

# P'ART-COURT : Visites et Transports Touristiques Personnalisés

Projet répondant à l'appel d'offre « Agorantic : Culture, Patrimoine, Sociétés Numériques »

## Porteurs du projet :

Juan-Manuel Torres-Moreno<sup>1,3,4</sup>  
juan-manuel.torres@univ-avignon.fr

Marie-Sylvie Poli<sup>2,3</sup>  
marie-sylvie.poli@univ-avignon.fr

## Équipe :

Fen Zhou<sup>1,3</sup>  
fen.zhou@univ-avignon.fr

Didier Josselin<sup>1,3,6</sup>  
didier.josselin@univ-avignon.fr

Andréa C Linhares<sup>5</sup>  
andrea.linhares@ufc.br

Mayeul Mathias<sup>1</sup>  
mayeul.mathias@alumni.univ-avignon.fr

<sup>1</sup> Laboratoire Informatique d'Avignon (LIA/UAPV) - France

<sup>2</sup> Équipe Culture et Communication / Centre Norbert Elias UAPV - France

<sup>3</sup> FR Agorantic (UPAV) - France

<sup>4</sup> École Polytechnique de Montréal - Canada

<sup>5</sup> Université Fédérale du Ceará - Brésil

<sup>6</sup> UMR ESPACE UAPV CNRS- France

**Objectifs** : créer un système de recommandation qui propose aux visiteurs des parcours touristiques personnalisés dans une ville. Le système doit s'appuyer sur les acquis du Traitement Automatique de la Langue Naturelle, des techniques d'optimisation et des Réseaux de communication.

**Mots-clés** : Publics de la culture, Pratique culturelles, Système de recommandation, Traitement Automatique de Langues, Recherche Opérationnelle, Routage dans les réseaux, Transport à la demande, Analyse spatiale.

## 1. Introduction

Le projet « P'ART-COURT » consiste en l'accompagnement du visiteur au sein d'une ville afin qu'il profite au maximum de sa visite.

1. Avant la visite : sélection automatique de textes permettant au visiteur de préparer sa visite dans sa langue et selon ses propres critères.
2. Pendant la visite : en fonction du trajet effectué en début de visite, conseiller au visiteur une suite adaptative tenant compte de plusieurs contraintes (temps qu'il souhaite y consacrer, groupe de personnes, les horaires, la disponibilité des monuments, handicaps, etc).
3. Après la visite : permettre au visiteur d'exprimer son avis sur ce qu'il a vu ainsi que de consulter celui des autres (par exemple tripadvisor.com). Ceci permet d'améliorer le système en affinant la connaissance des attentes des visiteurs mais aussi de constituer un corpus de données pour de futurs algorithmes et enquêtes sociologiques.

Ce projet présente des caractéristiques d'un projet éminemment fédérateur et structurant pour les différentes entités qui y sont associées. Le fait que trois laboratoires différents de l'Université d'Avignon et des Pays de Vaucluse soient impliqués démontre clairement son caractère interdisciplinaire. La dimension internationale est aussi présente du fait de la participation d'une chercheuse d'une université brésilienne.

De plus, ce projet s'inscrit dans un élargissement significatif des problématiques explorées par les projets @MUSE et InfoMuse précédemment financés par la FR Agorantic sur les parcours personnalisés dans les musées d'art (qui ont abouti à une publication dans une conférence internationale ainsi qu'une thèse pour

poursuivre ces travaux).

## 2. Description du Projet « P'ART-COURT »

Afin de faciliter les visites dans une ville, nous proposerons des parcours de visite personnalisés pour chaque visiteur en fonction de ses envies et contraintes. Nous supposons que les visiteurs installeront une application spéciale fournie par l'office de tourisme sur leur smart phone ou tablette. Avec cette application, le visiteur est en mesure d'obtenir les informations sur les sites culturels de la ville comme leur description, leur popularité, les collections présentées (dans le cas des musées), le nombre de personnes présentes à un moment donné, etc. La prise en compte d'événements culturels spontanés peut modifier les parcours proposés. L'application peut proposer une solution statique (sans changement d'avis de l'utilisateur pendant la visite) ou dynamique (changement du parcours en fonction du retard pendant le parcours, des évolutions des envies du visiteur, etc). Grâce aux fonctionnalités de géolocalisation, l'application peut guider le visiteur à travers la ville et permettre d'indiquer la position des différents membres d'un groupe à un guide ou un responsable des publics.

