

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК [002.6:004.421]:91

Жишко  
Юрий Петрович

**Алгоритмы для оптимизации географического поиска**

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 04 «Математическое моделирование, численные  
методы и комплексы программ»

Научный руководитель  
Стержанов Максим Валерьевич  
доцент, кандидат технических наук

Минск, 2015

## ВВЕДЕНИЕ

Быстрое развитие информационных технологий выдвигает все больше и больше требований к аппаратному и программному обеспечению современных систем. Так как основной единицей на рынке информационных технологий и услуг является информация, то одним из ключевых факторов развития данного сегмента является ее быстрый поиск.

В последнее время активно развивается рынок мобильных услуг. Люди все больше и больше прибегают к использованию мобильных устройств. Денежный оборот в этой сфере составляет сотни миллиардов долларов в год. Порядка 30% процентов этого оборота составляет мобильная контекстная реклама (англ. *Mobile Advertisement*). Преимуществом мобильных устройств, для рекламодателей, является то, что можно отслеживать перемещение конкретного пользователя используя не только данные специфических приложений, загружаемых пользователем, историю посещения интернет сайтов, но и другие источники: координаты GPS и сотовая триангуляция. На основе статистики посещения сайтов, географического положения и истории перемещения пользователя мобильного устройства поставляется контекстная реклама, наиболее релевантная для данного пользователя, происходит информирование о новинках на рынке товаров и услуг.

Одним из ключевых механизмов и методов подбора подобного рода рекламы является геотаргетинг (англ. *geo targeting*) — в веб-разработке и интернет-маркетинге, метод выдачи посетителю содержимого, соответствующего его географическому положению. Приложения и веб-сервисы использующие геотаргетинг должны обладать возможностью предоставлять метаданные на основе географического положения пользователя (дом, магазин, адрес, штат, страна, парковка и т. д.) за короткий промежуток времени и передавать эти метаданные другим сервисам.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

При использовании метода геотаргетинга приходится решать ряд задач по поиску объектов на географической местности. В частности такие как: *по заданной широте и долготе определить объект, которому принадлежит эта точка, определение близлежащих объектов в заданном радиусе поиска, находится ли данный объект внутри другого объекта или какую площадь он занимает.*

В общем случае подобные задачи сводятся к классу задач классической алгебры и пространственной геометрии с использованием дискретных машинных алгоритмов. В частности, задача принадлежности точки с указанными координатами многоугольнику произвольного размера и формы (полигону), задача поиска ближайших точек в окружности заданного радиуса, определение площади перекрытия двух полигонов.

При использовании стандартных математических алгоритмов перебора, решение данных задач потребует значительное количество времени на обработку данных, что делает их непригодными в использовании в режиме реального времени. По этой причине оптимизация подобных вычислений играет ключевую роль в геотаргетинге.

На данный момент существует множество оптимизаций решения подобных задач, позволяющих избежать прямого перебора. С использованием различных индексных структур, к примеру, бинарных деревьев и их разновидностей (B+-дерево, BST – дерево, R – дерево), удастся значительно снизить время поиска. Подобные структуры активно используются в различных базах данных и поисковых системах для реализации пространственного поиска.

Но, при детальном изучении, было выявлено, что, в конечном счете, они не гарантируют максимальную производительность. Идея использовать координатную сетку для индексации пространственных объектов зарекомендовала себя как лучшее решение в плане производительности.

*Целью* данной работы является создание и разработка алгоритмов оптимизации географического поиска с использованием координатной сетки. В связи с этим были поставлены следующие *задачи*:

- I. Проанализировать существующие алгоритмы поиска в двумерном евклидовом пространстве;
- II. Определить возможности применения данных алгоритмов на географической местности с использованием географических координат;
- III. Предложить версию алгоритмов с использованием координатной сетки решения следующих задач:
  - a. Принадлежности точки с заданными двумерными координатами (широта и долгота) многоугольнику (полигону);
  - b. Определение ближайших точек (географических объектов) в окрестности по заданному расстоянию;
  - c. Определение зоны (площади) перекрытия двух многоугольников;
- IV. Разработать метод разбиения географической местности на координатную сетку;
- V. Выполнить программную реализацию алгоритма разбиения на координатную сетку;
- VI. Выполнить программную реализацию алгоритмов с использованием координатной сетки для вышеперечисленных задач;
- VII. Дать оценку полученным результатам.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Работа состоит из семи глав.

