
Recommandation Dynamique dans les Graphes Géographiques

Elie Guàrdia-Sebaoun¹, Vincent Guigue¹, Patrick Gallinari¹

Laboratoire d'informatique de Paris 6 (LIP6)
CNRS UMR 7606 Sorbonne-Universités, UPMC, Paris 6, France
{elie.guardia-sebaoun,vincent.guigue,patrick.gallinari}@lip6.fr

ABSTRACT. Le domaine de la recommandation gagne peu à peu la vie réelle : à leur arrivée dans un musée ou dans une ville à visiter, beaucoup de touristes consultent encore divers guides. Au cours des dernières années, l'apparition et la démocratisation des appareils numériques nomades (smartphones, tablettes) ont poussé les touristes vers des guides dématérialisés. Ainsi, les musées se sont équipés de nouveaux terminaux et d'applications de personnalisation de visite (par exemple le Louvre-Lens, le Guggenheim et la Cité des Sciences et de l'Industrie). Au coeur de ces applications se trouve un système de recommandation. Cependant, ce nouveau domaine applicatif comporte plusieurs spécificités : il s'agit de recommander non seulement des oeuvres ou des points d'intérêt mais aussi des parcours thématiques et personnalisés. Du point de vue scientifique, les enjeux sont nombreux : transposer le problème de la recommandation à des séquences, prendre en compte les aspects dynamiques pour affiner les profils au fil des visites, intégrer des contraintes de temps de visite ou d'engorgement des certaine salles.

Dans ce travail, nous nous intéressons à la recommandation dynamique de points d'intérêt (POI) dans des graphes géographiques. Nous travaillons sur plusieurs jeux de données : un jeu issu d'un musée et trois jeux de photos géolocalisées collectées sur Flickr. Dans tous les cas les données sont séquentielles, elles décrivent les parcours touristiques pour des utilisateurs. Nous proposons trois modèles de références basés sur une modélisation probabiliste des déplacements observés puis nous introduisons un modèle latent permettant de mieux passer à l'échelle en projetant les graphes dans un espace continu. Etant donnée l'originalité de la tâche, nous introduisons également plusieurs métriques d'évaluation dédiées à ce cadre applicatif.

KEYWORDS: Recommandation dans les graphes, Dynamique, Projection latente de graphes

DOI:10.3166/RMPD.22.1-11 © 2014 Lavoisier

1. Introduction

A son arrivée dans un musée ou une nouvelle ville, le visiteur peut être dépassé par la masse de contenus proposée. Les musées offrent à leurs visiteurs des parcours thématiques et des audioguides, les guides touristiques et les offices du tourisme recensent les points d'intérêt régionaux et les visiteurs définissent leur parcours à venir en fonction de ces sources d'information. Pour les assister dans leur visite, de nouvelles applications de recommandation sont en cours de développement (Setten *et al.*, 2004 ; Picot Clemente, Bothorel, 2013). En effet, le problème de la recommandation de parcours touristique présente plusieurs particularités : il s'agit de recommander non seulement des oeuvres ou des points d'intérêt mais aussi des itinéraires thématiques et personnalisés. Du point de vue scientifique, les enjeux sont nombreux : transposer le problème de la recommandation à des séquences, prendre en compte les aspects dynamiques pour affiner les profils au fil des trajets effectifs des utilisateurs, intégrer des contraintes de temps de visite ou d'engorgement de certaines salles.

Au cours des dernières années, l'apparition et la démocratisation des appareils numériques nomades (smartphones, tablettes) ont poussé les musées à revoir leurs assistants de visite et de nouveaux outils sont proposés aux visiteurs du Louvre-Lens, des musées Guggenheim ou de la Cité des Sciences et de l'Industrie par exemple. Au coeur de ces dispositifs, un système de recommandation propose une expérience personnalisée. Parallèlement, la démocratisation des smartphones et de l'Internet mobile a donné naissance à des applications comme Foursquare¹ ou Flickr². La première est un réseau social permettant de *check in* dans des lieux, de les noter et de laisser son avis. La seconde est un réseau social de partage de photos, souvent géolocalisées. L'une comme l'autre génèrent donc des traces d'utilisateurs dans un graphe géographique. L'intérêt de ces données pour les applications de recommandation a été mis en évidence dans (Van Canneyt *et al.*, 2011 ; Baraglia *et al.*, 2013).

