



**HAL**  
open science

## Une Approche de recommandation proactive dans un environnement mobile

Imen Akermi, Mohand Boughanem, Rim Faiz

► **To cite this version:**

Imen Akermi, Mohand Boughanem, Rim Faiz. Une Approche de recommandation proactive dans un environnement mobile. 33eme congrès INFormatique des Organisations et Systèmes d'Information et de Decision (INFORSID 2015), May 2015, Biarritz, France. pp.301-316. hal-01914043

**HAL Id: hal-01914043**

**<https://hal.science/hal-01914043>**

Submitted on 6 Nov 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.



## Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in:  
<http://oatao.univ-toulouse.fr/19144>

**To cite this version:** Akermi, Imen and Boughanem, Mohand and Faiz, Rim *Une Approche de recommandation proactive dans un environnement mobile*. (2015) In: 33eme congrès INFormatique des Organisations et Systemes d'Information et de Decision (INFORSID 2015), 26 May 2015 - 29 May 2015 (Biarritz, France).

Any correspondence concerning this service should be sent to the repository administrator: [tech-oatao@listes-diff.inp-toulouse.fr](mailto:tech-oatao@listes-diff.inp-toulouse.fr)

# Une approche de recommandation proactive dans un environnement mobile

Imen Akermi<sup>1,2</sup> — Mohand Boughanem<sup>1</sup> — Rim Faiz<sup>2</sup>

<sup>1</sup> Institut de Recherche en Informatique de Toulouse, SIG  
118, route de Narbonne, F-31062 Toulouse Cedex 9, France  
{imen.akermi,mohand.boughanem}@irit.fr

<sup>2</sup> Laboratoire de Recherche Opérationnelle, de Décision et de Contrôle de Processus.  
IHEC, Université de Carthage, 2016 Carthage Présidence, Tunisie  
rim.faiz@ihec.rnu.tn

---

*RÉSUMÉ.* Les systèmes de recommandation contextuelle visent à combiner un ensemble de technologies et de connaissances sur le contexte de l'utilisateur pour lui fournir une information pertinente au moment où il en a le plus besoin, c'est ce qu'on appelle la recommandation proactive. Dans cet article nous proposons une approche de recommandation contextuelle et proactive dans un environnement mobile qui apprend implicitement les préférences de l'utilisateur. Nous avons évalué notre approche dans le cadre de la tâche "Contextual Suggestion Track" de TREC 2014. Les résultats que nous avons obtenus sont prometteurs.

*ABSTRACT.* The Context Aware Recommender Systems aim to combine a set of technologies and knowledge about the user context in order to deliver the most appropriate information to the user need at just the right time. It is called Proactive Recommendation. In this paper, we propose an approach for context-aware proactive recommendation in a mobile environment that aims at providing users with relevant information tailored to their own needs and preferences. The user preferences are implicitly identified and learned. We evaluated our approach through the 2014 TREC Task "Contextual Suggestion Track". The results that we obtained are promising.

*MOTS-CLÉS :* modélisation du contexte, système de recommandation contextuelle, recommandation proactive.

*KEYWORDS:* context modeling, context-aware recommender systems, user modeling, proactive recommendation.

## 1. Introduction

L'accès aux informations pertinentes, adaptées aux besoins et au profil de l'utilisateur est un enjeu majeur dans le cadre actuel caractérisé par une prolifération massive des ressources d'information hétérogènes. Le développement d'appareils mobiles équipés de plusieurs fonctionnalités telles que la caméra, le WIFI, la géo-localisation et bien d'autres permettent aux systèmes mobiles actuels de recherche d'information d'être hautement contextualisés, pouvant ainsi fournir à l'utilisateur des informations pertinentes au bon moment quand il en a le plus besoin, sans attendre qu'il établisse une interaction avec son appareil. Un système proactif peut prendre la forme d'un guide touristique personnalisé qui se base sur la localisation et les préférences de l'utilisateur pour suggérer à ce dernier des endroits intéressants sans qu'il fournisse sa préférence ou une requête explicite. Cela permettra de réduire les efforts, le temps et l'interaction de l'utilisateur avec son appareil mobile et de présenter les informations pertinentes au bon moment et au bon endroit.

L'efficacité de ce type de système requiert une bonne modélisation des dimensions du contexte et plus particulièrement la modélisation du profil de l'utilisateur. En effet, comme mentionné par Mizzaro et Vassena (2011), plusieurs dimensions de contexte, tels que la localisation, le temps, les activités des utilisateurs, leurs besoins, la lumière, le bruit, le mouvement, etc., doivent être pris en compte pour bien capturer tous les facteurs pouvant influencer la recommandation. La véracité et la multiplicité de ces facteurs rend les modèles de contexte trop complexes à implémenter. D'autre part, l'utilisation de peu de dimensions mène à des modèles de contexte incapables de comprendre l'environnement de l'utilisateur (Mizzaro et Vassena, 2011).

Dans cet article, nous proposons une approche de recommandation contextuelle et proactive dans un environnement mobile qui intègre les technologies mobiles et les dimensions spatio-temporelles pour la modélisation d'un profil situationnel de l'utilisateur. En effet, on vise à identifier la bonne information à recommander en fonction de la situation dans laquelle se trouve l'utilisateur.

