



HAL
open science

Recommandation de ressources éducatives en contexte d'apprentissage parascolaire sur application mobile

Anaëlle Badier

► To cite this version:

Anaëlle Badier. Recommandation de ressources éducatives en contexte d'apprentissage parascolaire sur application mobile. Neuvièmes Rencontres Jeunes Chercheur×e×s en EIAH, RJC-EIAH, May 2022, Lille, France. pp.118-125. hal-03846074

HAL Id: hal-03846074

<https://hal.science/hal-03846074>

Submitted on 9 Nov 2022

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.

Recommandation de ressources éducatives en contexte d'apprentissage parascolaire sur application mobile

Anaëlle Badier (D2)

Université Lyon 1, LIRIS, UMR 5205, F-69622, France
Nomad Education, Paris, France
`anaelle[dot]badier[at]liris.cnrs.fr`

Abstract. Dans cet article nous proposons un modèle de recommandation personnalisée de ressources pédagogiques afin d'orienter l'apprenant vers des ressources de remédiation ou d'approfondissement sur application mobile à usage parascolaire. Le contenu pédagogique est structuré par des notions prérequisées ou attendues, et la maîtrise de chaque chapitre est estimée par la méthode statistique de la théorie de réponse aux items. Après un pré-filtrage des ressources éligibles pour un utilisateur et une thématique donnés, un score de recommandation est calculé à partir d'un indicateur de cohérence pédagogique, pénalisé par l'historique de l'apprenant et encourageant le renouvellement des recommandations. Le système est actuellement en phase de test.

Keywords: Système de recommandation · Apprentissage Adaptatif · Application mobile parascolaire.

1 Introduction

L'augmentation des effectifs d'élèves rend difficile la mise en place d'un suivi personnalisé adapté aux besoins de chaque apprenant en classe. Les plateformes pédagogiques utilisées en parascolaire viennent en complément de la classe et permettent grâce aux nouvelles technologies de proposer à l'apprenant un suivi personnalisé de son parcours. En travaillant sur une application mobile pluridisciplinaire qui propose des ressources éducatives de tous les niveaux du collège au bac+5, nos travaux de recherche visent à identifier les besoins de l'apprenant pour le réorienter si nécessaire vers des ressources de niveau scolaire antérieur ou au contraire à l'encourager à aller étudier des ressources de niveau plus avancé, toujours en lien avec le chapitre sur lequel il souhaite travailler. Les ressources pédagogiques à disposition sont constituées de QCM de 5 à 20 questions et de mini-cours, organisés en chapitres au sein d'une discipline. Le processus de construction du système de connaissances ayant été décrit dans un précédent article [1], nous nous focaliserons ici sur les indicateurs permettant la construction d'un score de recommandation qui répond à plusieurs contraintes.

Après avoir présenté le système de recommandation existant (section 2) et introduit le cadre scientifique dans lequel s'inscrit cette nouvelle contribution

(section 3), nous développerons l'architecture et la paramétrisation d'une extension de ce modèle (section 4) pour conclure sur les travaux à venir.

2 Contexte de travail et précédente contribution : Préfiltrage des ressources par modélisation en arbre de connaissances et méthode IRT

L'application mobile dans laquelle s'intègre notre travail contient un nombre de ressources limitées pour chaque thématique abordée. L'objectif est de proposer à l'apprenant du contenu en lien avec le chapitre qu'il souhaite étudier.

Les ressources pédagogiques à recommander sont structurées *via* un "arbre de connaissances" qui fait le lien entre les différents chapitres du programme enseigné, grâce à des tags de type *prérequis* ou *attendu* (figure 1). Les notions qualifiées de *prérequis* sont des notions qu'il est souhaitable de maîtriser avant de pouvoir aborder un chapitre particulier. Les *attendues* sont des notions qui ont été découvertes ou dont le niveau de maîtrise a augmenté ou à l'issue de l'étude d'un chapitre. Une même notion peut à la fois être *prérequis* et *attendue* si l'étude du chapitre permet de monter en compétence sur cette notion. Nous ne nous inscrivons pas ici dans une approche par compétences à strictement parler mais par thématiques abordées. Les notions sont apposées sur les différents chapitres par les concepteurs des ressources pédagogiques.