Prenant appui sur ces informations, notre application proposera le circuit de visite optimisé en fonction des besoins des visiteurs.

Nous envisageons deux scénarii :

1. Dans le premier scénario, nous supposons que les visiteurs choisissent librement les monuments qu'ils veulent voir en fonction de leurs goûts et de la popularité de ces derniers (l'application peut également suggérer une liste de sites qui correspondent à ses goûts pour l'aider dans sa décision).
  - L'application fournira un circuit de visite en s'adaptant aux envies exprimées mais aussi aux contraintes des monuments (disponibilité, horaires, type de transport).
  - Lors de sa visite, le visiteur peut changer d'avis et visiter un site qui n'est pas inclus dans le circuit initialement proposé. Quand cela se produit, notre application va récupérer la position actuelle du visiteur et calculer un nouveau circuit de visite incluant les sites qui n'ont pas encore été visités.
2. Dans le second scénario, nous supposons que la ville est très grande alors que l'emploi du temps du visiteur est très serré (par exemple, un visiteur est en transit à Paris et doit prendre son train dans 3 heures, mais il ne veut pas manquer l'occasion de visiter quelques endroits culturels importants). Il y a tellement de monuments qu'il ne lui est pas possible de tous les visiter.
  - L'application classe les monuments en fonction des goûts du visiteur puis propose un circuit de visite qui peut être parcouru dans le temps imparti tout en maximisant la satisfaction globale apportée par la visite des monuments.

## 3. Objectifs et résultats attendus

### 3.1 Définition de la problématique

Le problème de la personnalisation du circuit touristique (*Personalised Sightseeing Tour*, PST) peut être considéré comme une variante du problème du voyageur de commerce (*Travelling Salesman Problem*, TSP), un problème NP-difficile (avec une complexité super-polynomiale - Gutin, 2006) qui est le suivant : compte tenu d'un ensemble de villes et des distances entre chaque villes, le problème est de trouver le plus court chemin qui passe par chaque ville exactement une fois et revient à la ville de départ. Dans notre problème, chaque monument peut être considérée comme une ville et l'objectif est soit de réduire le temps de trajet entre les monuments (dans le 1<sup>er</sup> scénario) soit de maximiser la satisfaction globale des monuments visités ou même plusieurs critères pris en compte simultanément (2<sup>me</sup> scénario). A ce problème s'ajoute la prise en

compte des éléments d'intérêt à différentes échelles. En effet, il est possible de gérer des granularités différentes selon l'échelle du problème. Les parcours optimisés peuvent ainsi reposer sur l'attrait des musées ou des sites culturels, eux-même composés d'œuvres plus ou moins cotées, se trouvant dans des salles spécifiques (typés par exemple pour les peintres impressionnistes). Tout ces niveaux peuvent intervenir dans les choix prônés des recommandations, selon les deux scenarii considérés. Le problème du PST est différent de celui du TSP sur quatre aspects :

1. Ce ne sont pas tous les monuments qui doivent être visités et les points d'entrée et de sortie de la ville peuvent être différentes.
2. Les monuments peuvent avoir des horaires d'ouverture (les musées par exemple) Pour cela, à notre problème s'ajoute une autre dimension d'optimisation : l'ordonnancement.
3. La fonction objectif est différente puisqu'il faut trouver le sous ensemble des monuments qui puissent être visités (contrainte de temps limité) et apporte la meilleure satisfaction pour le visiteur (contrainte sur les envies).
4. Le PST est multi-granularité (ou multi-échelle) et fait appel à des structures de graphes emboîtés relativement complexes
5. Pendant la visite, le visiteur peut visiter des monuments qui ne sont pas proposés par le SDR. Dans ce cas, la méthode sera en mesure de proposer dynamiquement un circuit de visite en tenant compte de l'état actuel de la visite. Par contre, TSP et PST ne sont pas applicables pour un scénario dynamique.