Cet article est organisé comme suit. Nous présentons dans la section 2 un état de l'art se focalisant sur la recommandation contextuelle. La section 3 présente l'approche proposée. Dans la section 4, nous présentons et discutons les résultats des expérimentations de notre approche sur la collection TREC Contextual Suggestion Track 2014. Nous concluons et énonçons quelques perspectives en section 5.

## 2. Etat de l'art

Les systèmes de recommandation proactifs (SRP) comme décrit par Melguizo *et al.* (2007), récupèrent de grandes quantités de documents, décident quelle information disponible est probablement en rapport avec les besoins des utilisateurs, et suggèrent ces informations sans une demande explicite de l'utilisateur.

Récemment, plusieurs approches ont été développées pour soutenir la recommandation proactive. On distingue les approches qui se basent sur l'historique de comportement de l'utilisateur pour déterminer ses intérêts. Le comportement de l'utilisateur englobe l'historique de navigation Web/clics (Mizzaro et Vassena, 2011); les comportements antérieurs de visite de lieux pour les systèmes basés sur la localisation (Li *et al.*, 2012; Pu *et al.*, 2012) et l'historique des articles lus pour les systèmes de recommandation d'articles d'actualité (Intema *et al.*, 2010; Arora et Shah, 2011; Athalye, 2013; Dumitrescu et Santini, 2012).

D'autres approches de recommandation traitent le contexte de l'utilisateur d'un point de vue centré sur l'activité. Les activités communes utilisées pour construire le profil de l'utilisateur dans ces systèmes pourraient prendre la forme de pages Web ou des documents ouverts (Budzik et Hammond, 2000; Finkelstein *et al.*, 2001; Rhodes et Maes, 2000; Prekop et Burnett, 2002; Dumais *et al.*, 2004; Karkali *et al.*, 2013); de conversation ou l'activité en cours, telles que les messages texte, les appels téléphoniques (Popescu-Belis *et al.*, 2011) et de l'activité de l'utilisateur à travers les réseaux sociaux tels que le contenu des tweets de l'utilisateur sur Twitter (Phelan *et al.*, 2011; De Francisci *et al.*, 2012; O'Banion *et al.*, 2012).

Toutes les approches mentionnées ci-dessus et beaucoup d'autres ont essayé de relever le défi d'inférer le contexte de l'utilisateur et ses préférences, au bon moment et sans son intervention. Cependant, la plupart de ces systèmes rencontrent certaines limites. En effet, diverses approches (Bedi et Agarwal, 2012; Liu *et al.*, 2010; Saranya et Sadhasivam, 2012) reposent sur l'historique de comportement passé ou actuel de l'utilisateur tel que son historique de navigation sur internet ou son historique de visite de lieux, pour déterminer ses intérêts. Cependant, on ne peut pas se référer seulement au comportement passé de l'utilisateur car il peut ne pas contenir suffisamment d'informations sur ses intérêts. Un utilisateur peut naviguer sur une page web par hasard ou visiter un endroit particulier sans pour autant porter un intérêt à cet endroit.

Egalement, de nombreuses approches de recommandation (Budzik et Hammond, 2000; Finkelstein *et al.*, 2001; Rhodes et Maes, 2000; Prekop et Burnett, 2002; Dumais *et al.*, 2004; Karkali *et al.*, 2013) traitent le contexte de l'utilisateur d'un point de vue centré sur l'activité qui peut prendre la forme d'une conversation en cours ou des documents ouverts. Néanmoins, on ne peut pas réduire le profil de l'utilisateur à ces activités. On peut ouvrir un document pour travailler dessus ou avoir une conversation sur un sujet sans pour autant vouloir recevoir des recommandations qui y sont liées. De plus, l'activité sociale d'un utilisateur fournie par les commentaires exprimés à travers ses comptes sociaux ne peut pas fournir suffisamment d'informations pour construire un profil équilibré, car elle pourrait être limitée, voire absente. Ce genre de profilage des utilisateurs contient beaucoup de bruit à gérer et pas assez d'informations pour établir un profil assez complet.

Il existe aussi d'autres limites communes entre les systèmes de recommandation contextuels qui doivent être abordées, telle que, la dépendance à un domaine. En

effet, de nombreux systèmes (Pu et al., 2012 ; Yeung and Yang, 2010 ; Quercia et al., 2010) dépendent d'un seul domaine (tourisme, cinéma, nouvelles ...) et ont des dimensions de contexte spécifiques à appliquer en fonction du domaine. Cependant, la plupart de ces systèmes reposent presque sur la même combinaison de dimensions qui comprend essentiellement les préférences de l'utilisateur, la localisation, et le temps avec une légère différence sur la façon d'aborder ces dimensions. Par conséquent, de nombreux services liés à des domaines différents peuvent profiter des mêmes informations de contexte sans pour autant encombrer le mobile de l'utilisateur. Notre approche traite ces limites et contribue aux travaux existants en impliquant et en combinant les caractéristiques suivantes:

- La non-dépendance à un domaine particulier: nous couvrons divers domaines dans le processus de recommandation.
- Un système proactif: nous proposons un système de recommandation contextuel et proactif qui peut aider l'utilisateur à faire face au problème de la surcharge d'information en recommandant une information pertinente qui correspond à ses intérêts personnels au moment où il en a le plus besoin sans attendre que l'utilisateur pose une requête.

Dans notre approche, on vise à identifier l'information adéquate à recommander selon la situation de l'utilisateur.

