報告番号	※甲　　　第　　　　号

主　論　文　の　要　旨

論文題目　Similarity Metrics in Neighborhood-Based Methods for the Completion of Partially Observed Data Matrices（近接解析手法による部分観測データ行列補完のための類似度指標）

氏　　名　ORTAL VITE Patricia Isabel

論 文 内 容 の 要 旨

Matrix completion is a widespread task in the machine learning field because of the often occurrence of missing data in real-world scenarios. It involves the estimation of unobserved values of partially observed data represented in tabular form. As part of the collaborative filtering techniques, neighborhood-based methods are widely used to address the data matrix completion problem. While neighborhood-based methods heavily depend on similarity metrics to compute accurate estimations, few studies have focused on the design of similarity measures and instead rely on conventional metrics regardless the problem to solve. This dissertation presents two neighborhood-based methods to solve two different practical data matrix completion problems respectively. The proposed approaches consist on leveraging the known information available from different sources, computing clusters that exhibit close patterns or trends, and then estimating the unknown values based on these groups and similarities. Furthermore, the metrics to establish the degree of similarity between agents are designed to be domain-specific, and free from the limitations of conventional similarity measures. In particular, this work focuses on evaluating the effects that custom measures have in terms of prediction performance. The first proposed method encodes products as a sequence of attributes, each of which represents a different dimension of the consumer perception. The second method is a switching hybrid recommender that estimates item ratings addressing the sparsity problem that affects the performance of collaborative filtering techniques. The results of experiments conducted using real-world and synthetic data indicate that the proposed methods have superior performance compared to conventional approaches in terms of mean absolute error (MAE) and root mean square error (RMSE).