

TITOUAN VAYER

7 rue Trarieux, 69003 Lyon, France

(+33)682325570 ♦ titouan.vayer@ens-lyon.fr ♦ <https://tvayer.github.io/>

PARCOURS UNIVERSITAIRE

Doctorat en informatique

2017 - 2020

Université Bretagne Sud, IRISA, Vannes

Titre: *Une contribution au Transport Optimal sur des espaces incomparables*

Supervisée par Nicolas Courty, Laetitia Chapel et Romain Tavenard

Défendue le 5 Novembre 2020 devant le jury composé de:

- *Rapporteurs*: Gabriel Peyré (DR CNRS DMA), Marc Sebban (PR Univ. Jean Monnet)
- *Examineurs*: Filippo Santambrogio (PR Univ. Claude Bernard), Julie Delon (PR Univ. Paris Descartes), Pierre Vanderghenst (PR EPFL), Rémi Flamary (MCF CMAP École Polytechnique)

Diplôme d'ingénieur

2013 - 2017

Ecole Centrale de Lille

Spécialités : Mathématiques, informatique et science des données.

Licence de Mathématiques

2013 - 2015

Université de Lille

Spécialités : Probabilités et statistiques.

EXPÉRIENCES PROFESSIONNELLES

Ecole Normale Supérieure, Lyon

nov 2020 - présent

Post-Doctorant.

Laboratoire de l'informatique et du Parallélisme (LIP, UMR 5668). Équipe DANTE

Thématiques de recherche: apprentissage compressif, optimisation sous contrainte de parcimonie, apprentissage PAC.

Université Bretagne Sud, Vannes

2017 - 2020

Doctorant.

Institut de recherche en informatique et systèmes aléatoires (IRISA, UMR 6074). Équipe OBELIX

Thématiques de recherche: transport optimal sur graphes, séries temporelles, données hétérogènes.

Crédit Agricole Consumer Finance, Lille

2016 - 2017

Data Scientist dans le service modélisation

Création et industrialisation de modèles de machine learning pour le crédit.

Analyse de jeux de données massifs (web) avec Spark et Hadoop.

Régression logistique et forêt d'arbres décisionnels.

Michelin, Clermont-Ferrand

May 2015 - Dec 2015

Stagiaire

Analyse de données et statistiques.

Création d'un modèle de prévision des flux de pneus.

Réseau bayésien, analyse de séries temporelles.

Skins International Trading, Zürich, Suisse

Jan 2013 - Feb 2013

Stagiaire

Optimisation du stock de gestion de la société.

Programmation linéaire.

ACTIVITÉS D'ENSEIGNEMENT

Mes enseignements ont été effectués lors de ma thèse à l'IUT de Vannes dans le département STID (Statistique et Informatique Décisionnelle) pour des étudiants de deuxième année DUT STID ainsi qu'au sein de l'École Centrale de Nantes pour des étudiants en dernière année option Mathématiques appliquées. Cette section présente les enseignements effectués de l'année scolaire 2017/2018 à l'année scolaire 2019/2020.

Un résumé des enseignements année par année est donné dans le tableau suivant:

Discipline	Année	Établissement	Public	eqTD	Nature	Responsabilités
Régression Logistique	2017/2018	IUT Vannes	DUT STID (L2)	27h	TD/TP	Création sujets TD/TP/Exams
	2018/2019	IUT Vannes	DUT STID (L2)	27h	TD/TP	Création sujets TD/TP/Exams
	2019/2020	IUT Vannes	DUT STID (L2)	39h	CM(3h) + TD/TP(36h)	Création du cours + création sujets TD/TP/Exams
Apprentissage profond	2018/2019	Ecole Centrale de Nantes	Ingénieur (M2)	15h	cours/TD/TP	Création du cours + création sujets TD/TP
	2019/2020	Ecole Centrale de Nantes	Ingénieur (M2)	11.5h	cours/TD/TP	Création du cours + création sujets TD/TP

Ainsi qu'un résumé quantitatif dans le tableau suivant :

Discipline	Public	CM	TD/TP	Total	Rédaction
Régression Logistique	Bac + 2	3	90	93	Cours, TD, TP, Examen
Apprentissage profond	Bac + 5	0	26.5	26.5	Cours, TD, TP
Bilan		3	116.5	119,5	

Apprentissage profond. Ecole Centrale de Nantes (ECN) (26.5h) 2018-2020

Lors de ma thèse, j'ai effectué 26.5h EqTD d'enseignement de l'apprentissage profond à l'ECN sous la responsabilité de Bertrand Michel.