### **3. Une approche de recommandation contextuelle et proactive dans un environnement mobile**

L'idée principale que nous proposons est l'intégration des informations relatives à un utilisateur, sa tendance de navigation et les technologies mobiles (GPS, accéléromètre, etc.) afin de recommander de manière proactive des informations pertinentes à l'utilisateur sans avoir à attendre que ce dernier initie une interaction avec son appareil. Nous décrivons dans ce qui suit le modèle de contexte proposé ainsi que le processus de recommandation.

#### **3.1. La modélisation du contexte**

Nous considérons le contexte comme une représentation multidimensionnelle *Contexte* = (*profil, localisation, temps, activité*), composée de quatre dimensions :

- Le profil de l'utilisateur est défini par les préférences de l'utilisateur et ses données démographiques
- Localisation: La position de l'utilisateur définie par des coordonnées GPS
- Temps: défini par des étiquettes numériques ou temporelles (matin, soir, ...)
- L'activité de l'utilisateur: l'activité de l'utilisateur en cours

Les instanciations de ces dimensions à un moment donné  $t$  forment une situation  $S$  où la recommandation sera jugée comme nécessaire ou pas. Nous extrayons implicitement les informations relatives à chaque dimension sans l'intervention de l'utilisateur. Nous définissons dans les sections suivantes les différents éléments du processus de modélisation de contexte.

### 3.1.1. *Le profil de l'utilisateur*

Nous utilisons un modèle de profil classique (Daoud et al. 2009) défini par deux attributs, les données démographiques; des informations relatives à l'utilisateur telles que le nom, l'âge, etc. ; et les préférences de l'utilisateur par rapport à des centres d'intérêt spécifiques ( $C$ ). Un centre d'intérêt est représenté par une catégorie et son poids, désignant l'intérêt que porte l'utilisateur à la catégorie en question. Les catégories sont prédéfinies en utilisant « The Open Directory Project Dmoz<sup>1</sup> ».

$$profil = \{C_i, w_i\} ; i=1..n$$

Une catégorie est représentée par des termes pondérés extraits à partir de son historique de navigation et les pages qu'il suit sur les réseaux sociaux :

$$C_i = \{t_{j,i}^{(i)}, w_{j,i}^{(i)}\} ; j=1..m$$

### 3.1.2. *Le temps*

Selon le dictionnaire Merriam-Webster<sup>2</sup>, le temps est défini comme "*un continuum non-spatial qui est mesuré en termes d'événements qui se succèdent du passé par le biais du présent au futur*". Le temps peut aussi être exprimé comme la période mesurée ou mesurable au cours de laquelle une action, un processus ou une situation existe ou continue (Whitrow 1988). Selon la deuxième définition, plusieurs unités de temps sont définies, par exemple, heures, jours, mois et années.

La flexibilité de la conception du temps et de sa mesure implique différentes représentations. En effet, le temps peut être représenté comme une variable continue dont les valeurs déterminent les moments précis au cours desquels des items sont évalués par un utilisateur donné. Exemple : *un utilisateur A a donné un score à un objet O à un temps  $t = 1$  juin 2010 à 18:05:00*

Une autre manière de modéliser le temps est d'identifier des valeurs catégorielles, pour des périodes de temps spécifiques. Par exemple, dans le domaine du tourisme, la variable "saison" peut être exprimée ainsi: *saison = {saison chaude, saison froide}*.

Le temps peut également être modélisé de façon hiérarchique qui permet de définir le degré de granularité de la dimension temporelle.

Exemple:

$$Jour\_semaine = \{lundi, mardi, \dots, Dimanche\} ; temps = \{matin, midi, \dots, nuit\}$$

<sup>1</sup> <http://www.dmoz.org/docs/en/about.html>

<sup>2</sup> <http://www.merriam-webster.com/dictionary/time>

Nous adoptons le modèle hiérarchique. Dans notre approche, nous divisons une journée en intervalles de temps (time slots) d'une certaine longueur. C'est selon ces intervalles que nous choisissons entre autre le type d'information à recommander.

On représente le temps en 2 niveaux :

1. Temps de la journée ( $D_t$ ): la journée est divisée en cinq périodes qui sont le matin, le midi, l'après-midi, le soir et la nuit. Ces périodes sont encadrées selon des intervalles de temps répartis sur 24 heures.

$$D_t = \{ \text{morning } [07:00, 12:00], \text{ midday } [12:00, 14:00], \text{ afternoon } [14:00, 18:00], \text{ evening } [18:00, 22:00], \text{ night } [22:00, 00:00] \cup [00:00, 07:00] \}$$

2. Jour de la semaine ( $D_w$ ): on distingue deux classes qui sont les jours de travail (du lundi au vendredi) et le reste des jours (weekend, vacances et jours fériés)

### 3.1.3. La localisation

Le système de positionnement global (GPS) intégré ou installé dans le dispositif mobile permet de définir l'emplacement de l'utilisateur. Cet emplacement est défini en termes de latitude et de longitude. Ces coordonnées GPS ne sont pas les seules caractéristiques que nous pouvons considérer lors de la définition d'un emplacement. En effet, comme indiqué par Dobson (2005), il existe différentes façons de caractériser l'emplacement de l'utilisateur: *Position absolue* ; *Position Relative* (à côté de, ...) ; *Un nom de lieu* ; *Une classe nommée* qui représente le type du lieu, par exemple, musée, école, etc.

