



HAL
open science

Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire

Anaëlle Badier, Mathieu Lefort, Marie Lefevre

► To cite this version:

Anaëlle Badier, Mathieu Lefort, Marie Lefevre. Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire. 10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain, Marie Lefevre, Christine Michel, Jun 2021, Fribourg, Suisse. pp.294-299. hal-03287742

HAL Id: hal-03287742

<https://hal.science/hal-03287742v1>

Submitted on 21 Jul 2021

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.

Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire

Anaëlle Badier, Mathieu Lefort et Marie Lefevre

Université Lyon 1, LIRIS, UMR 5205, F-69622, France
`prenom.nom@liris.cnrs.fr`

Résumé Nous présentons dans cet article les principes d'un système de recommandation de ressources pédagogiques dans le cadre d'une application mobile à usage parascolaire. Le système proposé permet de fournir des recommandations à travers différents niveaux scolaires et d'orienter les apprenants vers des ressources de remédiation, de continuité du programme ou d'approfondissement. Notre système s'appuie sur l'organisation des ressources pédagogiques disponibles dans l'application en un "arbre de connaissances", et sur l'évaluation du niveau réel d'un apprenant par la méthode de la théorie de réponse aux items (*IRT*). Les premières validations des experts pédagogiques semblent confirmer notre approche et nous encouragent à poursuivre par la validation du système en situation réelle ouvrant ainsi de nombreuses perspectives de recherche.

Mots-clé: Système de recommandation · Apprentissage adaptatif · Plateforme pédagogique mobile · *IRT*

Abstract. In this article, we present the principles of a learning resources recommendation system in a mobile application for extracurricular use. The system provides recommendations across different grade levels to direct learners toward resources for remediation, school curriculum continuity or in-depth study. Our system is based on the organisation of the pedagogical resources available in the application structured with a "knowledge tree", and on the assessment of the learner's actual level using the item response theory (*IRT*) method. The first feedback of pedagogical experts seem to validate the approach that will be next tested in real-life situations and open up numerous research perspectives.

Keywords: Recommendation system · Adaptive Learning · Mobile educational platform · *IRT*

1 Introduction

Les avancées technologiques de ces dernières années ont permis de proposer de nombreux systèmes d'apprentissage en ligne (MOOC, applications de e-learning,

etc.) [2]. La démocratisation de ces outils, jusqu'alors restreints au contexte du laboratoire, a augmenté la quantité de données utilisateurs en contexte réel disponibles et permet d'utiliser les techniques d'intelligence artificielle pour améliorer les systèmes d'apprentissage adaptatif [7].

Nous travaillons avec une application mobile qui propose du contenu éducatif pour tous les niveaux, du collège au supérieur et toutes les matières y compris les enseignements professionnels ou de spécialité. L'application est construite pour être un outil parascolaire volontaire, de révision ou de découverte, plus que d'apprentissage. L'enjeu est de pouvoir naviguer facilement entre les différents niveaux scolaires afin de compléter l'enseignement vu en classe.

Notre objectif est donc de proposer un système de recommandation qui identifie automatiquement les besoins de l'apprenant et propose le contenu de l'application le plus adapté. Nous nous focaliserons dans cet article sur les connaissances exploitées pour notre système de recommandation. Le modèle apprenant et l'interface d'interaction feront l'objet des travaux à venir.

Dans un premier temps, nous exposerons les travaux de recherche existants sur lesquels s'appuie notre projet pour contextualiser et justifier nos travaux (section 2) avant de présenter les principes sur lesquels s'appuie notre système (section 3). Nous détaillerons ensuite un protocole de validation du système proposé (section 4) pour conclure sur les perspectives de recherche (section 5).

2 État de l'art

Les systèmes de recommandation existants appliqués aux EIAH combinent plusieurs approches et peuvent être basés sur les profils des utilisateurs, sur la structure du contenu à recommander ou sur une approche axée compétences. Yago et al. proposent une revue des différents systèmes de recommandation [11].