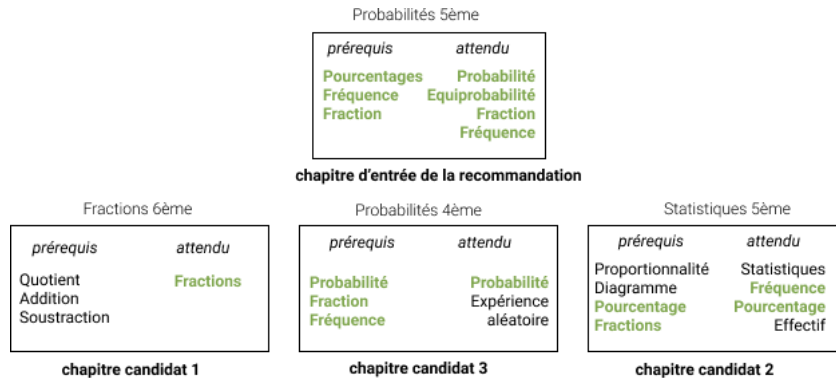


Fig. 1. Extrait de l'arbre de connaissances de notre système de recommandation pour la discipline Mathématiques. Les notions surlignées en vert sont celles que les chapitres candidats ont en commun avec le chapitre d'entrée du système.

L'utilisation de l'*Item Response Theory* (IRT) [2] permet de définir le niveau de maîtrise du chapitre qui vient d'être terminé par l'apprenant ("chapitre d'entrée de la recommandation"). Si le chapitre d'entrée est mal maîtrisé, nous considérons que l'apprenant doit retravailler les prérequis à ce chapitre, par

conséquent les chapitres candidats à la recommandation seront ceux qui ont au moins une notion en commun avec les prérequis du chapitre d'entrée, de niveau scolaire inférieur ou égal. Si le chapitre d'entrée est très bien maîtrisé, nous encourageons l'apprenant à aller explorer les chapitres de niveau supérieur ou égal à son niveau actuel qu'il peut appréhender avec ses nouvelles compétences acquises, les chapitres candidats seront ceux qui mobilisent au moins une notion en commun avec les attendus du chapitre d'entrée. Pour les apprenants de niveau moyen, tous les chapitres de niveau actuel et de niveau n-1 ayant une notion en commun avec les notions du chapitre d'entrée seront candidats. Cela laisse le choix à l'apprenant de continuer à travailler sur son année en cours ou de renforcer ses connaissances un peu fragiles avec des chapitres de niveau inférieur. Les détails et choix d'implémentation sont présentés dans l'article [1].

Le modèle développé permet donc de proposer des ressources en lien avec le niveau de maîtrise estimé de l'élève des notions mobilisées sur le dernier chapitre. Nous affinons cette recommandation en prenant également en compte l'historique de travail, les oublis dans le temps et les recommandations déjà suivies tout en conservant l'intérêt de l'apprenant pour le système en renouvelant régulièrement les contenus proposés.

3 Etat de l'art

Les revues systématiques qui répertorient les travaux en systèmes de recommandations pédagogiques montrent que les contributions concernent majoritairement la précision et la cohérence pédagogique de la recommandation proposée. Ainsi, de nombreux systèmes sont fondés sur des ressources web [3,9] pour trouver celle qui conviendra le mieux au sujet en cours d'étude. Dans notre contexte, les ressources sont internes à l'application et ne bénéficient pas de méta-données de description sur lesquelles s'appuient la plupart des systèmes dans le domaine du web pour structurer les ressources en ontologies. Une autre approche consiste à appuyer les systèmes de recommandation sur des compétences cibles comme dans les travaux de Clement et. al [5]. L'identification des compétences cibles peut s'appuyer sur des experts. L'inconvénient de ces systèmes est le domaine d'application restreint du système de recommandation, qui sera limité à la base de connaissance définie par l'expert, souvent à une discipline particulière, par exemple l'apprentissage de la programmation [11] ou des notions précises de sciences [8]. Dans notre contexte, nous cherchons à développer un système de recommandation le plus générique possible qui permette d'accompagner l'élève dans toute sa scolarité à travers toutes les disciplines, voire à aller vers des recommandations transdisciplinaires.