Dans notre approche, nous considérons la position absolue et la représentation de classe nommée pour caractériser la dimension spatiale de l'utilisateur. Nous définissons deux niveaux pour cette dimension utilisés dans le processus d'extraction de l'information à recommander:

- L'emplacement actuel: il s'agit de l'emplacement actuel de l'utilisateur à un moment donné.
- Les endroits liés à l'utilisateur: Ce sont les endroits liés à la vie quotidienne de l'utilisateur (travail, domicile, ...) et qui pourraient être déterminés en étudiant les endroits dans lesquels se trouve l'utilisateur pendant un certain laps de temps.

Le type de l'emplacement actuel peut être déterminé à l'aide de plusieurs outils tels que geonames<sup>3</sup> ou le réseau social Foursquare qui attribuent une catégorie à une combinaison de coordonnées GPS données. Nous utilisons Foursquare pour déduire les informations de localisation: La catégorie du lieu et son nom.

### 3.1.4. L'activité de l'utilisateur

L'activité de l'utilisateur actuelle se résume en ce que l'utilisateur fait à un moment donné et peut être déterminée à travers son agenda ou l'application ouverte sur son téléphone portable. Cette dimension est importante dans notre approche et est utilisée pour évaluer si nous pouvons initier le processus de recommandation ou

---

<sup>3</sup> <http://www.geonames.org/>



pas. En effet, nous ne pouvons pas recommander des informations, d'une manière imposante, alors que l'utilisateur est toujours à une réunion par exemple.

### 3.2. *Le processus de recommandation*

#### 3.2.1. *Inférence de la catégorie d'information à recommander*

Nous considérons que le processus de recommandation implique un jeu de situations qui reflètent une catégorie d'intérêt spécifique caractérisée par les instanciations des dimensions spatio-temporelles et qui sont organisées dans une base de connaissances. Une situation est alors représentée par deux dimensions spécifiques: le temps, plus précisément le jour de la semaine ( $D_w$ ), et le temps de la journée ( $D_t$ ); et la catégorie du lieu ( $D_l$ ). Une situation est formellement représentée comme  $S = (D_l, D_w, D_t)$ .

En effet nous estimons que le besoin de l'utilisateur en information change selon le moment de la journée. Une étude<sup>4</sup> réalisée en 2013 révèle que 58% de personnes lisent avant toute chose leurs emails et les actualités le matin. Au déjeuner, les gens préfèrent naturellement trouver le restaurant le plus proche. Par exemple, la situation « déjeuner » est décrite par :

- $D_l$  : n'importe quel lieu
- $D_w$  : n'importe quel jour
- $D_t$  :  $t \in [12,14]$

La catégorie de l'information à recommander est inférée de la situation actuelle (Exemple : « Restaurant » pour la situation « déjeuner »).

#### 3.2.2. *Extraction de l'information à recommander*

Afin de récupérer les informations à recommander, une requête  $q$  est formulée comme suit :  $q = (\text{latitude}, \text{longitude}, \text{catégorie\_d'intérêt})$

La requête est ensuite envoyée à un service basé sur la géo-localisation et choisi en fonction de la catégorie d'intérêt. Par exemple si la situation implique la recommandation d'actualités, on utilise le réseau social Feedly<sup>5</sup> pour en extraire les actualités. Si la situation implique la recommandation d'un film, on interroge le réseau social Flixster<sup>6</sup>. L'intérêt de l'utilisation des réseaux sociaux pour l'extraction de l'information à recommander réside dans le fait que le filtrage collaboratif y est implicite ce qui assure non seulement la recommandation d'une information pertinente par rapport aux préférences de l'utilisateur mais aussi une information que d'autres utilisateurs recommandent.

Le résultat de la requête envoyée est exprimé par un ensemble d'items  $I$  :

$$I = \{i_1, \dots, i_n\}.$$

---

<sup>4</sup> <http://blog.marketo.com/2013/07/email-wanted-dead-or-alive-infographic.html>

<sup>5</sup> <http://feedly.com/index.html#discover>

<sup>6</sup> <https://fr.flixster.com/>

Si on prend l'exemple de la recommandation d'actualités, l'ensemble  $I$  sera représenté par un ensemble d'articles d'actualités :

$$I = \{A_1 (\text{titre}, \text{contenu}), \dots, A_k (\text{titre}, \text{contenu})\} ; k=1..p$$

Ensuite, nous filtrons à partir de  $I$ , les items convenant le mieux aux préférences de l'utilisateur en calculant un score d'utilité de l'item. Ce score combine un score de pertinence thématique de l'item vis-à-vis des préférences thématiques de l'utilisateur et un score géo-pertinence, évaluant l'accessibilité de l'item vis-à-vis de la position actuelle de l'utilisateur. Plus précisément, concernant la pertinence thématique, on considère qu'un item est représenté par un vecteur de mots pondérés (mots existants à partir des éléments textuels décrivant l'item).