Les systèmes développés récemment dans les musées (Bohnert, 2010 ; Karimi *et al.*, 2012) reposent fortement sur la recommandation de contenu : l'idée est d'élargir le champ autour d'une oeuvre en proposant à l'utilisateur les ressources thématiquement proches, en lui montrant aussi des oeuvres actuellement non exposées et en lui fournissant des informations connexes sur l'auteur de l'oeuvre. Dans ce travail, nous nous intéressons à la recommandation dynamique de points d'intérêt (POI) dans les graphes géographiques : nous intégrons la notion de parcours et de personnalisation. La question devient : "*étant donné un début de parcours, qu'est ce qui est le plus pertinent pour la suite de la visite connaissant les séquences des utilisateurs passés?*" (Baraglia *et al.*, 2013).

Le système est intrinsèquement dynamique puisque les propositions évoluent au fil de la visite en tirant parti de l'historique grandissant pour affiner le profil de l'utilisateur. Nous travaillons sur plusieurs bases de données : une base issue d'un musée et

1. <http://fr.foursquare.com/>

2. <https://www.flickr.com/>

plusieurs bases de photos géolocalisées collectées sur Flickr. Dans tous les cas les données sont séquentielles, elles décrivent les parcours touristiques pour des utilisateurs. Nous proposons plusieurs modèles de références basés sur une modélisation probabiliste des déplacements observés puis nous introduisons un modèle latent permettant de mieux passer à l'échelle en projetant les graphes dans un espace continu. Etant donnée l'originalité de la tâche, nous introduisons également plusieurs métriques d'évaluation dédiées à ce cadre applicatif.

La seconde section de cet article présente les travaux connexes et la troisième, les données utilisées. Les modèles que nous avons utilisés ainsi que les métriques d'évaluation originales sont décrits dans la quatrième partie. Enfin, la cinquième contient l'analyse des résultats obtenus lors des expériences.

2. Travaux Connexes

Les systèmes de recommandation exploitent classiquement les traces virtuelles des utilisateurs sur internet en listant les produits que ce dernier a visionnés, notés ou achetés. Ici l'espace décrit est concret, les traces sont géographiques et les cibles à mettre en valeur sont des points d'intérêt (POI) dans des villes ou des musées. Les premiers travaux scientifiques dans les musées se focalisent sur les systèmes de localisation. (Bohnert, 2010) s'intéresse au musée de Melbourne et liste par exemple les verrous techniques à la personnalisation des parcours dans les musées en insistant sur le potentiel des smartphones pour dépasser les schémas traditionnels de visite. (Karimi *et al.*, 2012) propose une architecture matérielle de localisation basée sur les puces RFID.

Nous distinguons usuellement trois familles d'algorithmes de recommandation (Burke, 2007) : ceux basés sur les similarités de contenu, ceux basés sur la connaissance (adéquation entre item et utilisateur à base de règles) et les systèmes collaboratifs. Le contenu est utile pour élargir le champ autour d'une oeuvre en proposant des ressources thématiquement proches et en fournissant des informations connexes sur l'auteur de l'oeuvre. Les approches de la seconde famille demandent un profil utilisateur très fin, reposant souvent sur de longs questionnaires incompatible avec les visites de musées d'après les conservateurs que nous avons interrogés. Les techniques de factorisation matricielle utilisées en recommandation permettent de représenter les utilisateurs et les items dans le même espace latent, décrivant ainsi de la même façon les caractéristiques de l'item et les goûts de l'utilisateur (qu'ils soient positifs ou négatifs) (Koren *et al.*, 2009). Ces approches, plébiscitées ces dernières années, proposent plusieurs raffinements efficaces comme l'introduction d'un biais sur les utilisateurs (certains étant plus sévères que d'autres) ou sur les items (ce qui équivaut à prendre en compte leurs réputations). Il est aussi possible d'intégrer des informations contextuelles (e.g. localisation, âge et sexe de l'utilisateur). Cependant, la nature même des données sur lesquelles nous travaillons pose problème pour implémenter des techniques collaboratives : le démarrage à froid est particulièrement difficile (Schafer *et al.*, 1999) et pour proposer une recommandation pertinente au fur et à mesure de la

visite, il faudrait entièrement recalculer la factorisation après l'activation de chaque POI.