Les systèmes de filtrage par contenu (*content-based filtering*) proposent de recommander un contenu proche du contenu déjà étudié donc plus susceptible de satisfaire les attentes de l'utilisateur. Les contenus, pour la plupart issus du web, sont structurés en ontologies [9] ou réseaux sémantiques. Les recommandations sont ensuite effectuées par un système de correspondance entre une requête (correspondant au contenu récemment étudié) et la base de ressources. L'association peut par exemple être réalisée en utilisant des techniques de web-sémantique qui s'appuient sur des métadonnées descriptives de la ressource [8]. L'utilisation de techniques de filtrage par contenu est particulièrement efficace lorsque l'on dispose d'une grande diversité de ressources qui traitent du même sujet.

Les méthodes de recommandation axées utilisateurs peuvent proposer des systèmes de filtrage collaboratif (*collaborative filtering*) en identifiant des similitudes entre utilisateurs pour proposer à un apprenant un contenu étudié par un autre apprenant au profil similaire, *via* différentes méthodes de construction de profils utilisateurs [7]. Les activités utilisateurs peuvent également être analysées par des algorithmes de fouille de données pour définir des règles d'associations entre un certain profil et une ressource à étudier [4]. Ces algorithmes rencontrent toutefois des limites computationnelles lorsque les jeux de données sont très im-

portants. De plus, dans le cas de notre application volontaire et parascolaire, il serait difficile de caractériser des profils utilisateurs sans avoir de suivi régulier dans le temps, ce qui limite l'efficacité de ces techniques dans notre contexte.

Les méthodes de recommandation par compétences sont liées aux mécanismes cognitifs de l'apprentissage. Ils introduisent une chronologie en proposant d'organiser les ressources à étudier en représentant leur accessibilité par compétences acquises et à acquérir. La structuration des ressources en compétences est utilisée comme paramètre des modèles de traçage de connaissances dans l'article de Vie et Kashima [10]. La prise en compte des précédentes tentatives de l'apprenant pour maîtriser un sujet permet une meilleure prédiction de réussite. Cette influence des cours déjà étudiés sur la réussite aux cours suivants est également mobilisée dans les travaux de Jiang et al [6], qui introduisent la notion de Zone de Développement Proximal (*ZPD*) développée par Lev Semyonovich Vygotsky. Clément et al. basent également leurs modèles sur la *ZPD* et propose de limiter les recommandations aux contenus de niveau légèrement supérieur au niveau réel de l'apprenant [3]. En effet, suggérer des ressources de niveau trop inférieur au niveau de l'apprenant ne lui permet pas d'augmenter ses connaissances et des ressources de niveau trop élevé risquent de provoquer une frustration.

La caractérisation du niveau de connaissance de l'apprenant est donc essentielle pour pouvoir utiliser une approche par acquisition de compétences. La méthode statistique de la théorie de réponse aux items (*Item Response Theory* - *IRT*) formalisée par Baker [1] permet d'évaluer la compétence d'un apprenant en s'appuyant sur les réponses des autres apprenants pour caractériser les items auxquels l'apprenant est confronté. L'application mobile sur laquelle nous travaillons étant utilisée par un million d'utilisateurs, nous disposons de nombreuses données qui permettent d'utiliser cette méthode statistique d'évaluation de compétences.

3 Combiner *Connaissances* et *IRT* pour recommander

L'approche que nous proposons est une méthode de type "Support de navigation adaptatif" [2]. Le système doit proposer à la fin de chaque ressource consultée la ressource qui serait la plus adaptée pour poursuivre l'apprentissage. Afin de proposer un contenu adapté au niveau de l'apprenant, nous proposons dans un premier temps une structuration des ressources en "arbre de connaissance" qui permet de les lier à travers différents niveaux scolaires (section 3.1), puis une évaluation du niveau de connaissance de l'apprenant en utilisant l'*IRT* (section 3.2) pour l'orienter vers une ressource de difficulté correspondant à son niveau réel, indépendamment de son niveau scolaire.

3.1 Arbre de connaissances

Organisation des ressources. Les ressources de l'application sont organisées par niveaux scolaires. Chaque niveau comporte des disciplines, structurées en sous-disciplines, elles-mêmes séparées en chapitres. Un chapitre comporte de 1 à 3 mini-cours et de 1 à 5 quizz de 10 questions en moyenne.

