Video Game Recommender
Research Internship @ Inria Rennes, University of Rennes, France

1 Supervision

Advisers	Paolo Cremonesi, DEIB - Politecnico di Milano `paolo.cremonesi@polimi.it`, 0223993517
	Davide Frey, Inria Rennes, University of Rennes `davide.frey@inria.fr`, +33299847565
Lab	Centre Inria Rennes, IRISA (UMR 6074) / Blacknut
Team	WIDE (The World Is DistributEd)
	Équipes-Projet Inria / Département D1 IRISA

2 Context

Blacknut is a start-up company that provides a video-game streaming service. Blacknut subscribers can enjoy a selection of more than 360 video games on their TV screen, mobile device, or laptop without requiring expensive dedicated hardware. The video game runs on the cloud on Blacknut's servers and streams content directly to user devices.

Today users can manually browse a catalog of 360 video games, but Blacknut (`http://www.blacknut.com`) aims to offer a recommendation service that will allow its customers to experience new content at each visit. This masters thesis leverages the creation of this recommendation service to study its impact on the behavior of Blacknut's users and to understand how one can predict the performance of a recommendation system in a production setting.

3 Background

Recommender system have found application in a variety of online settings including e-commerce [5], social networks [1] or news websites [4, 2]. As a result both academia and industry have proposed a variety of algorithms in an effort to provide more and more accurate recommendations. Two families of techniques exist to evaluate the accuracy of recommendation [3]. On the one hand, deployed recommender systems can be evaluating by directly measuring their impact on the target application in terms of click rate, sales, or other application-specific metrics through a technique known as A/B testing. In short, different sets of users unknowingly test different recommendation algorithms, providing performance measures in terms of click/purchase rates.

On the other hand, most recommendation algorithms receive an offline evaluation on existing datasets with a variety of metrics. Unfortunately, recent research has shown that these offline metrics often offer contradicting results when comparing different algorithms with one another. Early recommenders were often evaluated using error metrics such as RMSE (root mean squared error) or MAE (mean average error), while more recent research favors accuracy metrics such as precision, recall, or fallout. Not only do these families of metrics offer diverging perspectives, but even the results of a single metric differ depending on the evaluation protocol employed in the evaluation [6]. Finding the best way to evaluate a recommender system without access to a production environment therefore remains an important open research question.

4 Objective

The goal of this masters thesis consists in answering the question of whether it is possible to establish an offline benchmark that correlates well with the performance of a recommendation system in a production environment. To this end, we will leverage the creation of Blacknut's new recommendation service in order to explore whether existing or novel offline metrics can predict the impact of a recommendation algorithm on online metrics such as the click rate.

5 Tasks

- The first task will consist in analyzing datasets that collect the behavior of Blacknut's users in order to identify which recommender algorithms are most likely to provide good results.

- Then the student will implement a recommendation framework that will be able to operate both on the datasets and in a production environment.

- Next, the student will conduct an experimentation program that tests a variety of recommendation algorithms with different sets of users over a period of time.

- Finally, the student will analyze the results of the experimentation using online metrics and study their correlation with those of offline metrics applied to datasets that capture user behavior without and with the recommender system.

6 Candidate Profile

The student is expected to have reasonable knowledge about the main recommendation algorithms and their evaluation. He or she is also expected to be proficient in Object-Oriented Programming with good software engineering skills. Blacknut's backend platform being written in Go, knowledge of this language can be a plus although it is not strictly required. Fluency in English or French is required.

7 Information about Inria Rennes and Logisitics

This thesis project will be hosted by the WIDE team of Inria Rennes. Established in 1967, Inria is the only French public research body fully dedicated to computational sciences. Inria's missions are to produce outstanding research in the computing and mathematical fields of digital sciences and to ensure the impact of this research on the economy and society. The WIDE team is located in the Inria center of Rennes, Brittany. It consists of a 15+ strong group of researchers, academics, and PhD students focusing on the design, evaluation and implementation of large-scale computing systems. A group of researchers in WIDE has been focusing on recommender systems in the last 10 years. This effort led to the creation of Mediego (`http://www.mediego.com`), a company offering news recommendation services. Interactions with this company will be possible particularly in the final parts of the project. The majority of the work will be carried our in collaboration with Blacknut, another startup company, which instead focuses on video-game streaming. The student will normally be based at Inria Rennes, but we expect he/she will spend between 10% and 20% of his/her time in Blacknut's premises.

Inria Rennes is located in the Beaulieu campus of the University of Rennes. It is easily reachable in a few minutes from the city center by bus, bicycle, or car. Blacknut is also easily reachable and is only a 10-minute bicycle ride from the University campus.

References

[1] Recommending items to more than a billion people. `https://code.facebook.com/posts/861999383875667/recommending-items-to-more-than-a-billion-people/`. Accessed on 19/09/2017.

[2] Antoine Boutet, Davide Frey, Rachid Guerraoui, Arnaud Jégou, and Anne-Marie Kermarrec. Whatsup: A decentralized instant news recommender. In *Parallel & Distributed Processing (IPDPS), 2013 IEEE 27th International Symposium on*, pages 741–752. IEEE, 2013.

[3] Paolo Cremonesi, Franca Garzotto, and Roberto Turrin. User-centric vs. system-centric evaluation of recommender systems. In Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2013*, pages 334–351, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[4] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*, pages 271–280. ACM, 2007.

[5] Greg Linden, Brent Smith, and Jeremy York. Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(1):76–80, 2003.

[6] Daniel Valcarce, Alejandro Bellogín, Javier Parapar, and Pablo Castells. On the robustness and discriminative power of information retrieval metrics for top-n recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, RecSys '18, pages 260–268, New York, NY, USA, 2018. ACM.