### 3.2 Méthodologies de résolution et verrous sous-jacents

Pour résoudre notre problème du PST dans une petite ville, nous pouvons utiliser l'optimisation linéaire pour modéliser le problème et le résoudre de manière optimale avec l'algorithme du SIMPLEXE. Les contraintes comme celles du jour de la semaine, l'horaire de visite ou la mobilité réduite seront modélisées comme des pénalités à être associées aux monuments. Les projets @MUSE et InfoMuse ont démontré la possibilité d'utiliser ces outils pour résoudre des problématiques similaires (Mathias et al. 2014) dans les musées, nous nous baserons sur ces résultats pour développer de nouveaux modèles applicables à la problématique du PST.

Dans le cas de grandes villes ou de parcours dynamiques, l'utilisation de l'optimisation linéaire conduit à des temps de calcul trop importants et des algorithmes heuristiques devront être imaginés pour résoudre le problème PST avec des solutions sous-optimales acceptables générées dans des temps convenables. Les travaux réalisés dans le cadre des projets @MUSE et InfoMuse ont permis le développement d'heuristiques satisfaisantes et montrent que ces méthodes peuvent apporter des solutions au problème du PST. Plus précisément, nous envisageons l'utilisation dans un premier temps d'heuristiques gloutonnes puis de méta-heuristiques. Ces méthodes sont définies comme un ensemble de concepts algorithmiques et de structures de données génériques pour le développement et l'application d'algorithmes heuristiques et sont très utilisées pour la résolution satisfaisante des problèmes NP-difficiles (Dorigo et Stützle, 2004) dans un temps polynomiale (Arora, 1998). Dans le contexte de TSP, la recherche Tabou a donné de bons résultats dans les problèmes de routage et pourrait être appliquée. Nous nous intéresserons également aux méthodes bio-inspirées, comme les colonies de fourmis (*Ant Colony Optimization*) et les algorithmes génétiques (Chevrier et al, 2008). Ces méta-heuristiques, grâce à leur structure générique ont été utilisées avec succès pour résoudre des problèmes complexes et variés.

Il y a également plusieurs problématiques liées au traitement du langage naturel écrit. Il s'agit dans un premier temps d'analyser les préférences du visiteur, qui seront exprimées en langage naturel. Cette phase d'analyse comporte des caractéristiques d'analyse et de Recherche d'Information dans des corpus multilingues. Une fois les éléments saillants détectés, on pourra proposer des résumés personnalisés pour chaque visiteur. Afin de s'affranchir des limites technologiques imposés par la taille des dispositifs (petits écrans), les résumés peuvent être produits en faisant appel aux algorithmes de compression de phrases. Enfin, la fouille de texte et la détection d'opinion permettront d'analyser et de classer les avis exprimés par les visiteurs sur les monuments vus au cours de leur visite.

En complément, des solutions de transport adéquat à la demande sont aussi à prévoir pour acheminer les visiteurs d'un site à un autre (Castex & Josselin, 2007). Un suivi des visiteurs peut également être assuré et analysé afin d'observer les points d'attractivité en fonction de la structure urbaine et de l'organisation des sites (Space Syntax). Il s'agit en quelque sorte d'un espace micro-géographique culturel dont la compréhension du fonctionnement peut servir à terme à mieux organiser les déplacements culturels dans la ville mais également à améliorer l'usage de l'information au cours de la visite et après celle-ci.

Nous comptons produire au moins un article en congrès pour montrer les résultats préliminaires des algorithmes développés. Un autre résultat attendu est la préparation au montage d'un projet plus ambitieux à un niveau national ou international.

#### **4. Caractères innovants du projet « P'ART-COURT »**

Comme vu dans l'introduction, ce problème peut être posé comme une généralisation du problème de visites personnalisées dans les musées avec un changement de granularité, la ville comme un musée à ciel ouvert. Ce changement d'échelle amène de nouvelles problématiques dans plusieurs domaines :

- La personnalisation des parcours en fonction des goûts mène à la problématique de l'optimisation de critères subjectifs et donc à une réflexion sur la captation, l'interprétation et la modélisation des attentes exprimées des visiteurs.
- Contrairement aux musées, il existe une multitude de chemins menant d'un site culturel à un autre dans une ville, deux itinéraires peuvent amener le visiteur à découvrir la ville différemment. Ainsi un des points innovants de ce projet est de considérer toute la ville comme un lieu culturel (et non pas seulement les sites et monuments).
- La dimension dynamique des parcours proposés, via les scénarii, ajoute un verrou supplémentaire, car elle nécessite de disposer d'information de localisation des visiteurs ou une interaction avec ceux-ci permettant d'intégrer des modifications de contexte réorientant les parcours préconisés.
- Le caractère multi-scalaire des informations amène à manipuler des graphes multi-niveaux, d'une complexité plus grande
- Le fait que la visite se déroule dans une ville amène aussi des problématiques de transport optimisé qui doivent agir en combinaison avec les parcours thématiques culturels proposés.