La pertinence du contenu est calculée par le score de similarité cosin:

$$\text{Topic}_{\text{rel}}(VC_i, It) = \frac{\sum_{j=1}^n VC_i^{(j)} \times It_j}{\sqrt{\sum_{j=1}^n (VC_i^{(j)})^2} \times \sqrt{\sum_{j=1}^n (It_j)^2}} \quad (2)$$

Où:

$VC_i$ : Le vecteur termes préférence relatif à la catégorie  $C_i$

$It$ : Le vecteur termes de l'item  $i$

La géo-pertinence n'est utilisée que dans le cas où l'utilisateur doit se déplacer pour se rendre à l'item proposé. Elle est exprimée par un score qui mesure l'accessibilité à l'emplacement de l'item suggéré. L'accessibilité est mesurée comme la distance entre deux coordonnées GPS correspondant à l'emplacement actuel de l'utilisateur et l'emplacement de l'item proposé: (Point<sub>1</sub>(lat1,long1) et Point<sub>2</sub>(lat2,long2)) :

$$\text{accessibility} = R \times c \quad (3)$$

Où :

$R$ : The earth radius=6,371Km

$c = 2 \times \text{atan2}(\sqrt{a}, \sqrt{1-a})$

$a = \sin^2((\text{lat2} - \text{lat1})/2) + \cos(\text{lat1}) \times \cos(\text{lat2}) \times \sin^2((\text{long2} - \text{long1})/2)$

La pertinence globale d'un élément à recommander est calculée comme suit :

$$\text{Rel} = \alpha \times \text{Topic}_{\text{rel}}(VC_i, It) + (1 - \alpha) \times \text{accessibility} \quad (4)$$

Les items sont classés selon leurs scores de pertinence globale.

#### 4. Expérimentations

Nous avons évalué notre approche dans le cadre de la tâche « *Contextual Suggestion Track*<sup>7</sup> » de TREC 2014. Nous présentons, dans cette section, une description générale de cette tâche, ensuite nous exposerons les résultats que nous avons obtenus.

<sup>7</sup> <https://sites.google.com/site/trecontext/trec-2014>

#### 4.1. Description de la tâche « Contextual Suggestion Track - CST » de TREC 2014

Cette tâche étudie les techniques de recherche pour les besoins d'information complexes qui dépendent fortement du contexte et des intérêts des utilisateurs.

Le but de la tâche est de recommander à des utilisateurs se trouvant dans un contexte donné des lieux qui pourraient les intéresser. Un profil est associé à chaque utilisateur et indique la préférence de l'utilisateur par rapport à un ensemble de suggestions de lieux. Chaque suggestion comprend un titre, une description et une URL associée. Par exemple, une suggestion pourrait être une recommandation d'avoir une bière à la Dogfish Head Alehouse. Un profil est défini par un ensemble de préférences négatives ou positives par rapport à l'ensemble des suggestions de lieux. Chaque contexte correspond à un lieu géographique particulier (une ville). Par exemple, le contexte pourrait être Gaithersburg, Maryland. Le but de la tâche est de retourner pour chaque pair profil/contexte, une liste de lieux intéressants allant jusqu'à 50 lieux. Chaque suggestion doit être adaptée au profil (en fonction des préférences de l'utilisateur) et au contexte (selon l'emplacement).

##### 4.1.1. Construction des profils

Les profils sont construits à partir de la liste de suggestions jugées par l'utilisateur. Chaque suggestion est évaluée selon deux notes, une note pour le titre et la description du lieu et une note pour son site web. Les notes sont données sur une échelle de cinq points sur la base du degré d'intérêt que peut porter cet utilisateur pour le lieu en question:

4 (très intéressé), 3 (intéressé), 2 (Neutre), 1 (Pas intéressé), 0 (Fortement pas intéressé), -1 (Site web ne s'est pas chargé ou pas de score)

Pour construire le profil thématique de l'utilisateur, on identifie, pour chaque suggestion, sa catégorie dans l'API Google Places<sup>8</sup> en utilisant l'url qui est associée à la suggestion en question. A l'issue de cette étape, un profil sera alors modélisé comme un ensemble de catégories pondérées qui comportent les termes associés aux suggestions que l'utilisateur préfère :  $profil = \{C_i(w_i)\}; i=1..n$

Le calcul du poids assigné à chaque catégorie, pour chaque profil, prend en considération les scores que ce profil a attribués aux suggestions relevant de cette catégorie d'intérêt :

$$weight(category) = \frac{\sum_{s \in C} R_{td} + R_w}{N_p} \quad (5)$$

Où :

$\forall s \in C$ : pour chaque suggestion  $s$  appartenant à cette catégorie d'intérêt  $C$

$R_{td}$ : le score donné au titre et à la description du lieu d'attraction

$R_w$ : le score du site web du lieu d'attraction

$N_p$ : le nombre de suggestions appartenant à cette catégorie d'intérêt

---

<sup>8</sup> <https://developers.google.com/places/documentation/?hl=fr>

Les catégories sont représentées par des termes pondérés qu'on extrait à partir des descriptions des suggestions associées à chaque catégorie.

#### 4.1.2. Sélection des suggestions

La tâche TREC définit le contexte en fonction des coordonnées GPS fournies dans cette tâche :  $Contexte = \{latitude, longitude\}$ . Pour notre part, on instancie le contexte en exploitant directement les données fournies par la tâche. Pour chaque contexte, on extrait les lieux intéressants en soumettant une requête aux géo-services Google Places et Foursquare<sup>9</sup>, formulé comme suit :

$q = \{latitude, longitude, categorie\}$

Suite à cette requête, on obtient un ensemble de lieux. Un lieu, est modélisé par un objet ayant des attributs spécifiques et qui appartient à une catégorie d'intérêt donné :  $lieu = \{nom, url, description, accessibilité, categorie\}$

L'*accessibilité* : représente la distance qui sépare le lieu du contexte représentés par leurs coordonnées GPS ( $P_1(lat1, long1)$  et  $P_2(lat2, long2)$ ) et calculée par la formule (3) . Pour le total des 50 contextes donnés, on a pu extraire à peu près 15 000 lieux d'attraction ce qui fait en moyenne 300 lieux pour chaque contexte qui sont répertoriés sous des catégories.