Первая глава посвящена методам прямого поиска и перебора для следующего круга математических задач:

**Задача 1:** Из множества многоугольников определить те, которые содержат точку с указанными координатами.

**Задача 2:** Из множества точек на плоскости (географической местности) определить те, которые лежат в окрестностях заданных точек на определенном расстоянии.

**Задача 3:** Определить площадь перекрытия двух многоугольников произвольной формы.

Во второй главе приведены методы и алгоритмы оптимизации решения поставленных задач с использованием  $R$ -деревьев, а также произведенные вычисления роста производительности.

В третьей и четвертой главах рассмотрено использование координатных систем  $UTM$  и  $MGRS$  для решения вышеприведенных задач.

В четвертой главе представлена разработанная в ходе исследования координатная система  $JGRS$  для оптимизации географического поиска.

В пятой и шестой главах подробно описаны алгоритмы построения пространственной сетки и пространственного индекса в  $JGRS$ , разобрано решение задач 1, 2 и 3 в координатной системе  $JGRS$ .

В завершающей седьмой главе приведена сравнительная характеристика производительности  $JGRS$  и  $MGRS$ ,  $UTM$ , а также дано сравнение с методами на основе  $R$ -деревьев для решения однотипных задач географического поиска.

## ЗАКЛЮЧЕНИЕ

В результате выполнения работы были проанализированы основные проблемы, с которыми приходится сталкиваться при решении задач географического поиска. Также были рассмотрены существующие методы и алгоритмы для решения типичных математических задач, к которым сводятся основные задачи геотаргетинга. Среди них можно выделить методы простого перебора и использование структур, обеспечивающих логарифмическое время поиска координат в пространстве и на плоскости.

В работе были предложены новые методы и алгоритмы географического поиска с использованием координатной сетки и сеточных указателей. В результате анализа существующих координатных систем, используемых для локализации объектов на географической местности, была разработана принципиально новая координатная система *JGRS*, которая обладает рядом преимуществ, по сравнению с уже существующими координатными системами *MGRS* и *UTM*.

В рамках исследования была выполнена программная реализация алгоритма разбиения географической местности на координатную сетку в координатной системе *JGRS*. Разработаны методы и алгоритмы решения задач геотаргетинга в координатной системе *JGRS*. Была выполнена программная реализация этих методов и алгоритмов на нескольких языках программирования. Разработаны программные модули, которые были интегрированы в различные современные базы данных и поисковые системы.

В завершающей части были рассмотрены преимущества и недостатки, разработанных алгоритмов для оптимизации географического поиска при использовании их в различных системах. Проведен сравнительный анализ с уже существующими алгоритмами и методами. Даны рекомендации по возможностям их использования в вычислительных системах.

Использование координатной системы *JGRS* позволяет получить максимальную производительность при решении задач географического поиска. Разработанные алгоритмы индексации и поиска легко распараллеливаются на большое число вычислительных машин. Однако на построение пространственного индекса *JGRS* требуются значительные аппаратные ресурсы.

В качестве продолжения исследования можно реализовать следующие идеи:

- Расширить применение географической системы *JGRS* на пространства с размерностью больше двух (широта, долгота и высота);
- Использовать *JGRS* в GPS приемниках;
- Создать структуру поиска по сеточным ссылкам различной точности, требующую только один проход по всему дереву поиска без распараллеливания.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

На момент написания диссертации все основные положения и разработки, представленные в работе, находились в состоянии проверки в американском патентном бюро без права на публикацию.

Библиотека БГУИР