This research paper presents a new approach for optimizing big data analytics in parallel distributed systems. The existing works in this field primarily focus on reducing the communication cost of complex join queries, but ignore the computational intensity of the query. To address this issue, the authors propose the Adaptive Distributed Join (ADJ) approach, which co-optimizes communication, pre-computing, and computation costs by exploring cost-effective partial results through cost-estimation by sampling. The results of their experiments indicate that the ADJ approach outperforms existing multi-way join methods by a significant margin.

The paper starts with a clear introduction that outlines the problem being addressed and the solution proposed. The authors provide a comparison between the state-of-the-art HCubeJ algorithm and their own approach. The authors then present a detailed description of their proposed method and the results of their experimental study. The experimental study compares the performance of ADJ with four other state-of-the-art methods and investigates the effectiveness of attribute order pruning and co-optimization. The results show that ADJ provides near-linear speed up on two out of three queries and has limited scalability on the third query.

The conclusion of the paper provides a clear summary of the research performed, however, it could benefit from a statement on the significance of the results and a comparison with existing methods. The related work section provides a good overview of previous works in this field, but could benefit from a more detailed analysis on the influence of previous works on the design of the authors' approach.