Self-Tuning Approach for Implementing a Multidimensional Recommendation System Using PID

Saurabh Chaudhari
Dept. of Computer Engineering
STES SKNCOE
Pune, India
Email: sach16795@gmail.com

Aditya Sawant
Dept. of Computer Engineering
STES SKNCOE
Pune, India
Email: aditya33sawant@gmail.com

Reeta Patil
Dept. of Computer Engineering
STES SKNCOE
Pune, India
Email: reetapatil1795@gmail.com

***Abstract*—The exponential growth in the online retail services and products has resulted in the generation of huge amounts of data. In an online retail portal, analysis of the data plays an important role in optimization of recommendation systems, which helps offer better user experience. As large amount of data is consistently updated in the databases of online retail portals, the data quantity keeps on increasing and it becomes more difficult to carry out analysis for the process of recommendation. Recommendation algorithms have been designed to take some input information and generate relevant recommendations. The performance of traditional recommendation systems and analysis techniques degrade while processing large data. New recommender system technologies are needed that can quickly produce high-quality recommendations, even for very large-scale problems. Here is a newly proposed algorithmic multidimensional approach that can be deployed to improve recommendation systems performance. PID is a closed loop, self-tuning algorithm which is predominantly implemented in mechatronics instruments where manual supervision is not feasible and it functions for correction of errors quantifiable in physical measures. If we were to map errors in physical dimensions to irrelevancy in computer systems it can be approximately stated to be higher the error, higher the irrelevancy. Also, the availability of multiple factor assessment in PID algorithms can be used to implement the multidimensional approach. As proposed and analyzed the implementation is expected to provide with faster as well as big data compliant analytics option.**

***Keywords*— Data Mining, Information Filtering, PID Controller, Recommendation System, Hybrid Recommendation System, Attribute Extraction.**

I. Introduction

With a large number of items being sold on online retail services, it would almost be impossible and impractical to browse the entire catalog of items which we want to buy. By making the necessary searches, one can find the necessary items pertaining to the search query. Instead of constantly using the search query, recommendation systems are built in order to provide the user with a list of items without having to make another search query; thereby enhancing user experience.

Much work has been done both in the industrial and research fields on developing new approaches to implementing recommender systems [3] [4] [5]. The interest in this area is huge because it constitutes a problem-rich research area and because of the abundance of practical applications which could help users deal with the paradox of choice created by information overload and provide personalized recommendations for items [1]. Recommender systems should thus be able to list products that will likely be preferred by estimating user preferences. So far, two major techniques have been proposed: collaborative and content-based filtering and they have complementary properties [4].

To mitigate the shortcomings of the above techniques, we propose a hybrid recommendation system that is multidimensional and uses parameters such as both the users rating, discount, price, item specifications and site demographic data.

II. Related Topics

A. Recommender Systems

Recommender systems solve the problem of information overload for a user by searching through large volumes of dynamically generated data, to provide users with personalized content and services [1]. The main objective of recommendation systems for an online service portal is to rank and list items which have not been seen by the user or which the user has not searched or rated or seen on the online shopping portal. Using additional metadata about the items or transactions can help build the multi-dimensional recommendation system when integrated together with the PID algorithm.

Recommender systems are beneficial to both online retail services and users because they reduce transaction costs of finding and selecting items in an online shopping environment while presenting the appropriate items as recommendations. In an e-commerce setting, recommender systems enhance revenues as they are effective means of selling more products. From the perspective of E-commerce, recommendation systems can be viewed as a tool that helps users search through

978-1-5090-5682-8 /17/$31.00

records of knowledge which is related to user's interests and preferences.

In recent years, various approaches for building recommendation systems have been developed, which can utilize either collaborative filtering, content-based filtering or hybrid filtering[1][3][4][5].

Collaborative methods recommend products by considering how other users with similar taste and it uses their opinion to recommend items to the active user. Collaborative approaches exhibit issues such as cold-start, sparsity, and scalability problems [4] [5].

Content-based filtering techniques normally base their recommendations on users information, and they ignore contributions from other users. Some of the issues associated with content-based filtering techniques are limited content analysis and overspecialization- which results in sparsity of data [4] [5].

The hybrid approach combines content-based and collaborative methods in order to mitigate the problems faced by the two filtering techniques individually. Hybrid recommendation systems can also be augmented by various algorithms and knowledge-based techniques to further improve the accuracy of recommendation systems [4]. When we specifically consider knowledge-based systems, there arises a need for knowledge acquisition, a well-known bottleneck for many artificial intelligence applications. In the proposed system, control loop theory is integrated into the core algorithm in order to achieve higher more accurate recommendation lists.