Le cours était destiné à des étudiants ingénieurs en option Mathématiques Appliquées en dernière année.

Ce cours avait pour objectif d'approfondir les connaissances des étudiants en apprentissage profond. J'ai créé les supports de cours ainsi que ceux des TD/TP.

- *Thèmes abordés*: rappels sur les distributions de probabilités, sur la descente de gradient et la différentiation automatique. Théorie et pratique des modèles génératifs (GAN). Introduction au transport optimal et à son utilisation pour les GAN. Auto-Encodeur Variationnel. Fine-Tuning d'un réseau préapprenant sur des données réelles.

La mise en pratique de ces notions se faisaient en Python lors de différents TP et TD.

Les supports de cours ainsi que les différents TP sont accessibles aux pages suivantes:

- 2018: https://github.com/tvayer/tvayer.github.io/tree/master/cours/deep_ecn_2018
- 2019: https://github.com/tvayer/tvayer.github.io/tree/master/cours/deep_ecn_2019

Lors de ma thèse, j'ai effectué 93h EqTD d'enseignement de la régression logistique à l'IUT de Vannes sous la responsabilité de Laetitia Chapel.

Le cours était destiné à des étudiants de deuxième année en DUT *Statistique et Informatique Décisionnelle*.

J'étais chargé des TD/TP en 2018 et 2019 et responsable du cours en entier en 2020 (montage du cours, création des supports de TD/TP et de l'examen final pour une centaine d'étudiants).

- *Thèmes abordés* : théorie et pratique de la régression logistique, classification supervisée binaire, maximum de vraisemblance et estimation des paramètres du modèle, odds et odds-ratio, tests statistiques de significativité, critère AIC et sélection de variables.

La mise en pratique de ces notions se faisaient en R lors de différents TP et TD.

Le support de cours ainsi que les différents TP sont accessibles à la page suivante:

- 2020: https://github.com/tvayer/tvayer.github.io/tree/master/cours/reg_log_stid

ACTIVITÉS DE RECHERCHE

Dans cette section je détaille mes différents travaux de recherche effectués lors de ma thèse sous la direction de Nicolas Courty, Laetitia Chapel et Romain Tavenard. Ils se divisent principalement selon trois axes centrés autour du transport optimal:

- L'apprentissage statistique sur des données structurées telles que les graphes et les séries temporelles.
- L'apprentissage statistique sur des données hétérogènes qui intervient lorsque les données d'apprentissage proviennent de sources diverses (images de différentes résolutions par exemple).
- Un versant plus théorique consacré au développement des outils de transport optimal et plus précisément concernant la théorie de Gromov-Wasserstein.

Voici un tableau récapitulant les publications aux termes de la thèse :

	2018	2019	2020	Total
Revue internationale avec comité de relecture (Algorithms)	0	0	1	1
Conférences internationales avec actes (NeurIPS/ICML)	0	2	1	3
Conférences nationales avec actes (CAP/GRETSI)	1	1	0	2

Publications principales

[1] Titouan Vayer, Ievgen Redko, Rémi Flamary et Nicolas Courty. *CO-Optimal Transport*. In Advances in Neural Information Processing Systems 33 (NeurIPS). 2020. Article court. Top 2 des conférences en Machine Learning & Intelligence Artificielle. IF: 33.49.

url: <https://papers.nips.cc/paper/2020/hash/cc384c68ad503482fb24e6d1e3b512ae-Abstract.html>

