

# BIG DATA PROCESSING AND IMPROVING DATA CLUSTERING METHOD FOR AGENT-BASED RECOMMENDATION SYSTEM ON MOBILE PLATFORMS

XỬ LÝ DỮ LIỆU LỚN VÀ NÂNG CAO HIỆU QUẢ PHƯƠNG PHÁP PHÂN NHÓM DỮ LIỆU CHO HỆ THỐNG KHUYẾN NGHỊ AGENT-BASED TRÊN HỆ THỐNG DI ĐỘNG

Vương Xuân Hương<sup>1\*</sup>, Nguyễn Trung Tuấn<sup>1</sup>

## ABSTRACT

Today, with new technologies and technical trends like social media, mobile devices and Internet of Things, much more data are created than we would ever amass. The massive amount of data mix of structured data and unstructured data come from every sensor on every piece of equipment that's connected. The sheer amount of information provides an opportunity for an organization to identify interesting trends but it also creates difficulties for business to process with existing tools. In this paper we present an approach on processing big data and an agent-based recommender for mobile platform. The inputs for our big data process will be a test data set from a big social type forum. From that we can apply the same algorithm for mobile data. Later we created a social media platform which provides inputs for our recommender and stored in a semantic database. We propose a method to increase the precise for data clustering, from that we will directly enhance the recommendation results.

**Keywords:** *Big data, recommendation system, data clustering.*

## TÓM TẮT

Ngày nay, với sự ra đời của các trào lưu công nghệ mới như mạng xã hội, các thiết bị di động và mạng lưới vạn vật kết nối internet (Internet of Things), đã có một lượng dữ liệu khổng lồ được sinh ra nhiều hơn rất nhiều so với tất cả dữ liệu đã tích lũy từ trước tới nay. Khối lượng khổng lồ của cả dữ liệu có cấu trúc và không có cấu trúc đến từ mọi cảm biến trên từng bộ phận của các thiết bị kết nối với nhau. Có rất nhiều dữ liệu trong tay sẽ cung cấp cơ hội cho các tổ chức phát hiện ra những xu hướng thú vị nhưng nó cũng có những thách thức cho các doanh nghiệp khi xử lý bằng những công cụ sẵn có. Trong bài báo này, chúng tôi trình bày một phương án tiếp cận việc xử lý dữ liệu lớn và hệ thống khuyến nghị agent-based cho hệ thống di động. Đầu vào cho xử lý dữ liệu lớn chúng tôi sẽ sử dụng bộ dữ liệu thử nghiệm của diễn đàn mạng xã hội lớn. Từ đó, chúng tôi áp dụng cùng giải thuật đó cho dữ liệu di động. Sau đó, chúng tôi tạo ra hệ thống mạng xã hội trên di động để thu thập dữ liệu cho hệ thống khuyến nghị và lưu trong cơ sở dữ liệu ngữ nghĩa. Chúng tôi đề xuất một phương pháp để tăng độ chính xác cho việc phân nhóm dữ liệu, từ đó, chúng ta sẽ cải thiện trực tiếp kết quả của hệ thống khuyến nghị.

**Từ khóa:** *Dữ liệu lớn, hệ thống khuyến nghị, phân nhóm dữ liệu.*

---

<sup>1</sup>Viện Công nghệ thông tin kinh tế, Trường Đại học Kinh tế quốc dân

\*E-mail: huongvx@neu.edu.vn

Ngày nhận bài: 08/07/2016

Ngày nhận bài sửa sau phản biện: 25/07/2016

Ngày chấp nhận đăng: 15/08/2016