

Заключение. Вышеописанный алгоритм был реализован на языке программирования Scala и протестирован на больших объемах данных (в качестве исходных данных использовалась база логов известного веб-сервиса Last.fm). Примеры рекомендаций можно увидеть в презентации.

ЛИТЕРАТУРА

1. Рекомендательные системы [ИСТОЧНИК URL: https://ru.wikipedia.org/wiki/Рекомендательная_система].
2. Коллаборативная фильтрация [ИСТОЧНИК URL: https://ru.wikipedia.org/Коллаборативная_фильтрация].
3. Коллаборативная фильтрация [ИСТОЧНИК URL: <https://habrahabr.ru/post/150399/>].

ВИЗУАЛИЗАЦИЯ ГРАФОВ С ПОМОЩЬЮ СИЛОВЫХ АЛГОРИТМОВ

М.В. Демешко, А.Ю. Дёмин
(г. Томск, Томский политехнический университет)
e-mail: demeshkomaria@gmail.com

GRAPH VISUALIZATION USING FORCE ALGORITHMS

M.V. Demeshko, A.Y. Demin
(Tomsk, Tomsk Polytechnic University)

This article describes theory and realization of force-based graph visualization. There is an overview of aesthetic criteria of visualization. Article contains examples of visualization using force-based algorithms and describes the need to use different metrics.

Keywords: graph, force-directed, graph-visualization, graph drawing, graph metrics.

Задача визуализации графов встречается в таких областях как картография, анализ сетей (в том числе социальных), анализ программных зависимостей, анализ цитируемости публикаций и многих других. Не смотря на то, что в каждой из этих областей используются выборки данных разного объема, а также преследуются различные конечные цели, существует несколько эстетических критериев, которые позволяют оценить качество визуализации графа для любой задачи.

К этим критериям можно отнести отсутствие наложения вершин. В зависимости от задачи, вершины могут плотно примыкать друг к другу или отстоять на некоторое расстояние, но их наложение в любом случае затрудняет чтение графа. Следующий критерий – это минимизация числа пересечений ребер. Для его соблюдения в некоторых алгоритмах используются дуговые или изломанные ребра.

Для некоторых задач принципиальна укладка графа на минимально возможную площадь, то есть длины ребер не должны быть избыточными, а сам граф не должен выходить за область визуализации. Кроме того, одним из самых важных критериев является отражение топологии графа.

Соблюдения перечисленных критериев можно достичь использованием силовых алгоритмов визуализации. Они относятся к классу алгоритмов, использующих в своей работе физические аналогии.

Силовые алгоритмы описывают функцию, которая определяет идеальные параметры, к которым должен стремиться граф. Соответственно, при каждой итерации система прибли-

жается к заданным параметрам, а после заключительной итерации считается находящейся в равновесии.

На каждой итерации необходимо вычислять следующую позицию вершин. Для решения этой задачи используется метод численного интегрирования Верле [1]. С его помощью следующая позиция может быть вычислена на основе текущей и прошлой.

Силовая функция оперирует понятием заряда. Каждой вершине присваивается отрицательный заряд, что позволяет им отталкиваться друг от друга, критерий отсутствия наложений вершин соблюдается. Но при таком подходе возможна ситуация, когда компоненты связности будут отталкиваться друг от друга и удаляться на избыточное расстояние, что может сделать граф слишком разреженным или даже вынести часть графа за пределы области визуализации.

Для того чтобы соблюдался критерий использования минимально возможной площади, необходимо ввести дополнительную силовую функцию, которая будет присваивать положительное значение заряда центру области визуализации [2]. Таким образом, вершины будут отталкиваться друг от друга, но притягиваться к центру.

Использование силовых алгоритмов позволяет настраивать множество параметров. Чтобы эффективно их использовать, необходимо осуществлять предварительный анализ графа. Он заключается в расчете некоторых метрик, которые позволяют установить конфигурацию графа.

Самой простой метрикой является отношение числа ребер к числу вершин. Чем оно больше, тем большая длина ребер должна использоваться. Это позволяет визуализировать граф с достаточной степенью разреженности.

Следующая метрика – число компонент связности. Чем больше компонент связности, тем больше должен быть гравитационный параметр, чтобы компоненты не удалялись за границы области воспроизведения, а также располагались максимально компактно.