Pour être pertinente, une recommandation doit également prendre en compte l'historique de travail et d'apprentissage d'un apprenant. Nous pouvons par exemple citer les travaux de Choffin et. al [4] qui étudient l'aspect historique de l'apprentissage en ciblant l'effet de la répétition sur la mémorisation à long terme, et ceux de Jiang et. al [7] qui introduisent la notion de Zone de Développement Proximal (*ZPD*) développée par Vygotsky. En effet, une ressource peut être

pédagogiquement pertinente mais si l'élève s'est déjà confronté à un échec face à cette ressource, il serait décourageant de l'y confronter à nouveau trop rapidement. De même, il y a peu d'intérêt à recommander un ressource pertinente mais déjà bien maîtrisée.

Enfin, un système de recommandation doit recommander des ressources pertinentes mais également renouveler ses propositions. Si dans les algorithmes de recommandation de ressources de divertissement l'aspect sérendipité des contenus proposés est important [13], il est souvent négligé au profit de la pertinence pédagogique dans le cadre éducatif. Cet aspect est développé dans les travaux de Yu et. al [12] où la satisfaction de l'utilisateur du système est étudiée *via* des critères de sérendipité, nouveauté, inattendu et succès de la recommandation. Le fait de recommander de nouvelles ressources, permet de susciter la curiosité de l'apprenant, ce qui est un moteur de motivation intrinsèque comme démontré dans les travaux de Oudeyer et.al [10], d'autant plus dans notre contexte d'apprentissage parascolaire.

Ainsi, notre contexte de travail (un nombre de ressources internes et limitées, en apprentissage volontaire non encadré et parascolaire) soulève des contraintes qui, même si elles ont été abordées individuellement dans la littérature, posent dans notre cas la difficulté de leur prise en compte dans leur globalité pour fournir des recommandations pertinentes.

4 Un modèle de recommandation combinant différents indicateurs impactant la recommandation

Afin de proposer des recommandations adaptées, les chapitres pré-sélectionnés par le système existant (voir section 2) sont classés par pertinence selon un score de recommandation, construit comme un score de pertinence pédagogique pénalisé par un indicateur historique et un indicateur de nouveauté (formule 1). L'indicateur pédagogique priorise les chapitres de thématique et de niveau proches du chapitre qui vient d'être étudié, l'indicateur historique prend en compte ce que l'apprenant a déjà fait et son niveau de maîtrise et l'indicateur nouveauté maintient la curiosité de l'apprenant en renouvelant les recommandations. Le score pédagogique est relatif à l'arbre de connaissances et au niveau de l'utilisateur tandis que les indicateurs historique et de nouveauté sont dynamiques selon l'activité de l'utilisateur. L'importance de chacun des indicateurs dans le score de recommandation est définie par des variables de poids qui ont été initialisées de façon à accorder plus de poids à la pertinence pédagogique. Cette pertinence pédagogique a été soumise à la validation d'enseignants par deux protocoles. Dans le premier, il leur était demandé de noter la pertinence du chapitre le plus proche pédagogiquement proposé par notre système pour un chapitre d'entrée donné, et si besoin de proposer une meilleure suggestion. Dans le deuxième, les enseignants devaient sélectionner dans les chapitres proposés par l'application le plus pertinent pour un chapitre et un niveau de réussite d'entrée donnés. Cette sélection a été par la suite comparée à celle de notre système. Ce protocole a permis de valider la cohérence de notre score pédagogique. Un plus

faible poids a été attribué aux modules historiques et nouveauté, ce qui permet d'autoriser la suggestion de contenus jugés moins pertinents mais pas encore abordés ou mal maîtrisés.