[2] Titouan Vayer, Rémi Flamary, Nicolas Courty, Romain Tavenard et Laetitia Chapel. *Sliced Gromov-Wasserstein*. In Advances in Neural Information Processing Systems 32 (NeurIPS). 2019. Article court. Pages = 14753–14763. Top 2 des conférences en Machine Learning & Intelligence Artificielle. IF: 33.49. Taux d'acceptation: 21.6%.

url: <https://papers.nips.cc/paper/2019/hash/a9cc6694dc40736d7a2ec018ea566113-Abstract.html>

[3] Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard et Nicolas Courty. *Optimal Transport for structured data with application on graphs*. In International Conference on Machine Learning (ICML), volume 97, 2019. Article court. Pages = 6275–6284. Top 6 des conférences en Machine Learning & Intelligence Artificielle. IF: 18.48. Taux d'acceptation: 22.6%.
url : <http://proceedings.mlr.press/v97/titouan19a.html>

Transport optimal pour les données structurées.

Thèmes de recherche: transport optimal, théorie des graphes, programmation quadratique, optimisation non-convexe, apprentissage supervisé et non supervisé.

Article correspondant [3]

Le transport optimal est un outil mathématique qui permet de trouver des relations entre des ensembles de points et qui définit une notion géométrique de distance entre des distributions de probabilité. Grâce à ses propriétés le transport optimal est devenu très populaire ces dernières années pour l'analyse de données. Cependant, dans sa formulation originelle, il reste assez limité aux applications dans lesquelles les points font partie d'un même espace métrique. Cette limitation empêche notamment son utilisation pour une variété de tâches dans lesquelles il existe une information de structure supplémentaire sur les données.

Je me suis intéressé dans ma thèse à adapter et étendre les outils classiques du transport optimal dans des contextes d'apprentissage avec des graphes ce qui nécessite, entre autres, d'intégrer la théorie de Gromov-Wasserstein.

J'ai ainsi proposé une méthode basée sur le transport optimal, appelée Fused Gromov-Wasserstein (FGW), qui définit une notion de distance entre des graphes. FGW a été appliquée pour la comparaison de signaux sur graphes ainsi que pour la classification, le clustering et la simplification de graphes. En terme de résolution numérique j'ai proposé une variante intéressante en terme de complexité de l'algorithme de Frank-Wolfe pour la résolution du problème de FGW et de Gromov-Wasserstein.

Ces différents travaux ont été publiés et présentés à ICML 2019 [3] (top 6 des conférences en Machine Learning) ainsi que dans les conférences nationales CAp 2018 et GRETSI 2019 [5,6]. Ils ont aussi donné lieu à une version étendue de FGW entre des objets structurés continus qui a été acceptée dans la revue Algorithms [1].

Développement de la théorie de Gromov-Wasserstein

Thèmes de recherche: transport optimal, assignment quadratique, optimisation non-convexe, optimisation sur l'espace des mesures

Article principal [2]

La théorie de Gromov-Wasserstein permet d'appliquer le transport optimal dans le cas où les points des supports des mesures de probabilités ne font pontentiellement pas partie d'un même espace métrique. Cette théorie permet de trouver des relations entre des ensembles de données hétérogènes ou encore de définir une notion de distance entre des données relationnelles ce qui est utile notamment pour la correspondance de formes.

Cependant en pratique la théorie de Gromov-Wasserstein est difficile à mettre en oeuvre: elle requiert de résoudre un problème quadratique non convexe qui est difficile à approximer.

La deuxième contribution de ma thèse a été de montrer qu'on pouvait résoudre simplement et rapidement le problème de Gromov-Wasserstein entre des mesures de probabilité uni-dimensionnelles. Plus précisément j'ai montré que le problème pouvait se résoudre de manière quasi-linéaire par un algorithme de tri, ce qui est équivalent à la complexité du transport optimal classique sur la droite réelle.