B. PID Control Theory

With its three-term functionality covering treatment to both transient and steady-state responses, Proportional-Integral-Derivative (PID) control offers the simplest and yet most efficient solution to many real-world control problems. Since its invention in 1910, owing to Elmer Sperrys ship auto-pilot and Zeigler-Nichols (Z-N) straightforward tuning methods in 1942 [7], the popularity of PID controllers has grown tremendously. With the consequent advances in digital technology, the science of automatic control now offers a wide spectrum of choices for control schemes. Its wide application has stimulated and sustained the development of various PID tuning techniques, hardware modules, and sophisticated software packages [6].

A feedback controller is designed to generate an output that causes some corrective effort to be applied to a process so as to drive a measurable process variable towards the desired value known as the set-point. The controller uses an actuator to affect the process and a sensor to measure the results. Feedback controllers determine the output by observing the error between the setpoint and a measurement of the process variable. Errors would occur when an operator changes the set-point intentionally or when some disturbance or load on the process changes the process variable accidentally. The controller's mission is to eliminate the error automatically.

A standard PID controller has the full transfer function [9].

$$u(t) = \overbrace{K_p e(t)}^{\text{Proportional}} + \overbrace{K_i \int_0^t e(\tau)d\tau}^{\text{Integral}} + \overbrace{K_d \frac{d}{dt} e(t)}^{\text{Derivative}}$$

Where Kp >> Kd > Ki.

Proportional term (Kp) - responsible for overall control action which is proportional to the error signal [10].

Integral term (Ki) - responsible for reducing the steady state errors through compensation by an integrator [10].

Derivative term (Kd) -responsible for improving response through compensation by a differentiator [10].

Modern control theory provides a powerful set of tools to deal with dynamical systems- such as recommendation systems. Hence, by applying the principles of PID algorithm, we can construct and maintain a stable and robust recommender system for dynamically evolving environments.

By considering the various attributes of items being sold and the portal's users, we propose mapping the error in PID system to the relevancy of an item, which can be recommended to a user- thereby improving accuracy owing to the simple and self-tuning nature of PID algorithm.

III. PROPOSED SYSTEM

To address the limitations of traditional methods such as collaborative and content-based techniques we have devised a hybrid approach which makes use of PID control theory along with common techniques such as clustering and collaborative filtering[11][5].

The Basic system architecture is designed keeping in view structure of a common online retail portal. It comprises of three main units namely Database unit, Business unit, and User. Database unit is a repository of all the item, user and transaction information. The Business unit is the core of the recommendation system where all the filtering and processing would be carried out.

An administrator is responsible for supervising the overall functioning of the business unit. In the business logic following operations will be carried out:

1. Import data from the database.

2. Cluster the items in their respective categories and subcategories.

3. Aggregate products into datasets which vary by some interval in parametric values such as price, specification or ratings.

4. Calculate the Product Equivalence Value (PEV).

5. Split the available datasets into training and testing data depending on factors such as volume or measure of relevance.

The Product Equivalence Value (PEV) is calculated over various factors ranging from attributes of the product such as price, rating, discount, item specification and perhaps even transactions and site demographic data. Various product attributes will be clubbed into either proportional, integral or derivative class depending on the importance of the parameter in determining the relevancy of the product for the recommendation. The presence of three classes would allow us to integrate multiple parameters in a single class, resulting in a level of abstraction which would provide appropriate recommendations [12]. Since it exhibits multi-dimensionality, the results would be more accurate recommendations obtained over a simple and self-tuning technique.

For instance, the Proportional class will contain most important characteristics such as price and discount. We assign these attributes to this class because it is effectively responsible for the purchase of a product. Once a normalized scale is proposed by the organization, the products can hence be evaluated on those parameters. This system allows the flexibility of choosing the attributes preference so as to suit organizational needs. Quite similar to the proportional parameter in PID systems for error reduction, it would assess the current status of the products. For e.g. price and offers on products which tend to change continuously.

The Integral class will contain secondary parameters such as ratings, reviews etc. Also, the significantly low value of integral constant will negate for the availability of a high number of ratings and hence reduce the influence of the mere presence of a high number of ratings. As the integral parameter in PID algorithm is supposed to check for the accumulated errors, it would, in this system check for accumulated data attributes of the products. For e.g. ratings and reviews for the products, demographic data of the various users gathered over time can be used for fine tuning of the recommendation lists to be generated.