Nous avons opté pour un positionnement de cette étude sur les aspects dynamiques de la recommandation : nous prenons en compte la notion de séquence dans la visite, à la manière d'une *playlist* musicale (Chen *et al.*, 2012), et nous cherchons à prédire la prochaine étape du visiteur pour maximiser son intérêt. Cette tâche a été proposée pour la première fois dans le domaine touristique dans (Baraglia *et al.*, 2013). Si notre tâche reste un problème de recommandation, elle peut être vue comme la diffusion au cours du temps d'un visiteur dans l'espace des POIs (i.e au pas de temps suivant, quels sont les POIs potentiels pour un visiteur). Dans la littérature, ces modèles sont principalement utilisés pour répondre au problème de *brand adoption* (Bass, 1969). Deux familles de méthodes ont émergé dans les années 2000 pour répondre à ce problème : les Cascades Indépendantes (Goldenberg *et al.*, 2001 ; Saito *et al.*, 2008) et les modèles à seuil linéaire (Kempe *et al.*, 2003). Cependant, ces deux familles de méthodes se sont heurtées aux mêmes types de problème : les calculs sont très lourds lorsque toutes les cascades sont considérées ensemble et l'obtention d'un graphe social fiable est une tâche quasiment impossible. Pour répondre à cette problématique, deux nouvelles familles de méthodes ont vu le jour. Nous avons d'un côté les algorithmes de prédiction de lien, d'abord avec des méthodes gloutonnes (Gomez Rodriguez *et al.*, 2010 ; Myers, Leskovec, 2010), puis plus récemment à l'aide de méthodes basées sur un calcul d'entropie de transfert (Ver Steeg, Galstyan, 2013). La seconde famille s'affranchit du graphe à proprement parler en utilisant des méthodes d'apprentissage statistique pour prédire le volume de diffusion après un temps donné (Yang, Leskovec, 2010). Plus récemment, (Bourigault *et al.*, 2014) propose de projeter le graphe dans un espace latent continu et de modéliser la diffusion de l'information à l'aide d'un modèle de diffusion de la chaleur.

Nous reprenons à notre compte cette dernière approche qui permet de beaucoup mieux passer à l'échelle que des méthodes de ranking dans les graphes (Baraglia *et al.*, 2013) : une fois le graphe des POIs projeté dans l'espace latent, il suffit à chaque instant de repérer l'utilisateur dans cet espace pour proposer très facilement une recommandation. Les aspects collaboratifs, la mémoire des parcours précédents, sont intégrés lors de l'apprentissage de la fonction de transfert des POIs et la formulation est très flexible : il est possible de définir des contraintes de proximité des POI latents basées sur leurs contenus, leurs enchaînements dans le temps etc...

3. Jeux de données

Nous utilisons deux jeux de données modélisant les déplacements de touristes dans des graphes géographiques, le premier correspond à leurs trajets dans un musée, le second dans des villes. Le jeu de données (AMMICO, 2012) contient les traces (ou trajets) des visiteurs dans l'exposition Great Black Music à la Réunion³. L'exposition

3. <http://www.greatblackmusic.fr/fr/lexposition/>

est divisée en six salles (dont 4 permettant la géolocalisation) s'enchaînant dans un ordre défini. Les traces définissent ainsi un graphe dont la structure est particulière. Nous faisons l'hypothèse qu'il s'agit d'un cas classique pour les traces de musées.

