

П. А. Якушевич
(ГрГУ им. Я. Купалы, Гродно)

СРАВНЕНИЕ АЛГОРИТМОВ ПОИСКА ПОДСТРОК

Реляционная база данных – база данных, основанная на реляционной модели данных. Реляционная модель данных (РМД) – логическая модель данных, которая включает следующие компоненты: Структурный аспект – данные в базе данных представляют собой набор отношений.

- Аспект целостности – отношения отвечают определенным условиям целостности. РМД поддерживает декларативные ограничения целостности уровня домена, уровня отношения и уровня базы данных.

- Аспект обработки – РМД поддерживает операторы манипулирования отношениями.

Постоянно растущий объем информации, накопленной человечеством, предъявляет повышенные требования к поиску среди этой информации. Поиск подстроки в строке – одна из простейших и при этом важнейших задач поиска. Кроме того, зачастую точные критерии поиска неизвестны или не могут быть заданы.

Были рассмотрены и проанализированы различные алгоритмы поиска подстрок, в том числе алгоритм Рабина-Карпа, алгоритм Бойера-Мура, двоичный алгоритм, алгоритм Ахо-Корасик, алгоритм Кнута-Морриса-Пратта. Была спроектирована система для оценки производительности алгоритмов поиска подстрок и проведено нагрузочное и стресс-тестирование.

Проект предлагает пользователю заранее подготовленные данные для тестов, но оставляет возможность ввода собственных данных и выбора алгоритмов, которые будут протестированы.

В качестве платформы разработки была выбрана Java. Проект реализован как web-приложение. Клиентская часть выполнена с помощью технологий JSP и AJAX. В качестве персистентного хранилища данных использована реляционная база данных MySQL.