#### 4.1.3. Appariement Profils/Suggestions

Le processus de sélection des lieux intéressants pour chaque profil-contexte est résumé comme suit :

Pour chaque  $p_i \in P$

    Pour chaque  $c_j \in C$

        - Calculer pour chaque suggestion  $S \in C_{j\_suggestions}$  la pertinence globale (selon la formule (4) )

        - Normaliser les scores de pertinence globale des suggestions entre 0 et 1

        - Extraire les suggestions dont le score de pertinence globale est  $\geq 0,5$  **Et** qui appartiennent aux catégories d'intérêt appréciées par le profil :

$category_{weight}^{normalisé} > 0.5$

Où:

$P$ : l'ensemble des profils

$C$ : l'ensemble de contextes

$C_{j\_suggestions}$  : L'ensemble des suggestions relatives au contexte  $C_j$

---

<sup>9</sup> <https://fr.foursquare.com/>

#### 4.2. Résultats:

L'évaluation des résultats que nous avons obtenus comporte deux volets : l'évaluation de la pertinence géographique et la pertinence profil. Pour mesurer la géo-pertinence des lieux qu'on a suggérés, on extrait un échantillon  $V_g$  qui consiste en l'intersection de notre collection de lieux avec les lieux qui ont été évalués géographiquement dans la tâche TREC pour chaque contexte. Nous avons obtenus  $|V_g| = 4802$  lieux en commun, tous contextes confondus. Parmi ces lieux, 4644 ont été évalués géographiquement pertinents. Ce qui implique une précision géographique totale égale à 0,97 et est calculée comme suit :

$$geo\_relevance = \frac{Nb\_Geo\_Relevant\_Venues}{|V_g|} \quad (6)$$

Le tableau 1 présente en détail les résultats.

**Tableau 1. La géo-pertinence** (Nb\_Total est le nombre total des lieux suggérés pour chaque contexte ; Nb\_GP est le nombre des lieux géographiquement pertinents ; Nb\_NGP est le nombre des lieux identifiés comme géographiquement non pertinents )

Context ID	Nb_Total	Nb_GP	Nb_NGP	Context ID	Nb_Total	Nb_GP	Nb_NGP
101	120	118	2	127	86	80	6
102	74	71	3	128	118	117	1
103	116	113	3	129	106	106	0
104	103	100	3	130	35	29	6
105	58	56	2	131	132	128	4
106	116	114	2	132	115	111	4
107	137	132	5	133	49	46	3
108	80	77	3	134	213	210	3
109	113	109	4	135	82	75	7
110	80	76	4	136	60	56	4
111	63	61	2	137	48	47	1
112	24	22	2	138	155	152	3
113	171	165	6	139	35	33	2
114	68	66	2	140	94	91	3
115	103	92	11	141	69	67	2
116	24	23	1	142	114	111	3
118	79	73	6	143	52	43	9
119	111	107	4	144	203	202	1
120	129	127	2	146	99	96	3
121	111	108	3	147	105	103	2
122	114	111	3	148	26	26	0
123	180	180	0	149	204	201	3
124	198	195	3	150	69	67	2
125	80	73	7				

Pour mesurer la pertinence profil, il y avait deux alternatives à considérer. Une première alternative consiste à considérer pour chaque contexte et pour chaque run<sup>10</sup>, l'intersection de notre collection de lieux avec les lieux proposés par le run en question, cependant cette intersection donnait quasiment l'ensemble vide vu que nos méthodes d'extractions de lieux sont différentes. Nous avons donc opté pour une solution intermédiaire qui consiste à considérer l'intersection de notre collection de lieux avec l'union des lieux que chaque run a proposés, tous profils confondus. La cardinalité de cette intersection est  $|V_p| = 889$  lieux. Ensuite nous avons mesuré pour chaque lieu le niveau d'intérêt qu'il a sollicité de la part des profils en se basant sur le nombre de profils qui ont évalué ce lieu comme pertinent par rapport au nombre d'évaluations totale du lieu en question. Par rapport à la collection obtenue, nous avons obtenu une précision moyenne de 0,56 calculée comme suit:

$$profile\_relevance = \frac{\sum_{l=1}^{|V_p|} (nb\_p_l / nb\_p\_tot_l)}{|V_p|} \quad (6)$$

Où :

$nb\_p_l$  : le nombre de profils qui ont évalué le lieu ( $l$ ) comme pertinent

$nb\_p\_tot_l$  : le nombre total de profils qui ont évalué le lieu ( $l$ )

Pour se situer approximativement par rapport aux autres participants, nous avons appliqué cette démarche d'évaluations sur chaque run. On a mesuré pour chaque run et pour chaque lieu que ce run propose, le degré d'intérêt sollicité par les profils en appliquant la formule (6). Le tableau 2 illustre les résultats obtenus pour chaque run.