Annotation et structuration des ressources. Nous proposons dans un premier temps de faire la recommandation au niveau des chapitres. Afin de lier les chapitres entre eux, chacun est caractérisé par un ensemble de notions décrivant les connaissances mobilisées dans le chapitre. Ces notions font référence aux autres chapitres disponibles et aux relations pédagogiques qui existent entre ces chapitres. En l'absence de métadonnées descriptives du contenu de chaque chapitre, l'apposition des notions est effectuée par des professeurs, en suivant les descriptifs du programme de l'éducation nationale¹. Par exemple, le chapitre de mathématiques de Seconde traitant des *Fonctions de référence* est annoté en indiquant qu'il est nécessaire de maîtriser le chapitre *Fonctions affines* de niveau 3^e.

Le chemin pédagogique se construit par similarité entre les chapitres en fonction des connaissances mobilisées. Nous utilisons un algorithme de similarité basé sur les liens entre notions et sur la répartition des différentes notions au sein de nos ressources.

3.2 Évaluation du niveau de capacité de l'apprenant par IRT

Chaque quiz étant conçu par un professeur différent, les quiz n'ont pas tous le même niveau de difficulté. De plus, les quiz sont tous à choix multiples, ce qui rend possible une bonne réponse par sélection aléatoire. Nous disposons d'un grand volume de données (de 2000 à 40 000 réponses enregistrées pour les quiz de mathématiques de niveau collège-lycée), ce qui nous permet d'utiliser une méthode statistique pour évaluer le niveau réel de l'apprenant.

Nous utilisons la méthode de l'IRT pour avoir une valeur de ce niveau réel de compétence pour un item donné. Le modèle sélectionné est le modèle à 3 paramètres (IRT 3PL) [1], qui estime le niveau de compétence réelle θ d'un apprenant par la formule :

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-a(\theta - b)}}$$

(avec θ : compétence réelle de l'apprenant, a : paramètre de discrimination, b : paramètre de difficulté et c : paramètre de hasard).

Face à des profils d'apprenants très disparates, l'IRT présente l'avantage d'indépendance entre le niveau des apprenants ayant servi à estimer les paramètres d'un item et la capacité d'un apprenant. Les apprenants sont ensuite répartis en 3 groupes de niveau : G_{inf} pour les élèves présentant des difficultés sur le chapitre étudié, G_{med} pour le groupe de niveau moyen et G_{sup} pour les élèves maîtrisant bien le chapitre étudié. Les groupes sont définis selon les valeurs de θ : par les propriétés de l'IRT, la répartition des élèves n'influence pas la constitution des groupes. Cette répartition des apprenants permet de les orienter vers des ressources de remédiation (apprenants de niveau G_{inf}), de continuité du programme (niveau G_{med}), ou de découverte (apprenants de niveau G_{sup}), en fonction des notions et relations disponibles dans notre arbre de connaissances.

1. Exemple de l'Annexe 3 relative aux cours de 5e, 4e et 3e : <https://www.education.gouv.fr/bo/15/Special11/MENE1526483Aannexe3.htm>

4 Évaluation du prototype

Les propositions de recommandations de l'algorithme ont été soumises à validation auprès de 3 experts pour les mathématiques de Seconde générale. Nous avons choisi ce niveau scolaire pour correspondre d'une part à une population d'utilisateurs majoritaire et d'autre part pour tester l'algorithme dans un contexte de transition importante de niveau scolaire (passage collège-lycée), où les besoins en remédiation/approfondissement seront plus importants.

La validation a été présentée sous la forme d'un questionnaire en deux parties présentant des situations d'apprenants venant de terminer un chapitre donné et affecté à un groupe de niveau G_{inf} , G_{med} , G_{sup} d'après les résultats IRT. La première partie du questionnaire présente la liste des ressources disponibles dans l'application et demande aux experts le chapitre qu'ils auraient conseillé dans la situation présentée. La deuxième partie présente la recommandation de l'algorithme et demande aux experts d'évaluer cette recommandation.

Les retours d'experts sont plutôt favorables vis-à-vis des recommandations pour les situations des groupes G_{inf} . Pour le groupe G_{sup} , alors que notre modèle recommandait des ressources de niveau Première Générale, les experts préfèrent des exercices de même niveau scolaire mais de niveau de difficulté légèrement supérieur. Pour le groupe G_{med} , ils soulignent la nécessité d'une chronologie entre les chapitres d'un même niveau (d'après la citation d'un expert : "il faut s'assurer que le chapitre B a bien été étudié avant de suggérer le chapitre C ").