Le jeu de données Flickr contient les traces de touristes dans trois villes italiennes (Pisa, Roma et Firenze) de tailles différentes (Baraglia *et al.*, 2013). Les trajets s'effectuant dans des villes, leur structure est intrinsèquement moins rigide que celle du jeu de données (AMMICO, 2012) : le graphe associé est moins parcimonieux (i.e. plus connecté).

Tableau 1. Descriptif des jeux de données explicitant leur nombre d'items (POI) et de parcours-utilisateur (traces)

Jeu de données	#POI	#traces
AMMICO	69	3713
Pisa	110	992
Roma	490	12565
Firenze	888	5984

Le tableau 1 montre que le nombre de traces dans les jeux de données n'est pas proportionnel au nombre de POI. Le faible nombre de traces du jeu de données Firenze en fait un cas fortement défavorable à l'apprentissage de n'importe quel modèle.

4. Modèles et Mesures

Dans la suite, nous utiliserons les notations suivante. Les items sont indicés par i et regroupé dans un ensemble : $I = \{x_i\}_{i=1,\dots,N}$. Une trace t_k est constituée d'une séquence d'item : $t_k = \{x_i\}$, et l'ensemble des traces sont regroupées dans : $\Theta = \{t_k\}_{k=1,\dots,K}$. Par analyse des traces, nous construisons la matrice de contingence des transitions $T \in \mathbb{N}^{N \times N}$. Chaque modèle repose sur une fonction de scoring des items $\Phi(x_i, x_j)$ qui donne le score de x_j sachant que le visiteur se trouve en x_i . La prédiction de l'item suivant correspond donc à :

$$x_{\text{suivant}} = \arg \max_{x_j \in I} \Phi(x_i, x_j) \quad (1)$$

Nous extrayons entre 1 et 3 items en fonction des métriques d'évaluation utilisées. Nous détaillons lesdites métriques à la fin de cette section.

4.1. Modèles probabilistes de référence

Next Ce modèle se base sur une hypothèse markovienne d'ordre un. Le prochain POI prédit est celui qui maximise la probabilité de transition apprise :

$$\Phi(x_i, x_j) = \frac{\text{Card}(T_{i,j})}{\text{Card}(T)} \quad (2)$$

Future Au lieu de se limiter à une transition x_i, x_j , nous augmentons l'horizon temporel. Nous introduisons une matrice de contingence $C \in \mathbb{N}^{N \times N}$ qui compte toutes les fois où x_j suit x_i : nous conservons la notion d'enchaînement mais éliminons l'hypothèse markovienne d'ordre 1. Le modèle devient :

$$\Phi(x_i, x_j) = \frac{\text{Card}(C_{i,j})}{\text{Card}(C)} \quad (3)$$

Le but de ce modèle est de le rendre plus robuste aux impasses faites durant une visite.

Cooccurrences Nous éliminons maintenant la notion séquentielle pour se baser sur les simples cooccurrences de x_i et x_j dans une trace :

$$\Phi(x_i, x_j) = \frac{\text{Card}(C_{i,j}) + \text{Card}(C_{j,i})}{\text{Card}(C)} \quad (4)$$

Ce modèle augmente la probabilité de sélectionner un items de transition "début de visite" vers "fin de visite".

4.2. *Modèle latent*

A l'instar de (Bourigault *et al.*, 2014), nous nous sommes basés sur un modèle de diffusion de chaleur. En d'autres termes cela revient à considérer qu'entre deux items, le plus proche de la source sera aussi le plus chaud. Dans notre cas, cela signifie que si deux items ont une représentation latente proche, alors ils seront aussi susceptibles d'intéresser un même visiteur.

Nous avons décidé d'ajouter une amélioration proposée par (Chen *et al.*, 2012). Celle-ci consiste à dédoubler chaque représentation d'un POI en deux points pour tenir compte de l'ordre dans lequel les POI s'enchaînent : le premier correspondant à l'entrée, le second à la sortie. Cette amélioration permet bien d'orienter la visite: si un trajet entre trois POI a, b et c est souvent fait dans cet ordre, alors l'entrée de b sera proche de a sa sortie de c, défavorisant ainsi le passage inverse. Dans le cas contraire, si le trajet a-b-c est aussi emprunté que c-b-a, l'entrée et la sortie de b seront proches, ne favorisant ainsi aucun sens.

