Recommending Functions in Spreadsheets from the Fuse Corpus

Shaown Sarker, Matthew Neal, Nisarg Vinchhi
Department of Computer Science
North Carolina State University
Raleigh, NC 27606
{ssarker,meneal,nvinchh}@ncsu.edu

Project Objective

Spreadsheets are the most common form of end-user programming. Although spreadsheets have a large array of functions built-in, spreadsheet users often do not exploit them to perform their tasks efficiently. To address this issue, in this project we investigate recommender system technologies and consider two distinct approaches to a collaborative filtering based function recommender system for spreadsheets.

Our main inspiration comes from prior research on recommending commands in AutoCAD [1] and Eclipse [2]. The system described in both papers utilized user-based and item-based collaborative filtering algorithms on a collective users' command usage history to recommend personalized commands given the usage history of an individual user as input. We intend to use the same algorithms on Fuse[1], the largest and most diverse spreadsheet corpora to date, to recommend personalized functions for an individual given her function usage in the form of a collection of spreadsheets.

One key difference in our approach is that unlike the prior command recommendation systems which had access to sequential command usage history, we do not have similar temporal information as spreadsheet files in the corpus are static and does not contain such function usage history. We will depend on function usage frequencies instead to utilize the recommender algorithms. Our work will involve feature vector extraction from the Fuse corpus, applying the collaborative filtering algorithms to recommend functions for the input user, and validate the efficacy of the recommendations by cross validation.

Team Members

Our team consists of three members: Shaown Sarker will be involved in feature extraction and implementation of one of the algorithms, Matthew Neal will be implementing the other algorithm, and Nisarg Vinchhi will be involved in the cross validation. By April 5th, we intend to have implemented both the algorithms and the framework required for the cross validation.

References

[1] Matejka, J., Li, W., Grossman, T., & Fitzmaurice, G. (2009). CommunityCommands: command recommendations for software applications. *In Proceedings of the 22nd annual ACM symposium on User interface software and technology (pp. 193-202).* ACM.

[2] Murphy-Hill, E., Jiresal, R., & Murphy, G. C. (2012). Improving software developers' fluency by recommending development environment commands. *In Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering (p. 42).* ACM.

[1] http://static.barik.net/fuse/