Abstract

A rising trend in data mining is a High utility item sets (HUIs) mining. It aims to find all item sets which have an utility which meets a client determined least utility edge min_util. But, for clients, it is an issue to set a min_util efficiently. So, it is not proper procedure for clients to find a least utility edge by experimentation. An excessive number of HUIs will be produced, in the case that min_util is set very low. Due to this the mining procedure may result wasteful. It is also possible that no HUIs be found, if min_util is set very high. So for addressing the above issues, we redefine the problem of high utility item sets (HUIs) mining by top-k high utility item sets (top-k HUI) mining. Here, desired number of HUIs to be mined is k. Two different algorithms which are named as TKU and TKO (mining Top-K Utility item sets in two stages, mining Top-K utility item sets in one stage, respectively) are proposed for mining the item sets without setting the value of min_util. We apply pre-evaluation strategy to algorithms to improve the performance.

Index Terms

Computer Science Algorithms

Keywords

Utility mining, high utility item set mining, top-k high utility item set mining, frequent item set, transactional database.