Soit z_i la représentation de l'entrée d'un POI $x_i \in I$ et z'_i sa sortie, l'optimisation de ce modèle est effectuée par minimisation par descente de gradient de la fonction de coût Δ suivante :

$$\forall t \in \Theta, \text{Argmin}_{z'_i, z_n} \Delta(t) = \sum_{(x_i, x_n) \in t} \sum_{j \in I \setminus x_n} (||z'_i - z_n|| - ||z'_i - z_j||) \quad (5)$$

avec x_n suivant immédiat de x_i

Cette fonction permet d'obtenir des représentations proches pour les POI successifs tandis que les autres POI sont éloignés du point de référence. En inférence, nous utilisons simplement une distance euclidienne dans l'espace latent :

$$\text{Argmin}_{x_j} \Phi(x_i, x_j) = \|z_i - z_j\| \quad (6)$$

En s'affranchissant de la structure de graphe et en passant à une méthode de traitement stochastique, ce modèle permet de s'attaquer à de très gros jeux de données, ce que les modèles précédents étaient incapables de faire. Ce modèle présente d'autres avantages : il est résistant au bruit et surtout particulièrement évolutif. En effet, il est possible d'imaginer diverses contraintes lors de l'apprentissage pour intégrer des aspects thématiques, des contraintes basées sur la proximité physique des items, des enchaînements forcés (ou fortement encouragés)...

4.3. Mesures d'évaluation

Les systèmes de recommandation classiques sont communément évalués sur leur capacité à retourner une liste d'items pertinents, indépendamment de leur rang d'apparition dans la séquence d'un utilisateur. Notre approche étant dynamique, il était nécessaire de prendre en compte la composante temporelle et la notion d'enchaînement, à la manière de (Baraglia *et al.*, 2013). Ainsi, en plus de la précision au Top-3, nous avons proposé deux autres mesures : le MeanRank et le Next-3. Toutes ces mesures sont détaillées ci-dessous.

Top-3 Comme dans le cas de l'évaluation d'un système de recommandation classique, cette mesure calcule la précision de l'apparition du prochain item de la trame dans les 3 premières propositions du système.

MeanRank Le MeanRank calcule la moyenne des rangs attribués par le système à l'objet optimal. Soit $rank(x_n)$ la fonction qui rend le classement de l'item n par notre système pour l'item de référence x_i . Pour toutes les traces de Θ , pour toutes les prédictions possibles dans chaque trace, nous étudions le rang moyen de la prédiction issue de notre système (1 correspond à un système idéal).

$$\text{MeanRank}(\Theta) = \sum_{t_k \in \Theta} \frac{\sum_{x_n \in t_k} rank(n)}{\text{Card}(\Theta) \text{Card}(t_k)}$$

Cette mesure offre une bonne vue d'ensemble des capacités du système de recommandation.

Next-3 Cette mesure calcule la précision pour l'apparition de la première proposition de système dans les 3 prochains items de la trace, lui laissant ainsi une marge d'erreur (oubli d'un POI) et permet donc d'étudier ses capacités dans un cadre dynamique.

5. Résultats

Nous présentons successivement les résultats obtenus par les modèles probabilistes et le modèle latent. Nous proposons ensuite une interprétation graphique de l'espace latent appris.

Les résultats obtenus sont la moyenne de dix expériences successives réalisées avec les mêmes hyperparamètres sur dix répartitions apprentissage/test (90%/10%) des datasets tirées aléatoirement.