La prochaine étape de validation est de rendre la fonctionnalité de recommandation disponible au public pour approfondir la validation du système en étudiant l'utilisation effective des recommandations par les apprenants et leur impact sur leurs connaissances et sur leur utilisation de l'application. Le grand nombre d'utilisateurs permettra d'avoir des données pour faire une validation "en situation réelle" [5].

5 Conclusion et perspectives

Notre approche combinant "arbre de connaissances" et IRT nous a permis de proposer une première version d'un algorithme de recommandation de contenu "étape par étape". Contrairement au paradigme classique où de nombreuses ressources sont disponibles pour étudier un sujet précis, on se place ici dans la situation inverse où l'apprentissage se veut exploratoire avec la possibilité de transiter facilement d'un niveau à un autre.

Les évaluations par les experts semblent valider notre approche, en soulevant toutefois le problème du manque de ressources dans l'application. À partir de ces premiers résultats, nous travaillons actuellement sur une adaptation continue du processus d'Adaptive Learning, *via* une découverte automatique de connaissances et en interaction avec les acteurs du processus d'apprentissage. Cette adaptation automatique permettrait de s'adapter aux besoins de chaque apprenant et à leur comportement lors de l'utilisation de l'application.

Nous souhaitons pour cela exploiter les techniques d'apprentissage automatique et de fouille de données pour mener une analyse diagnostic permettant de construire un profil de connaissances acquises par l'étudiant, mais également une analyse comportementale pour comprendre la manière dont il exploite l'application et son niveau d'implication dans son processus d'apprentissage. Cette analyse comportementale passe par l'aspect motivation de l'élève et s'étudie également au travers de la ludification de l'application. Ces analyses exploiteront les traces d'apprentissage de l'apprenant, mais également celles de l'ensemble des utilisateurs pour faire émerger des comportements et situer l'apprenant par rapport à ces comportements. Ce profil de connaissances associé au profil comportemental permettra d'adapter les recommandations fournies et de choisir quand et comment lui faire des recommandations.

Références

1. Baker, F.B. : The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, College Park, Md., 2nd ed edn. (2001)
2. Brusilovsky, P., Peylo, C. : Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education* **13** (2003)
3. Clément, B., Roy, D., Oudeyer, P.Y., Lopes, M. : Multi-Armed Bandits for Intelligent Tutoring Systems. *Journal of Educational Data Mining* **7**(2) (2015)
4. Daher, J.B., Brun, A., Boyer, A. : Multi-source Data Mining for e-Learning. In : 7th International Symposium "From Data to Models and Back (DataMod)". Toulouse, France (2018)
5. Erdt, M., Fernandez, A., Rensing, C. : Evaluating Recommender Systems for Technology Enhanced Learning : A Quantitative Survey. *IEEE Transactions on Learning Technologies* **8**(4), 326-344 (2015)
6. Jiang, W., Pardos, Z.A., Wei, Q. : Goal-based Course Recommendation. In : Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp. 36-45. Association for Computing Machinery, New York, USA (2019)
7. Markowska-Kaczmar, U., Kwasnicka, H., Paradowski, M. : Intelligent Techniques in Personalization of Learning in e-Learning Systems. In : *Computational Intelligence for Technology Enhanced Learning* (2010)
8. Nguyen, C., Roussanaly, A., Boyer, A. : Learning Resource Recommendation : An Orchestration of Content-Based Filtering, Word Semantic Similarity and Page Ranking. In : 9th European Conference on Technology Enhanced Learning (2014)
9. Nguyen, C., Roussanaly, A., Boyer, A. : Studying Relations Between E-learning Resources to Improve the Quality of Searching and Recommendation. In : 7th International Conference on Computer Supported Education. vol. 1 (2015)
10. Vie, J.J., Kashima, H. : Knowledge Tracing Machines : Factorization Machines for Knowledge Tracing. *Conf. Artificial Intelligence AAAI* (2019)
11. Yago, H., Clemente, J., Rodriguez, D. : Competence-based recommender systems : A systematic literature review. *Behaviour & Information Technology* **37**, 1-20 (2018)