la criminalité, se base sur le même schéma [8]. Donc si on considère les cambriolages par exemple, tout ceux qui se sont déjà fait cambrioler savent que les chances de se faire cambrioler à nouveau dans une courte période de temps, sont beaucoup plus fortes. Il y a plusieurs raisons qui peuvent expliquer cela. La première est que les cambrioleurs connaissent déjà le plan de la maison, ils

savent où sont gardés les objets précieux, ils savent plein de chose sur le quartier, ainsi, les chances de se faire cambrioler à nouveau augmentent, mais elle augmente aussi pour les voisins, les voisins des voisins, et ainsi de suite dans toute la rue. Ce processus qui considère les événements comme liés dans le temps, veut dire qu'on peut modéliser ce qui se passe avec les cambriolages de façon statistique.

## Problématique retenue

C'est à l'aide de la base de données recensant l'intégralité des crimes commis dans la ville de Chicago que se base ce projet de TIPE. Mon étude se limite à l'évaluation du risque des cambriolages à partir d'historiques de cambriolages signalés dans la ville de Chicago. Son intérêt est double : sur le plan pédagogique, cette étude demande une analyse en temps et en espace qui exige une approche plus originale que l'application de méthodes de régression classique ; sur le plan des enjeux, ce projet permet de tester et d'évaluer les performances de modèles statistiques récents, dans l'espoir de développer des méthodes pratiques et efficaces. Plus précisément, il m'a semblé intéressant de répondre à la question suivante : A l'aide des Processus de Hawkes, peut-on prédire de manière précise des zones à risques futurs à partir de données du passé ?

## Objectifs du TIPE

1. Modélisation : après quelques rappels théoriques, j'expliquerai le "Processus de Hawkes" ainsi que le modèle utilisé pour répondre à la problématique du sujet.
2. Simulation : En se basant sur la base de données mise à disposition. J'estimerai les paramètres du modèle à l'aide de plusieurs méthodes utilisées en statistiques comme celle du maximum de Vraisemblance. A cette fin, j'utiliserai également des algorithmes comme celui d'Ogata.
3. Validations et Observations : En utilisant les bases de données des différentes années, il sera assez facile de comparer les estimations du modèles ainsi conçu avec les données de la réalité afin de valider ou non ce modèle mathématiques. Des commentaires sur les résultats seront également exprimés.

## Références

- [1] Paul GERBER : Des crimes et des séismes, la police prédictive entre science, technique et divination. *RÉSEAUX*, 206:95–123, Juin 2017. <https://www.>

[cairn.info/revue-reseaux-2017-6-page-95.htm](http://cairn.info/revue-reseaux-2017-6-page-95.htm).

- [2] Guillaud HUBERT : Police prédictive : la prédiction des banalités. *Le Monde*, Juin 2015.
- [3] Florent MASSEGLIA : Person of interest : la surveillance des données individuelles pour anticiper les crimes. *Interstices*, Novembre 2013.
- [4] Thierry GOUDON : Gendarmes, voleurs et... mathématiciens! *Breves de Maths*, Septembre 2013.
- [5] Thomas MESSIAS : La théorie de la vitre brisée, ou comment les maths préviennent le crime. *Slate*, Juillet 2020.
- [6] NUMBERPHILE : The mathematics of crime and terrorism - numberphile. Vidéo Youtube, Février 2016. <https://www.youtube.com/watch?v=1CjspXB5F4A>.
- [7] Loubna Ben ALLAL : Processus de hawkes : Modélisation, simulation et inférence, Septembre 2020. <https://hal.inria.fr/hal-02947341>.
- [8] Alice Launay YANNICK BESSY-ROLAND : Modélisation du risque terroriste par les processus de hawkes, 2018. <https://www.youtube.com/watch?v=1CjspXB5F4A>.