5.1. Modèles probabilistes de référence

Les performances des modèles probabilistes sont regroupées dans le tableau 2. Plusieurs conclusions s'imposent sur cette expérience. Premièrement, le modèle *co-occurrences* est systématiquement battu par les autres : la prise en compte de la temporalité dans la décision apporte un vrai gain d'information. Ensuite, les résultats sont meilleurs sur la base AMMICO que sur les données Flickr ce qui s'explique par une plus forte structuration du graphe des trajets pour AMMICO. Les POI y sont regroupés par salles et l'arité des noeuds est beaucoup plus faible que dans les jeux de données Flickr, cela simplifie la décision. En d'autres termes, la visite d'un musée est plus structurée que le trajet d'un touriste dans une ville. Il est donc intuitif que les trajets y soient plus semblables d'un individu à l'autre. Dans le cas du jeu de données AMMICO, le modèle *future* est meilleur que les autres en tout point, ce qui s'explique encore une fois par le faible niveau de bruit dans les données et la structure forte de la visite. En revanche, pour les jeux de données Flickr, le modèle *future* offre de meilleurs résultats pour les mesures *MeanRank* et *Next3* alors que le modèle *next* a un meilleur score en *Top3*. Ceci s'explique par la nature des modèles : le *Top3* est une mesure plus stricte, favorisant le modèle *next* alors que les deux autres permettent de tirer avantage de la capacité du modèle *future* à accepter les impasses lors d'une visite.

Les différences de performances entre les jeux de données Flickr ont deux causes principales: tout d'abord, les méthodes d'évaluation ne tiennent pas compte des différences de taille entre les catalogues d'items. Elles sont donc plus favorables aux jeux de données à petit catalogue, comme Pisa. De plus, comme relevé précédemment, le nombre de traces n'augmente pas proportionnellement au nombre d'items, ce qui rend beaucoup plus dur l'apprentissage sur la base Firenze que sur la base Roma par exemple.

5.2. Modèle latent

Sur le jeu de données AMMICO, le tableau 3 montre que le modèle latent améliore significativement les performances par rapport au modèle *next*, ce qui valide notre approche du problème. Les contraintes sur la succession des POI dans la formulation (5) permettent d'exploiter les forces du modèle *next* tout en introduisant plus de robustesse. Sur les données Flickr, les conclusions sont moins nettes et le modèle latent

Tableau 2. Modèles probabilistes : résultats pour les différentes métriques proposées sur les données AMMICO (traces dans l'exposition Great Black Music) et les trois bases Flickr (traces de touristes dans les villes de Pisa, Roma, Firenze)

Jeu de données	AMMICO			Pisa		
Model	meanRank	Top3	Next3	meanRank	Top3	Next3
Next	10.59	0.32	0.23	23.4	0.44	0.37
Future	17.95	0.13	0.13	18.6	0.39	0.46
Cooccurrences	26.5	0.11	0.12	61.15	0.09	0.12
Jeu de données	Roma			Firenze		
Model	meanRank	Top3	Next3	meanRank	Top3	Next3
Next	78.5	0.34	0.26	288.6	0.18	0.14
Future	48.2	0.26	0.23	178.7	0.13	0.14
Cooccurrences	317.9	0.04	0.05	566	0.02	0.02

est systématiquement dépassé par au moins un modèle probabiliste. Il semble que les contraintes utilisées ne soient pas optimales pour la visite de villes: il sera sans doute nécessaire d'ajouter des contraintes sur la localisation des POI pour renforcer l'espace latent.

Dans tous les cas, le modèle latent apporte une réelle possibilité de passage à l'échelle en inférence: une fois l'espace appris, les calculs de prédictions sont négligeables. Cet aspect de la question est crucial étant donnée l'application cible: il est aisé d'implémenter ces stratégies sur de l'embarqué.

Tableau 3. Modèle latent : résultats pour les différentes métriques proposées sur les données AMMICO (traces dans l'exposition Great Black Music) et les trois bases Flickr (traces de touristes dans les villes de Pisa, Roma, Firenze) avec un espace latent de dimension 10

Jeu de Données	meanRank	Top3	Next3
AMMICO	8.8	0.38	0.30
Pisa	39.9	0.29	0.32
Roma	100.7	0.25	0.20
Firenze	222.4	0.12	0.09

L'étude de la répartition des POI dans un espace latent 2D proposée en figure 2 nous permet de déduire deux choses : premièrement la configuration de l'agencement des POI dans AMMICO apparait clairement, ce qui explique *a posteriori* les bons résultats de notre modèle sur ce jeu de données. De plus, cette étude permet de corroborer l'hypothèse faite précédemment selon laquelle le graphe géographique des villes limité aux chemins usités est moins structuré que celui d'un musée.