#### **5. Dimension interdisciplinaire**

Le projet « P'ART-COURT » pose des verrous scientifiques pluridisciplinaires importants :

1. Avant la visite : il faut une sélection automatique de textes permettant au visiteur de préparer la visite dans la langue ou dans un bouquet de langues de son choix et selon des critères qui lui sont propres. Le problème scientifique concerne entre autres, le résumé automatique personnalisé multi documents et multi lingue (Torres-Moreno et al 2002, Torres-Moreno, 2014).
2. Pendant la visite : en fonction du trajet effectué en début de visite il faudra conseiller au visiteur une suite adaptative tenant compte du temps qu'il souhaite y consacrer. La problématique scientifique abordée concerne les systèmes de recommandation, les méthodes d'optimisation et les réseaux de communication.
3. Après la visite : on devra permettre au visiteur d'exprimer son avis sur ce qu'il a vu, et aussi et surtout consulter celui des autres. La problématique scientifique concerne la fouille de textes et plus précisément, la détection automatique d'opinion. Ce pavé peut bien entendu être utilisé avant la visite sur le versant de la consultation. Il a des questions ouvertes concernant la constitution d'un corpus

d'apprentissage selon les préférences des visiteurs qui seront étudiés lors d'un projet ultérieur.

4. Le traitement de l'information en temps réel, les communications réseaux et les algorithmes d'optimisation et de transport à la demande. À ce niveau, on peut imaginer des échanges d'information entre les visiteurs et les système, mais également entre les visiteurs eux-mêmes (dans la ville) ; on peut ajouter également l'envoi d'informations clés au passage de certaines bornes, informations qui pourraient être des avis d'autres visiteurs, mis à jour dynamiquement.

## 6. Positionnement dans l'Agorantic

Le projet est situé dans l'Axe 1 de l'Agorantic : « Culture et numérique ». L'équipe du projet est constituée de chercheurs dans différents domaines :

- Juan-Manuel Torres, Maître de Conférences HDR et responsable de la thématique Traitement Automatique du Langage Naturel Écrit (LIA)
- Marie-Sylvie Poli, Professeur en Sciences de l'Information et de la Communication et chercheur en Muséologie et médiation muséale (ECC/CNE)
- Mayeul Mathias, doctorant boursier du MESR interviendra sur le domaine des systèmes de recommandation touristique.
- Fen Zhou, Maître de Conférences (LIA), sur le routage dans les réseaux.
- Didier Josselin, Directeur de Recherche CNRS (UMR ESPACE) à propos du transport optimisé et des multi-graphes

Le LIA est en effet spécialiste en Traitement Automatique du Langage Naturel ainsi qu'en recherche opérationnelle et travaille sur plusieurs applications telle que le transport à la demande et l'ordonnancement.

## 7. Partenariats extérieurs (en cours et à venir)

Du côté de collaborations internationales, Andréa Carneiro Linhares, MCF à l'Université Fédérale du Ceará (Brésil) et spécialiste en recherche opérationnelle est une des associées à ce projet. La coopération avec Mme Linhares a commencé dans le cadre de séjours scientifiques (2009-2011) ; un post-doctorat de août 2012 à juillet 2013 et les projets @MUSE (2013) et INFOMUSE (2014) financés par FR Agorantic, comme une première approche du sujet présenté dans le projet P'ART-COURS dans le contexte des musées. D'autres coopérations avec le Brésil sont à l'ordre du jour : un stage de Master II (UFC/UAPV, février -juin 2015) et le montage d'un projet internationale concernant le TALN avec la participation de M Thiago Pardo de l'Université de São Carlos (UFS/Brésil) dans le cadre du programme CAPES/COFECUB en 2015.