Метрика бинарного дерева позволяет установить, представляет ли из себя граф эту структуру. Для ее расчета достаточно установить количество компонент связности, а затем, для каждой из них, проверить соотношение числа ребер и вершин. Если число ребер на единицу меньше, чем число вершин, то эта компонента представляет собой бинарное дерево. Эту структуру следует визуализировать послойно, опуская потомков ниже родителей.

Для графов, отражающих сетевые структуры, рассчитывается метрика компонент сильной связности [3]. Смысл ее использования заключается в том, чтобы наглядно показать плотность связей между некоторыми компонентами графа. Фактически, метрика позволяет визуализировать кластеры сетевых единиц.

На рисунках представлены примеры визуализации графов с помощью силовых алгоритмов, использующих метрики.

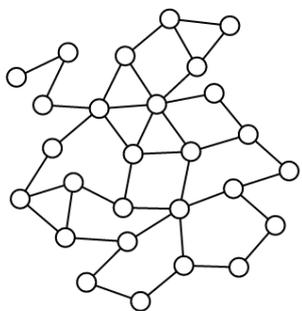


Рис. 1. Граф, визуализированный после расчета отношения числа ребер к числу вершин

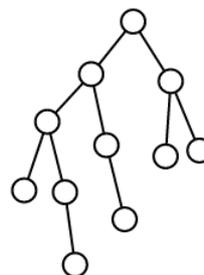


Рис. 2. Граф, визуализированный после расчета метрики бинарного дерева

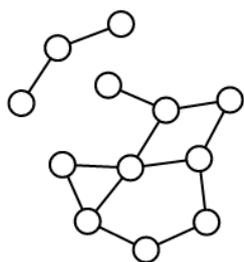


Рис. 3. Граф, визуализированный после нахождения числа компонент связности

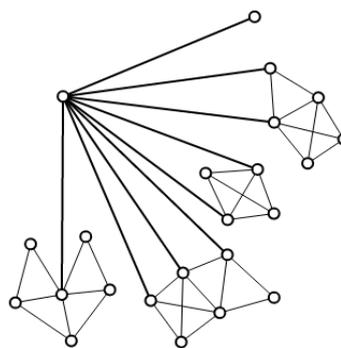


Рис. 4. Граф, визуализированный после анализа плотности связей

ЛИТЕРАТУРА

1. Jakobsen T. Advanced Character Physics. – Ю Interactive, Copenhagen, 2003.
2. Пупырев С.Н., Тихонов А.В. Визуализация динамических графов для анализа сложных сетей. – Моделирование и анализ информационных систем, 2010, №1.
3. Коломейченко М.И., Чеповский А.М. Визуализация и анализ графов больших размеров. – Бизнес-информатика, 2014, №4(30).

СЕРВЕРНОЕ ПРИЛОЖЕНИЕ ДЛЯ СБОРА И ХРАНЕНИЯ ФАЙЛОВ С МЕТЕОДААННЫМИ

Е.Л. Ерофеева, В.Н. Попов
 (г. Томск, Томский политехнический университет)
 e-mail: elizaveta.erofeeva@gmail.com

SERVER APPLICATION FOR GATHERING AND STORING OF FILES WITH METEOROLOGICAL DATA

Erofeeva E.L., Popov V.N.
 (Tomsk, Tomsk Polytechnic University)
 e-mail: elizaveta.erofeeva@gmail.com

This article describes the server application for automated gathering and storing of files with unstructured hydrometeorological data from sites on the Internet that provide such information. The software makes it easier to receive weather reports, including encoded in the KN-01, as well as to work with them.

Keywords: hydrometeorological data, server application, code operational data, database.

Наблюдение за климатом нашей планеты ведется на протяжении нескольких веков, по причине того, что своевременная и качественная гидрометеорологическая информация играет важную роль в обеспечении защиты жизни и имущества граждан, экономике регионов. Информацию о погодных условиях предоставляют гидрометеостанции, предоставляя количественные данные о текущем состоянии атмосферы, например, температура воздуха, скорость ветра, атмосферное давление, наличие осадков и т.д. Проблеме извлечения, структурированию, хранению и обработке файлов с метеоданными посвящена данная работа.

Наиболее распространённым кодом, содержащим гидрометеорологическую информацию, в котором гидрометеостанции хранят полученные данные в международном формате,