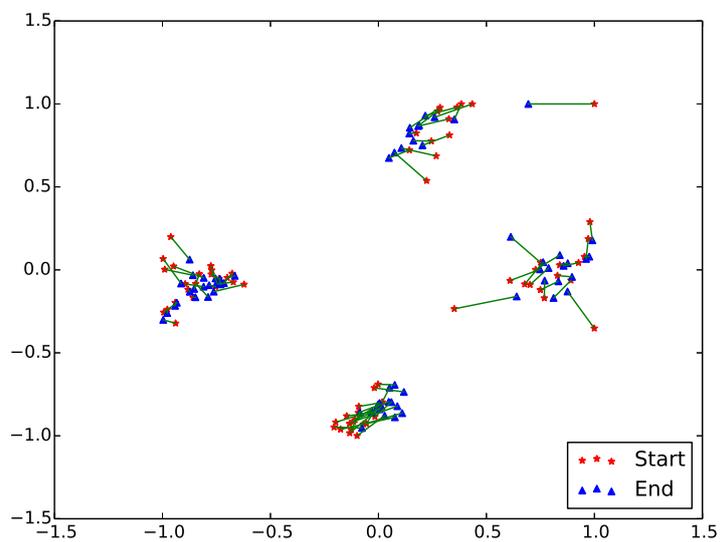


Figure 1. Représentation des POI du jeu de données AMMICO dans un espace latent 2D

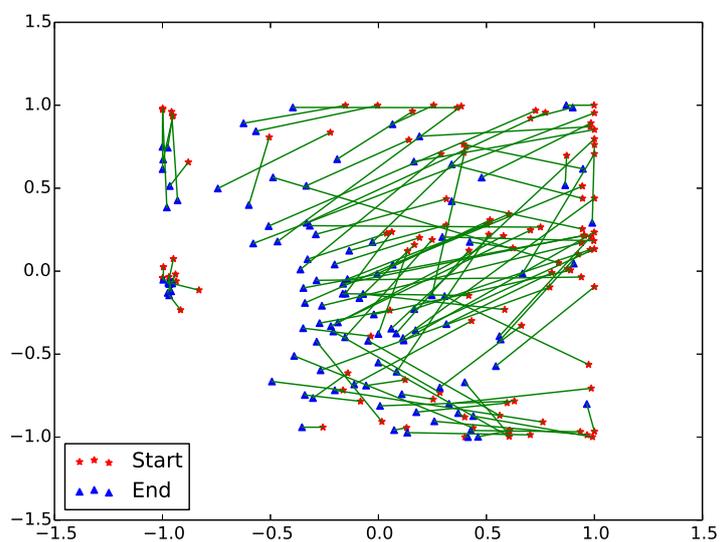


Figure 2. Représentation des POI du jeu de données Pisa dans un espace latent 2D

6. Conclusion et Perspectives

Dans cet article, nous avons cherché à appliquer les algorithmes de recommandation sur des graphes du monde réel. La tâche correspond à un besoin concret dans le domaine touristique, que ce soit dans les villes ou les musées, pourtant, elle reste peu explorée dans la littérature scientifique. Les nouveaux dispositifs dans les musées (type smartphone), et les nouveaux media sociaux permettent d'obtenir des données pour apprendre et tester de nouveaux algorithmes adaptés à ces besoins.

Nous avons développé des modèles de référence puis nous avons introduit une formulation latente qui permet d'obtenir de très bons résultats sur les données AMMICO. Ce rapprochement avec le domaine de la diffusion latente dans les graphes nous a permis de proposer une méthode efficace et légère, compatible avec les systèmes embarqués et les grandes masses de données.

Les perspectives sont doubles: nous voulons intégrer une composante réursive dans notre modèle et ainsi mieux prendre en compte le passé de l'utilisateur dans la décision. L'enjeu suivant consiste à prendre en compte les notes (ratings) attribuées à chaque POI en ajoutant des dimensions latentes spécifiques.