Cette propriété permet de définir une nouvelle distance de transport optimal, appelée Sliced Gromov-Wasserstein. L'idée centrale est de projeter nos données sur des droites aléatoires et de moyenner les distances de Gromov-Wasserstein uni-dimensionnelles qui sont simples à résoudre. Ces travaux ont été publiés et présentés à NeurIPS 2019 [2] (top 2 des conférences en Machine Learning) ainsi que lors du GdR ISIS lors du séminaire "Transport optimal en apprentissage statistique et traitement du signal".

Transport optimal pour les données hétérogènes

Thèmes de recherche: transport optimal, adaptation de domaines, programmation bilinéaire

Article principal [3]

Le transport optimal, dans sa formulation classique, demande de définir une fonction de distance entre les points des différentes distributions de probabilités. Cette restriction est importante lorsque les points ne font pas partie d'un même espace métrique et notamment lorsqu'ils sont dans différents espaces euclidiens. Dans ce cas une solution est d'utiliser la théorie de Gromov-Wasserstein.

La troisième contribution majeure de ma thèse a été de montrer que la théorie de Gromov-Wasserstein était limitée lorsqu'il s'agissait de données hétérogènes, comme par exemple des images de résolutions différentes. L'inconvénient principal de cette théorie est qu'elle ne permet pas de prendre en compte les relations entre les variables, ou les features, des points. J'ai donc proposé un nouveau problème de transport optimal, le *co-transport optimal*, afin de remédier à ces limitations.

J'ai montré que ce nouveau problème était intéressant pour des problèmes difficiles comme l'adaptation de domaines pour des données hétérogènes ainsi que la reconnaissance d'images semi-supervisé. Ces travaux ont été publiés et présentés à NeurIPS 2020 [3].

Projets de recherche

Je collabore sur plusieurs projets de recherche qui sont présentés ci-dessous:

- **MASSILIA**

Matrices spectral structures in graph learning and its applications

ANR Jeune chercheur (soumise en décembre 2020).

Expert externe pour le Working Package : *Graph signal processing with optimal transport on spectral features*.

Ce projet a pour but d'aborder les problèmes actuels liés à l'apprentissage des graphes et à ses applications d'une manière unifiée centrée sur la décomposition spectrale de la matrice laplacienne des graphes.

- **ACADEMICS**

2020-présent

Machine learning & data science for complex and dynamical models.

Projet ANR.

Ce projet a pour but de développer de nouveaux cadres théoriques et des algorithmes d'apprentissage dans deux domaines scientifiques impliquant des données complexes : la modélisation du climat et la compréhension quantitative des systèmes sociaux.

- **OATMIL**

2017-2020

Optimal transport for machine learning.

Projet ANR.

Ce projet a pour but de proposer de nouveaux concepts et de nouvelles méthodologies pour l'exploitation de grandes collections de données en croisant les outils et les idées issus du transport optimal et du machine learning.

Interventions orales

- *Neural information processing systems (NeurIPS), Vancouver, Canada*

2020

Présentation du papier *CO-Optimal Transport* lors de la conférence (session poster).

- *Neural information processing systems (NeurIPS), Vancouver, Canada*

2019

Présentation du papier *Sliced Gromov-Wasserstein* lors de la conférence (session poster).

- *International Conference on Machine Learning (ICML), Long Beach, USA*

2019

Présentation du papier *Optimal Transport for structured data with application on graphs* (spotlight de 5 min).

- *Transport optimal en apprentissage statistique et traitement du signal, GdR ISIS, Paris, France*

2019

Présentation des travaux du papier *Sliced Gromov-Wasserstein*.

Séminaire du GdR ISIS sur le Transport Optimal (intervention de 20 min).
url: <http://www.gdr-isis.fr/index.php?page=reunion&idreunion=395>

• *Qarma laboratory, Université Aix-Marseille, France* 2019

Présentation du Transport Optimal sur graphes lors d'un séminaire d'équipe.

Invité par Valentin Emiya (intervention de 45 min).

• *CAp Conference, Rouen, France* 2018

Présentation du papier *Fused Gromov-Wasserstein Distance* (intervention de 15 minutes).