Des contacts ont été pris avec des partenaires privés ou associatifs pour rendre ce projet plus visible et augmenter ses potentialités d'usage (réunion prévue le 18 décembre avec association Avignon le off, société Semaweb rencontrée le 16 décembre). Par ailleurs, le projet est identifié comme porteur par le label French Culture Tech et la SATT de Marseille, à laquelle l'UAPV est associée accompagne la création d'une entreprise portée par Didier Josselin sur l'optimisation en ligne des transports optimisés dynamiques. Le projet pourrait ainsi bénéficier prochainement d'un poste d'ingénieur en développement pour finaliser la partie transport optimisé qui concourt à cette application.

## 8. Budget prévisionnel

Pour mener à bien le projet « P'ART-COURT », nous demandons **12 990** euros qui seront utilisés comme suit :

- 4.790,00 (5 mois x 2 personnes x 479,00) pour deux stages de cinq mois d'étudiants de Master (Informatique et Sciences humaines).
- 2 500,00 € pour participer à un congrès scientifique international afin de publier nos résultats
- 1 200,00 € pour les différentes missions.
- 1 500,00 € pour un ordinateur portable.
- 3 000,00 € pour accueillir Andréa Carneiro Linhares de l'Université Fédérale du Ceará pendant deux mois qui viendrait au LIA pour le développement d'heuristiques de recherche de parcours.

## 9. Références

1. Castex E., Josselin D., Temporalités éclatées : la réponse des transports à la demande aux nouvelles formes de mobilité, *Espace Population et Sociétés*, n ° 2-3, pp. 433-447, 2007
2. Chevrier R., Castex E, Canalda P., Chatonnay P., Josselin D., Un algorithme génétique pour le transport à la demande en convergence., *Revue Internationale de Géomatique*, pp. 239-265, 2008
3. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L. & Stein, C., *Introduction to Algorithms*. 3a Ed. Cambridge, USA: MIT Press, 2009.
4. Dorigo, M. & Stutzle, T., *Ant Colony Optimization*. Cambridge, USA: MIT Press, 2004. Goldberg, D.E., *Genetic Algorithm in Search, Optimization and Machine Learning*. Reading, USA: Addison-Wesley Publishing Company, 1989.
5. Garaix T., Artigues C., Feuillet D., Josselin D., Optimization of occupancy rate in dial-a-ride problems via linear fractional column generation, *Computers and Operations Research*, Vol.38, N°10, pp.1435-1442, October 2011, Science Direct, 0305-0548.
6. Garaix T., Artigues C., Feuillet D., Josselin D., Vehicle routing problems with alternative paths: An application to on-demand transportation. *European Journal of Operational Research*, Vol. 204, N°1, pp.62-75, 2010.
7. Konak, A.; Coit, D.W. & Smith, A.E., *Multi-objective optimization using genetic algorithms: A tutorial*. *Reliability Engineering and System Safety*, 91(9):9921007, 2006.
8. Reeves, C.R., *Modern Heuristic Techniques for Combinatorial Problems*. New York McGraw-Hill, 1995.
9. Heitor Silvério Lopes and Luiz Carlos de Abreu Rodrigues and Maria Teresinha Arns Steiner, *Meta-Heurísticas em Pesquisa Operacional*, Curitiba, Ompipax, 2013
10. Gutin, G.; Punnen, A. P. (2006), *The Traveling Salesman Problem and Its Variations*, Springer, ISBN 0-387-44459-9.
11. Arora, S., Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems, *Journal of the ACM* 45 (5): 753–782, 1998.
12. Torres-Moreno, Juan-Manuel. *Automatic Text Summarization*, Wiley & Sons, 2014.
13. Poli, Marie-Sylvie & Gottesdiener, Hana. « Les titres d'expositions : sur quoi communiquent les musées », p.81-88, in *Culture et musées* 11, 2008.
14. Juan-Manuel Torres-Moreno, Patricia Velázquez-Morales, Jean-Guy Meunier. Condensés de textes par des méthodes numériques. *JADT*, pp 723-734, 2002.
15. Mayeul Mathias, Assema Moussa, Fen Zhou, Juan-Manuel Torres-Moreno, Marie-Sylvie Poli, Didier Josselin, Marc El-Bèze, Andréa Carneiro Linhares, Françoise Rigat. « Optimisation using Natural Language Processing: Personalized Tour Recommendation for Museums ». *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems* pp. 439–446