The Derivative class can be either used to negate the abstract overshooting effect caused by values in integral class or by the system administrator to fine tune the system according to organizational needs.

Once functioning, the system would use ratings, reviews, administrator's input and other values that change over a period of time as feedback for further fine tuning. For every transaction or rating given necessary changes can be made in the training data set which would slightly alter the PEV of the respective product. This provides the system with some self-tuning aspect or feedback mechanism. Finally, the results generated can be filtered in accordance with the user query and the recommendations can be displayed.

IV. Mathematical Model

Let the S be the set containing Input, Output, Functions.

S = {I, O, F}

Where,

I is the set of Input,

O denotes output set,

F denotes Functions used.

I = {T, D}

Where,

T is Training data set,

D is Data Value set,

Training data set consists of User information and inventory item details. Data Value set consists of Values for Kp, Ki, Kd.

O = {O1, O2, O3.., On}

Where,

O1–On is set of items recommended by system

F = {P, I, D, N}

Where,

P is a function to obtain proportional equivalent,

I is a function to obtain Integral equivalent,

D is function to obtain Derivational equivalent,

N is a function to obtain Final normalization equivalent.

V. Conclusion

Over the past few years, recommendation engines have made significant progress as various collaborative, content-based and hybrid methods have been proposed. The primitive methods are not as effective when used in the broader range of applications. Since the approach proposed in this paper makes use of PID algorithm, it can mitigate this drawback of primitive methods. As PID control theory can deal with dynamic systems, the proposed recommendation system can be used to enhance the operation of recommendation systems. Based on the concept of mapping the error in PID algorithm to the relevancy of an item in the recommendation list we are able to produce the items to be recommended to the user.

The system exhibits multi-dimensionality by providing various parameters of items such as price, ratings, discount, etc. thereby generating a more comprehensive recommendation list with optimum performance as it is coupled with PID algorithm. We hope that the proposed system presented in this paper can find its application in various recommendation problems in the future and contribute to discussions in the recommendation system community.

As future work, we intend to extend self-tuning attribute to define optimal value of constants in PID algorithm.

References

[1] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, " Recommendation systems: Principles, methods and evaluation", Egyptian Informatics Journal Volume 16, Issue 3, November 2015.

[2] Xing Bo-yang, Yu Li-ye, Zhou Zhong-Kai, " Composite single neural PID controller based on fuzzy self-tuning gain and RBF network identification", Control and Decision Conference (2014 CCDC), 2014.

[3] Gediminas Adomavicius, Alexander Tuzhilin, " Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art

and Possible Extensions", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 17, NO. 6, JUNE 2005.

[4] Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno, "An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model". IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 16, NO. 2, FEBRUARY 2008.

[5] Linden, G., Smith, B. and York, J. "Amazon.com recommendations: Item-to-item collaborative filtering" Industry Report-IEEE Internet Computing, Jan - Feb 2003, 7(1): pages 76 – 80.

[6] Kiam Heong Ang, Gregory Chong and Yun Li, "PID Control System Analysis, Design and Technology", IEEE Transactions On Control Systems Technology, Vol 13, No.4, July 2005.

[7] J.G Zeigler, N.B. Nichols, "Optimum Settings for automatic controllers", Trans ASME, vol. 64, pp. 759-768,1942.

[8] Vance J. VanDoren, Understanding PID Control, Control Engineering Online Magazine, February 1996.

[9] Fengying Ma, Fengying Ma, "An Improved Fuzzy PID Control Algorithm Applied in Liquid Mixing System", International Conference on Information and Automation, Hailar, China, July 2014.

[10] Kiam Heong Ang, Gregory Chong and Yun Li, "PID Control System Analysis, Design and Technology", IEEE Transactions On Control Systems Technology, Vol 13, No.4, July 2005.

[11] Kenneth Wai-Ting Leung, Wilfred Ng, and Dik Lun Lee, "Personalized Concept-Based Clustering of Search Engine Queries", Knowledge and Data Engineering IEEE Transactions on, vol. 20, pp. 1505-1518, 2008, ISSN 1041-4347.

[12] Tamas Jambor, Jun Wang and Neal Lathia "Using Control Theory For Stable and Efficient Recommender Systems", in Proceedings 21st international conference on World Wide Web, Pages 11-20, ACM New York, NY, USA ©2012, ISBN: 978-1-4503-1229-5