$$score_{reco} = w_{pedago} * score_{pedago} + w_{hist} * score_{hist} + w_{nouveauté} * score_{nouveauté} \quad (1)$$

Module pédagogique. Le module pédagogique calcule pour les chapitres pré-sélectionnés un indicateur de pertinence pédagogique. On cherche à faire travailler l'apprenant sur des ressources relativement proches de son chapitre actuel. L'indicateur de pertinence pédagogique (formule 2) prend en compte deux composantes : la similarité avec le chapitre qui vient d'être étudié et l'éloignement au niveau scolaire de ce chapitre.

$$score_{pedago} = score_{similarité} * (1 - pénalité_{distance}) \quad (2)$$

La similarité pédagogique est calculée en prenant en compte les notions communes au chapitre candidat et au chapitre récemment étudié pour le type de tag pré-sélectionné. On utilise la méthode de la similarité cosinus entre chaque ressource préalablement vectorisée en utilisant l'index Term-Frequency Inverse-Document-Frequency (TF-IDF). Cette méthode de vectorisation permet de prendre en compte la précision des notions apposées : les notions plus génériques auront moins de poids dans l'indice de similarité que les notions précises. Les programmes de l'éducation nationale étant structurés par cycles, la pénalité de distance au niveau scolaire est choisie de façon à pénaliser davantage une distance dite "inter-cycle" (de 5e à 6e) qu'une distance "intra-cycle" (de 4e à 5e).

Module historique. L'objectif du module historique est de prendre en compte ce que l'apprenant a déjà réussi pour ne pas proposer un contenu déjà maîtrisé, dont les compétences ne sont donc *a priori* pas des leviers pour s'améliorer, ou un contenu étudié trop récemment qui risquerait de le reconfronter trop vite à un même échec. Cet indicateur d'historique est constitué de 2 composantes: la complétion et le score obtenu sur les chapitres candidats (équation 3). Nous choisissons d'introduire un facteur de mémorisation pour faire évoluer le score historique avec le temps. Ce facteur est modélisé par l'allure de la courbe d'oubli de Ebbinghaus [6] dont les constantes ont été adaptées. En effet, on modélise une mémorisation sur un temps plus long avec des connaissances acquises, là où Ebbinghaus modélisait une mémorisation immédiate de syllabes sans signification. Avec l'adaptation des constantes, la courbe d'oubli calcule 70% de rétention des informations au bout de 7 jours, diminuée à 12% au bout de 30 jours. On ne pénalise donc pas un chapitre déjà étudié il y a longtemps, en supposant qu'un rappel des connaissances de ce chapitre peut aider à mieux maîtriser le chapitre en cours d'étude.

$$score_{historique} = 1 - [w(t) * score_{completion} * score_{note}] \quad (3)$$

avec $w(t)$: indice de mémorisation calculé selon la formule d'Ebbinghaus adaptée.

Module nouveauté. Du point de vue de l'apprentissage, les modules pédagogie et historique peuvent suffire à eux seuls pour fournir des recommandations consistantes. Néanmoins, nous travaillons dans un contexte d'utilisation volontaire et parascolaire, il est donc important de varier nos recommandations pour maintenir l'attractivité. De plus, les indicateurs sur lesquels s'appuient les recommandations peuvent être imprécis. Ainsi, nous introduisons un indicateur de nouveauté qui pénalise les recommandations qui ont déjà été proposées plusieurs fois par le système (recommandations redondantes). En considérant que les choix de l'apprenant peuvent enrichir nos connaissances, nous intégrons les propositions déjà soumises dans la recommandation pour que l'apprenant voit l'offre de recommandation évoluer, et palier ainsi un effet de lassitude. Comme pour l'indicateur historique, nous introduisons un facteur de temps pour se permettre de recommander à nouveau des ressources proposées il y a longtemps. Par souci de simplicité d'initialisation, les paramètres de la courbe d'oubli sont les mêmes que pour l'indicateur historique, ce qui pose le postulat suivant : la dynamique de l'oubli de connaissances suit celle de l'oubli des recommandations affichées.

