## A Bandit-Based Ensemble Framework for Exploration/ Exploitation of Diverse Recommendation Components: An Experimental Study within E-Commerce

## BJÖRN BRODÉN and MIKAEL HAMMAR, Apptus Technologies, Sweden BENGT J. NILSSON and DIMITRIS PARASCHAKIS, Malmö University, Sweden

This work presents an extension of Thompson Sampling bandit policy for orchestrating the collection of base recommendation algorithms for e-commerce. We focus on the problem of item-to-item recommendations, for which multiple behavioral and attribute-based predictors are provided to an ensemble learner. In addition, we detail the construction of a personalized predictor based on *k*-Nearest Neighbors (*k*NN), with temporal decay capabilities and event weighting. We show how to adapt Thompson Sampling to realistic situations when neither action availability nor reward stationarity is guaranteed. Furthermore, we investigate the effects of priming the sampler with pre-set parameters of reward probability distributions by utilizing the product catalog and/or event history, when such information is available. We report our experimental results based on the analysis of three real-world e-commerce datasets.

CCS Concepts: • Information systems  $\rightarrow$  Electronic commerce; Recommender systems; Collaborative filtering; Information retrieval query processing; • Computing methodologies  $\rightarrow$  Sequential decision making; Learning from implicit feedback; Ensemble methods;

Additional Key Words and Phrases: E-commerce recommender systems, streaming recommendations, multiarm bandit ensembles, session-based recommendations, Thompson Sampling, reinforcement learning

## **ACM Reference format:**

Björn Brodén, Mikael Hammar, Bengt J. Nilsson, and Dimitris Paraschakis. 2019. A Bandit-Based Ensemble Framework for Exploration/Exploitation of Diverse Recommendation Components: An Experimental Study within E-Commerce. *ACM Trans. Interact. Intell. Syst.* 10, 1, Article 4 (August 2019), 32 pages. https://doi.org/10.1145/3237187

## **1 INTRODUCTION**

Dynamically responding to user queries with top-*N* item ranking is a very common problem on the web, which pertains to search, advertising, and recommendations. A specific use case of interest is generating *item-to-item* (i2i) recommendations for every visited product page with the aim to maximize product sales. Such *non-personalized* recommendations are useful when user profiles are limited or non-existent. This is very typical of online shopping sessions, where most visitors

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

https://doi.org/10.1145/3237187

ACM Transactions on Interactive Intelligent Systems, Vol. 10, No. 1, Article 4. Publication date: August 2019.

The reviewing of this article was managed by special issue associate editors Mark Billinghurst, Margaret Burnett, and Aaron Quigley.

Authors' addresses: B. Brodén and M. Hammar, Apptus Technologies, Trollebergsvägen 5, 222 29, Lund, Sweden; emails: {bjorn.broden, mikael.hammar}@apptus.com; B. J. Nilsson and D. Paraschakis, Malmö University, Nordenskiöldsgatan 1A, 211 19, Malmö, Sweden; emails: {bengt.nilsson.TS, dimitris.paraschakis}@mau.se.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

<sup>2160-6455/2019/08-</sup>ART4 \$15.00