**Tableau 2. Résultats des autres participants**

id	run	Nb to	Nb in	Préc	id	run	Nb to	Nb in	Préc
1	BJUTa	1495	352	<b>0,24</b>	20	simpleScore	1496	244	<b>0,16</b>
2	BJUTb	1495	327	<b>0,22</b>	21	simpleScoreImp	1496	254	<b>0,17</b>
3	BUPT 01	671	81	<b>0,12</b>	22	tueNet	1497	129	<b>0,09</b>
4	BUPT 02	704	78	<b>0,11</b>	23	tueRforest	1497	138	<b>0,09</b>
5	cat	1496	318	<b>0,21</b>	24	UDInfoCS2014	1495	201	<b>0,13</b>
6	choqrun	1465	196	<b>0,13</b>	25	UDInfoCS2014	1495	360	<b>0,24</b>
9	dixlticmu	1496	367	<b>0,25</b>	26	uogTrBunSumF	1495	272	<b>0,18</b>
10	gw1	1466	58	<b>0,04</b>	27	uogTrCsLtrF	1496	199	<b>0,13</b>
11	lda	1496	163	<b>0,11</b>	28	waterlooA	1497	260	<b>0,17</b>
14	RAMARUN	1497	260	<b>0,17</b>	29	waterlooB	1497	286	<b>0,19</b>
15	RUN1	1494	229	<b>0,15</b>	30	webis 1	1477	269	<b>0,18</b>
16	run_DwD	1496	181	<b>0,12</b>	31	webis 2	1474	230	<b>0,16</b>
17	run_FDwD	1496	262	<b>0,18</b>					

<sup>10</sup> Un run est une application proposé par une équipe pour traiter la tâche « Contextual Suggestion » dans TREC 2014

D'après le tableau 2, on note que la précision de la pertinence profil de chaque run par rapport aux lieux qu'ils ont proposés ne dépasse pas **0,25**. Cela est expliqué par le fait, que pour chaque run, il y a un grand écart entre le nombre de profils qui ont jugé un lieu (appartenant à un seul contexte) comme pertinent et le nombre de jugements totaux. Généralement, les résultats que nous avons obtenus en appliquant notre approche sont prometteurs et montrent que l'utilisation de la classification par catégories des préférences d'un utilisateur implique une meilleure précision thématique par rapport à ses intérêts. Ces résultats montrent aussi que les choix des paramètres tels que le rayon que nous avons fixé pour l'extraction des lieux par rapport aux contextes sont efficaces. Cependant, cette évaluation n'a mis en examen qu'une partie de notre approche. En effet, la notion du temps n'est pas intégrée alors que cette dimension est considérée dans notre travail.

## **5. Conclusion**

Le but fondamental des systèmes de recommandations contextuels consiste à combiner le contexte et l'environnement de l'utilisateur dans une même infrastructure pour mieux caractériser les informations dont l'utilisateur a besoin afin d'améliorer le processus de recommandation. Nous avons proposé une approche de recommandation contextuelle et proactive dans un environnement mobile qui permet la recommandation d'éléments pertinents qui répondent aux préférences de l'utilisateur, juste au bon moment sans attendre que l'utilisateur initie une requête. Plus précisément, notre contribution est divisée en deux principales parties: La modélisation d'un profil situationnel de l'utilisateur et la définition d'un cadre d'agrégation pour la combinaison des dimensions contextuelles et sociales. Nous planifions actuellement d'appliquer notre approche au sein de l'entreprise Tunav<sup>11</sup> qui développe des produits et des systèmes de navigation et de tracking par GPS, des applications cartographiques et des géo-portails qui offrent à l'utilisateur l'accès à une masse importante d'information géographique. Ces systèmes se diversifient dans différents domaines d'activité dont le transport routier (public, marchandise, taxi, etc.), les travaux, les services, la distribution, la santé, le tourisme. Nous avons comme objectif d'apporter le « just-in-time information » et la dimension contextuelle aux systèmes qui en ont besoin afin de fournir aux utilisateurs des produits personnalisés adaptés à leurs propres besoins et préférences.

### *Remerciements*

*Ce projet est effectué dans le cadre d'une thèse de doctorat MOBIDOC du programme PASRI<sup>12</sup> financé par l'UE et administré par l'ANPR<sup>13</sup>.*

---

<sup>11</sup> <http://www.tunav.com/>

<sup>12</sup> <http://www.pasri.tn/> - Le Projet d'Appui au Système de Recherche et de l'Innovation (PASRI), un projet financé par l'Union européenne.

<sup>13</sup> <http://www.anpr.tn/> - Agence Nationale de la Promotion de la Recherche Scientifique en Tunisie.