Recherche reproductible et logiciel libre

• *Python Optimal Transport (POT)*

Librairie Python pour le transport optimal.

Contributeur et co-auteur d'un papier décrivant la librairie (sous review à JMLR, révision mineure demandée)

Plus de 160k téléchargements (Pypi+conda).

29 citations, 21 contributeurs, intégration continue Travis.

url: <https://github.com/rflamary/POT>.

• *Code et données du papier* Optimal Transport on structured data with application on graphs

Librairie Python. Mainteneur. <https://github.com/tvayer/FGW>

• *Code et données du papier* Sliced Gromov-Wasserstein

Librairie Python. Mainteneur. <https://github.com/tvayer/SGW>

• *Code du papier* Learning Convolutional Neural Networks for Graphs

Librairie Python pour le papier *Learning Convolutional Neural Networks for Graphs* par Mathias Niepert, Mohamed Ahmed and Konstantin Kutzkov. ICML. 2016. Mainteneur. <https://github.com/tvayer/PSCN>

Service à la communauté

• *Reviewer pour des conférences internationales* 2019-2020

J'ai été reviewer pour la conférence *Neural Information Processing Systems (NeurIPS)* en 2020 (top 10% des reviewers les mieux notés) et en 2019, ainsi que pour la conférence *the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)* en 2019.

• *Reviewer de journaux* 2020

J'ai occasionnellement été reviewer pour *Information and Inference (IMAI)* en 2020.

PUBLICATIONS

Cette section détaille les différentes publications obtenus lors de mes trois années de thèse.

Conférences internationales

[1] Titouan Vayer, Ievgen Redko, Rémi Flamary et Nicolas Courty. *CO-Optimal Transport*. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*. 2020. Article court. Top 2 des conférences en Machine Learning & Intelligence Artificielle. IF: 33.49.

url: <https://papers.nips.cc/paper/2020/hash/cc384c68ad503482fb24e6d1e3b512ae-Abstract.html>

[2] Titouan Vayer, Rémi Flamary, Nicolas Courty, Romain Tavenard et Laetitia Chapel. *Sliced Gromov-Wasserstein*. In *Advances in Neural Information Processing Systems 32 (NeurIPS)*. 2019. Article court. Pages = 14753–14763. Top 2 des conférences en Machine Learning & Intelligence Artificielle. IF: 33.49. Taux d'acceptation: 21.6%.

url: <https://papers.nips.cc/paper/2019/hash/a9cc6694dc40736d7a2ec018ea566113-Abstract.html>

[3] Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard et Nicolas Courty. *Optimal Transport for structured data with application on graphs*. In International Conference on Machine Learning (ICML), volume 97, 2019. Article court. Pages = 6275–6284. Top 6 des conférences en Machine Learning & Intelligence Artificielle. IF: 18.48. Taux d'acceptation: 22.6%.
url : <http://proceedings.mlr.press/v97/titouan19a.html>

Revues internationales avec comité de lecture

[4] Titouan Vayer, Laetitia Chapel, Remi Flamary, Romain Tavenard et Nicolas Courty. *Fused Gromov-Wasserstein distance for structured objects*. In: Algorithms, vol. 13, no. 9, page 212. 2020. Article long. IS: 2.31.
url : <https://www.mdpi.com/1999-4893/13/9/212>.

Conférences Nationales avec comité de lecture

[5] Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard et Nicolas Courty. *Transport Optimal pour les Signaux sur Graphes*. In: Groupe d'Etudes du Traitement du Signal et des Images (GRETSI). 2019.

[6] Titouan Vayer, Laetitia Chapel, Rémi Flamary, Romain Tavenard et Nicolas Courty. *Fused Gromov Wasserstein distance*. In: Conférence sur l'Apprentissage automatique (CAp). 2018.

Citations

Les informations ont été collectées *via* Google Scholar le 08/12/2020.

	Tout	Depuis 2015
Citations	92	92
h-index	3	3
i10	3	3

