

SEARCHING LARGE DATA SPACE FOR STATISTICALLY SIGNIFICANT PATTERNS

ABSTRACT

5 According to embodiments of the present disclosure, a method and a distributed processing system are provided to discover statistically significant patterns from arbitrarily large data set by statistical analysis. The present disclosure provides new distributed system and algorithm of detecting statistical patterns of different orders. Also, the present disclosure provides effectively traversing data domain for pattern candidate generation that supports
10 multi-agent distributed computation model. By increasing and decreasing the number of agents, the system is able to handle bigger or smaller problems. Further, the present disclosure provides a scheme of partitioning data in distributed storage more efficiently for statistical analysis.

15 FIG. 1D