## Références

- Arora A., Shah P. (2011). Personalized news prediction and recommendation. Unpublished doctoral dissertation, Stanford University.
- Athalye S. (2013). Recommendation system for news reader. Unpublished doctoral dissertation, San Jose State University.
- Bedi P., Agarwal S. (2012, Dec). A situation-aware proactive recommender system. In Hybrid intelligent systems (his), 2012 12th international conference on, p. 85-89.
- Budzik J., Hammond K. J. (2000). User interactions with everyday applications as context for just-in-time information access. In Proceedings of the 5th international conference on intelligent user interfaces, pp. 44–51. New York, NY, USA, ACM.
- Daoud M., Tamine L., Boughanem M. (2011). A personalized search using a semantic distance measure in a graph-based ranking model. Information Science, SAGE Publications, Vol. 37, No. 6, pp. 614-636.
- De Francisci Morales G., Gionis A., Lucchese C. (2012). From chatter to headlines: Harnessing the real-time web for personalized news recommendation. In Proceedings of the fifth acm international conference on web search and data mining, pp. 153–162. New York, NY, USA, ACM.
- DOBSON, S. (2005). Leveraging the subtleties of location. In Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies. ACM, pp. 189-193.
- Dumais S., Cutrell E., Sarin R., Horvitz E. (2004). Implicit queries (iq) for contextualized search. In Proceedings of the 27th annual international acm sigir conference on research and development in information retrieval, pp. 594–594. New York, NY, USA, ACM.
- Dumitrescu D. A., Santini S. (2012). Improving novelty in streaming recommendation using a context model. In Cars 2012 : Acm recsys workshop on context-aware recommender systems.
- Finkelstein L., Evgenly G., Yossi M., Ehud R., Zach S., Gadi W. et al. (2001). Placing search in context: the concept revisited. In Proceedings of the Tenth International World Wide Web Conference.
- Gershman A., Wolfe T., Fink E., Carbonell J. (2011). News personalization using support vector machines. Proceedings of the Workshop on Enriching Information Retrieval.
- IJntema W., Goossen F., Frasinca F., Hogenboom F. (2010). Ontology-based news recommendation. In Proceedings of the 2010 edbt/icdt workshops, pp. 16:1–16:6. New York, NY, USA, ACM.
- Karkali M., Pontikis D., Vazirgiannis M. (2013). Match the news: A firefox extension for realtime news recommendation. In Proceedings of the 36th international acm sigir conference on research and development in information retrieval, pp. 1117–1118. New York, NY, USA, ACM.
- Lee H., Park S. J. (2007). Moners: A news recommender for the mobile web. Expert Systems with Applications, Vol. 32, No. 1, pp. 143 - 150.



- Li W., Eickhoff C., Vries A. P. de. (2012). Want a coffee?: Predicting users' trails. In Proceedings of the 35th international acm sigir conference on research and development in information retrieval, pp. 1171–1172. New York, NY, USA, ACM.
- Lin C., Xie R., Li L., Huang Z., Li T. (2012). Premise: Personalized news recommendation via implicit social experts. In Proceedings of the 21st acm international conference on information and knowledge management, pp. 1607–1611. New York, NY, USA, ACM.
- Liu J., Dolan P., Pedersen E. R. (2010). Personalized news recommendation based on click behavior. In Proceedings of the 15th international conference on intelligent user interfaces, pp. 31–40. New York, NY, USA, ACM.
- Melguizo M. C. P., Bogers T., Deshpande A., Boves L., Bosch A. van den. (2007). What a proactive recommendation system needs - relevance, non-intrusiveness, and a new longterm memory. In J. Cardoso, J. Cordeiro, J. Filipe (Eds.), *Iceis (5)*, p. 86-91.
- Mizzaro, S., Vassena, L. (2011). A social approach to context-aware retrieval. *World Wide Web*, Vol. 14, No. 4, pp. 377–405.
- Nguyen T., Riedl J. (2013). Predicting users preference from tag relevance. In S. Carberry, S. Weibelzahl, A. Micarelli, G. Semeraro (Eds.), *User modeling, adaptation, and personalization*, Vol. 7899, p. 274-280. Springer Berlin Heidelberg.
- O'Banion S., Birnbaum L., Hammond K. (2012). Social media-driven news personalization. In Proceedings of the 4th acm recsys workshop on recommender systems and the social web, pp. 45–52. New York, NY, USA, ACM.
- Phelan O., McCarthy K., Bennett M., Smyth B. (2011). On using the real-time web for news recommendation and discovery. In Proceedings of the 20th international conference companion on world wide web, pp. 103–104. New York, NY, USA, ACM.
- Popescu-Belis A., Yazdani M., Nanchen A., Garner P. N. (2011). A speech-based just-in-time retrieval system using semantic search. In Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies: Systems demonstrations, pp. 80–85. Stroudsburg, PA, USA, Association for Computational Linguistics.
- Prekop P., Burnett M. (2003, July). Activities, context and ubiquitous computing. *Comput. Commun.*, Vol. 26, No. 11, pp. 1168–1176.
- Pu Q., Lbath A., He D. (2012). Location based recommendation for mobile users using language model and skyline query. *International Journal of Information Technology & Computer Science (IJITCS)*, Vol. 4, No. 10, pp. 19–28.
- Quercia D., Ellis J., Capra L. (2010). Nurturing social networks using mobile phones. *IEEE Pervasive Computing*, Vol.9, No. 3.
- Rhodes B., Maes P. (2000). Just-in-time information retrieval agents. *IBM Systems Journal*, Vol. 39, No. 3.4, pp. 685-704.
- Saranya.k.g, Sadhasivam G. S. (2012, November). Article: A personalized online news recommendation system. *International Journal of Computer Applications*, Vol. 57, No. 18, pp. 6-14.
- Tchunte, D., Canut, M.F., Jessel, N., Peninou, A., El Haddadi, A. (2010). Visualizing the Evolution of Users' Profiles from Online Social Networks. In Proceedings of the

International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp.370-374.

Yeung K. F., Yang Y. (2010, Sept). A proactive personalized mobile news recommendation system. In Developments in e-systems engineering (dese), 2010, p. 207-212.

Whitrow, G. (1989). Time in History: Views of Time from Prehistory to the Present Day. Oxford paperbacks, Oxford University Press.