



سید علیرضا طباطبایی پور، ۱۴۰۵

## بهینه‌سازی پرس‌وجوهای الحاق در پایگاه داده‌های توزیع شده به وسیله‌ی الگوریتم باغچه

در این تحقیق مسئله بهینه‌سازی اجرای پرس‌وجوهای الحاق در پایگاه‌های داده توزیع شده مورد بررسی قرار گرفته است. با افزایش داده‌ها و نیاز به پردازش هم‌زمان، عملکرد بهینه در این نوع پایگاه‌ها اهمیتی حیاتی دارد، زیرا پرس‌وجوهای الحاقی به دلیل تبادل حجم بالای داده میان گره‌ها، از پرهزینه‌ترین عملیات محسوب می‌شوند. هدف این پژوهش ارائه روشی کارآمد بر پایه الگوریتم باغچه است تا با الهام از فرآیند رشد طبیعی در گیاهان، ترکیبی متوازن از جست‌وجوی سراسری و بهبود محلی ایجاد شود. روش پژوهش به صورت ترکیبی از مدل‌سازی مفهومی، شبیه‌سازی الگوریتم در محیط MATLAB و اجرای واقعی در پایگاه داده PostgreSQL صورت گرفته است. در این فرآیند، جامعه‌ای از طرح‌های پرس‌وجو تولید، ارزیابی، و از طریق مراحل رشد، رقابت و به‌روزرسانی بهینه شده‌اند. داده‌های آزمایشی از مجموعه استاندارد TPC-H برای آزمون عملکرد استفاده شد. نتایج نشان داد که الگوریتم باغچه در مقایسه با الگوریتم‌های دیگر مانند جستجوی هارمونی (HS) و کلونی زنبور عسل مصنوعی (ABC) در شاخص‌های کلیدی از جمله کاهش هزینه ارزیابی پرس‌وجو (QEC)، کاهش زمان اجرا (ETF) و بهبود مصرف منابع سیستم (RUF) عملکرد بهتری داشته است. همچنین شاخص کیفیت طرح پرس‌وجو (QPQI) به طور میانگین ?? درصد بهبود یافت و نرخ همگرایی الگوریتم تا ?? درصد سریع‌تر از رقبا گزارش شد. این یافته‌ها نشان می‌دهد که الگوریتم باغچه گزینه‌ای کارآمد برای بهینه‌سازی پرس‌وجوهای پیچیده در سامانه‌های توزیع شده است.

**کلیدواژه‌ها:** کلیدواژه‌ها: بهینه‌سازی پرس‌وجو، پرس‌وجوهای الحاق، پایگاه داده توزیع شده، الگوریتم باغچه، الگوریتم

شماره‌ی پایان‌نامه: ۱۲۷۲۲۹۸۰۷۹۰۴۲۶۳۹۰۳۰۳۱۱۶۳۱۷۰۸۵۴

تاریخ دفاع: ۱۴۰۵/۰۳/۲۵

رشته‌ی تحصیلی:

دانشکده:

استاد راهنما: دکتر علی اکبر نقابی

### **Thesis:**

## Optimization of Join Queries in Distributed Databases Using the Orchard Algorithm

This research addresses the issue of optimizing the execution of join queries in distributed databases. With the increase in data volume and the need for concurrent processing, optimal



performance in such databases becomes critical, as join queries are among the most costly operations due to large data exchanges across nodes. The objective of this study is to propose an efficient method based on the Orchard Algorithm, inspired by natural plant growth processes, to create a balanced combination of global search and local improvement. The methodology combines conceptual modeling, algorithm simulation in the MATLAB environment, and real implementation in the PostgreSQL database. During this process, a population of query plans is generated, evaluated, and optimized through stages of growth, competition, and update. Test data from the standard TPC-H suite was used for performance evaluation. Results showed that compared to other algorithms such as Harmony Search (HS) and Artificial Bee Colony (ABC), the Orchard Algorithm performed better on key metrics including reduction of Query Evaluation Cost (QEC), execution time (Execution Time Function (ETF)), and system resource usage (Resource Usage Function (RUF)). Additionally, the Query Plan Quality Index (QPQI) improved by approximately 10%, and the convergence rate was up to 30% faster than competitors. These findings indicate that the Orchard Algorithm is an effective option for optimizing complex queries in distributed database systems.