Bibliographie

- AMMICO. (2012). *Projet fui-13: Assistant de visites de musées mobile intelligent et collaboratif*. (<http://ammico.fr>)
- Baraglia R., Muntean C. I., Nardini F. M., Silvestri F. (2013). Learnext: learning to predict tourists movements. In *Proceedings of the 22nd acm international conference on conference on information & knowledge management*, p. 751–756.
- Bass F. M. (1969, January). A New Product Growth for Model Consumer Durables. *Management Science*, vol. 15, n° 5, p. 215-227.
- Bohnert F. (2010). Personalising the museum experience. In *Proceedings of the 2010 workshop on pervasive user modeling and personalization (PUMP-10), held in conjunction with the 18th international conference on user modeling, adaptation, and personalization (UMAP-10)*, p. 33–36.
- Bourigault S., Lagnier C., Lamprier S., Denoyer L., Gallinari P. (2014). Learning social network embeddings for predicting information diffusion. In *Proceedings of the 7th acm international conference on web search and data mining*, p. 393–402. ACM.
- Burke R. (2007). Hybrid web recommender systems. In *The adaptive web*, p. 377–408. Springer.
- Chen S., Moore J. L., Turnbull D., Joachims T. (2012). Playlist prediction via metric embedding. In *Proceedings of the 18th acm sigkdd international conference on knowledge discovery and data mining*, p. 714–722. ACM.
- Goldenberg J., Libai B., Muller E. (2001). Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters*, vol. 12, n° 3, p. 211–223.

- Gomez Rodriguez M., Leskovec J., Krause A. (2010). Inferring networks of diffusion and influence. In *Proceedings of the 16th acm sigkdd international conference on knowledge discovery and data mining*, p. 1019–1028. ACM.
- Karimi R., Nanopoulos A., Schmidt-Thieme L. (2012). Rfid-enhanced museum for interactive experience. In *Multimedia for cultural heritage*, vol. 247, p. 192–205. Springer Berlin Heidelberg.
- Kempe D., Kleinberg J., Tardos E. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth acm sigkdd international conference on knowledge discovery and data mining*, p. 137–146. ACM.
- Koren Y., Bell R., Volinsky C. (2009, août). Matrix factorization techniques for recommender systems. *Computer*, vol. 42, n° 8, p. 30–37.
- Myers S. A., Leskovec J. (2010). On the convexity of latent social network inference. In *Nips*, p. 1741–1749.
- Picot Clemente R., Bothorel C. (2013). Un système de recommandation de lieux basé sur la mesure de Katz dans les réseaux sociaux géographiques. In *Marami 2013 : 4ième conférence sur les modèles et l'analyse des réseaux : Approches mathématiques et informatiques*.
- Saito K., Nakano R., Kimura M. (2008). Prediction of information diffusion probabilities for independent cascade model. In *Proceedings of the 12th international conference on knowledge-based intelligent information and engineering systems, part iii*, p. 67–75. Springer-Verlag.
- Schafer J. B., Konstan J., Riedl J. (1999). Recommender systems in e-commerce. In *Proceedings of the 1st acm conference on electronic commerce*, p. 158–166. ACM.
- Setten M. van, Pokraev S., Koolwaaij J. (2004). Context-aware recommendations in the mobile tourist application compass. *Adaptive Hypermedia and Adaptive Web-Based Systems*, p. 235–244.
- Van Canneyt S., Schockaert S., Van Laere O., Dhoedt B. (2011). Time-dependent recommendation of tourist attractions using flickr. In *23rd benelux conference on artificial intelligence (bnaic 2011)*.
- Ver Steeg G., Galstyan A. (2013). Information-theoretic measures of influence based on content dynamics. In *Proceedings of the sixth acm international conference on web search and data mining*, p. 3–12. ACM.
- Yang J., Leskovec J. (2010). Modeling information diffusion in implicit networks. In *Proceedings of the 2010 ieee international conference on data mining*, p. 599–608. IEEE Computer Society.