$$score_{nouveauté} = 1 - [w(t) * \overline{redondance}] \quad (4)$$

$$redondance = (3 - k)/3$$

avec $k \in [0, 3[$ position de la ressource dans le top des recommandations affichées

5 Conclusion et perspective

L'algorithme de notre moteur de recommandation est constitué de différents indicateurs pour fournir une recommandation pertinente en prenant en compte les spécificités de l'usage volontaire et parascolaire. Le système est actuellement soumis à un protocole de test en situation réelle d'utilisation auprès de 300 beta-testeurs de 3e et 2de aux profils variés. La suite des travaux consistera à analyser leurs traces d'interaction avec le système en étudiant par des méthodes statistiques l'impact du système sur des indicateurs comportementaux. N'ayant pas accès au parcours scolaire des apprenants, les critères d'évaluation de notre système ne concernent pas un gain d'apprentissage mais la modification du comportement des utilisateurs, comme le temps passé sur l'application, l'interaction avec les suggestions et l'étendue des chapitres abordés. Cela permettra de valider le système actuel et d'affiner la personnalisation par la suite en intégrant un mécanisme d'apprentissage actif. Les leviers de cet apprentissage actif seront les différents modules abordés en section 4, ainsi que leurs poids associés. Selon le comportement de l'utilisateur vis-à-vis de recommandations *a priori* moins pertinentes pédagogiquement, l'arbre des connaissances pourra être rectifié au besoin. Les valeurs des paramètres de la formule de recommandation seront modifiés dans un second temps selon l'analyse de ces traces, chaque poids étant un levier d'action. Par exemple, si l'on diminue le poids pédagogique au profit du poids de nouveauté, le système sera dans une démarche d'*enrichissement* de l'arbre de connaissances initial. La suite des travaux permettra donc d'enrichir le modèle en partant directement des données de son utilisation.

References

1. Badier, A., et al.: Système de recommandation de ressources pédagogiques pour un apprentissage sur application mobile parascolaire. pp. 294–299. EIAH (2021)
2. Baker, F.B.: The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, 2nd ed edn. (2001)
3. Brusilovsky, P., Peylo, C.: Adaptive and Intelligent Web-based Educational Systems **13** (2003)
4. Choffin, B., Popineau, F., Bourda, Y., Vie, J.J.: DAS3H: Modeling Student Learning and Forgetting for Optimally Scheduling Distributed Practice of Skills. In: Paris-Saclay Junior Conference on Data Science and Engineering (JDSE 2019) (2019-09-12)
5. Clément, B., Roy, D., Oudeyer, P.Y., Lopes, M.: Multi-Armed Bandits for Intelligent Tutoring Systems **7**(2), 30 (2015)
6. Ebbinghaus (1885), H.: Memory: A Contribution to Experimental Psychology **20**(4), 155–156 (2013-10)
7. Jiang, W., et al.: Goal-based Course Recommendation. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp. 36–45. LAK19 (2019-03-04)
8. Kinnebrew, J.S., Loretz, K.M., Biswas, G.: A Contextualized, Differential Sequence Mining Method to Derive Students’ Learning Behavior Patterns. Journal of Educational Data Mining **5**(1), 190–219 (2013-05-01)
9. Klačnja-Milićević, A., Ivanović, M., Nanopoulos, A.: Recommender systems in e-learning environments. Artificial Intelligence Review **44**(4), 571–604 (2015-12)
10. Oudeyer, P.Y., Gottlieb, J., Lopes, M.: Intrinsic motivation, curiosity, and learning. Progress in Brain Research **229**, 257–284 (2016)
11. Venant, R., et al.: Étude du comportement des apprenants dans les travaux pratiques et de sa corrélation avec la performance académique **25** (2019-02-08)
12. Yu, R., Pardos, Z., Chau, H., Brusilovsky, P.: Orienting Students to Course Recommendations Using Three Types of Explanation. pp. 238–245. ACM UMAP (2021-06-21)
13. Ziarani, R.J., Ravanmehr, R.: Serendipity in Recommender Systems. Journal of Computer Science and Technology **36**(2), 375–396 (2021-04-01)