

Полученные результаты моделирования показали, что алгоритм GIRS в сочетании с алгоритмами восстановления позволяет повысить точность классификации примеров при наличии шума в данных.

Литература

1. *Vagin V.N., Ereemeev A.P.* Methods and Tools for Modelling Reasoning in Diagnostic Systems. ISEIS 2009, Proceedings of the 11th International Conference on Enterprise Information Systems, vol. AIDSS, Milan, Italy, 2009. pp 271-276.
2. *Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В.* Достоверный и правдоподобный вывод в интеллектуальных системах / Под ред. В.Н. Вагина, Д.А. Поспелова. – 2-е издание дополненное и исправленное. – М.: ФИЗМАТЛИТ, 2008. – 712 с.
3. *Pawlak Z.* Rough sets and intelligent data analysis / Information Sciences, Elsevier Science. 2002. Vol. 147. iss. 1. pp. 1-12.
4. *Vagin V.N., Kulikov A.V., Fomina M.V.* The Development of the Generalization Algorithm based on the Rough Set Theory. International Journal INFORMATION Theories & Applications (ITHEA), 2006. Vol.13 № 3, pp. 255-262.
5. *Vagin V., Fomina M.* Problem of Knowledge Discovery in Noisy Databases. In Int. J. Mach. Learn. & Cyber. 2011. Vol.2, num. 3, Springer Ferlag, Berlin, pp. 135-145.
6. *Merz C, Murphy P.* 1998. UCI Repository of Machine Learning Datasets. Information and Computer Science University of California, Irvine, CA 92697-3425 <http://archive.ics.uci.edu/ml/>

Argumentation in inductive notion formation

*Vadim Nikolaevich Vagin, Doctor of Engineering, Professor, Chair of applied mathematics
Marina Vladimirovna Fomina, Associate professor, Candidate of Technical Sciences, Chair of computer facilities
National Research University Moscow Power Institute*

The methods of processing incomplete and inconsistent information in such subsystems of intelligent decision support systems as a decision searching subsystem of acquiring and storing knowledge are considered. In the decision searching subsystem, the main stress is made on the application of argumentation with justification degrees. In the acquiring and completion subsystem, the generalization problem in conditions of inconsistent and incomplete data is solved. For data generalization, the algorithm based on the rough set theory is developed. Research of noise influence on the work of the suggested generalization algorithm is carried out. The results of program modelling are presented.

Keywords: argumentation, defeasible reasoning, degrees of justification, rough sets, notion formation, generalization, noisy data.

УДК 519.673

СХЕМА ДЕЙСТВУЮЩЕГО ПРОТОТИПА МУЛЬТИАГЕНТНОЙ МОДЕЛИ РАСПРОСТРАНЕНИЯ ЭПИДЕМИЙ

Светлана Юрьевна Лапшина, научный сотрудник

Тел.: 8 916 518 65 80, e-mail: lapshina@jscc.ru

Федеральное государственное бюджетное учреждение науки Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН)

<http://www.jscc.ru>

В работе описана схема действующего прототипа мультиагентной модели распространения эпидемий с параллельным алгоритмом многократной маркировки перколяционных кластеров программно-аппаратного комплекса BIOCLUST МСЦ РАН.

Ключевые слова: перколяционный кластер, алгоритм многократной маркировки кластеров, имитационное моделирование.

В Межведомственном суперкомпьютерном центре Российской академии наук на основе теории формирования и роста перколяционных кластеров [1] как часть программно-аппаратного комплекса BIOCLUST [2-4] был разработан и опробован действующий прототип мультиагентной модели распространения эпидемий с параллельным алгоритмом многократной маркировки кластеров на языке Си и с использованием библиотеки MPI.

Прототип был построен для исследования возникновения и распространения широкомасштабных эпидемий. Для формирования анализируемой решетки использовались демографические данные о численности населения городов всего мира и некоторые предположения об интенсивности и характере взаимодействия между людьми. Для получения сведений о численности населения городов был разработан алгоритм обработки и анализа карт плотности населения (реализация на языке java). Сами карты были позаимствованы у Google при помощи Google Maps API, а для удобства восприятия результатов имитационных экспериментов — разработаны соответствующие



средства визуализации.

На рис. 1 представлена схема программной реализации действующего прототипа мультиагентной модели распространения эпидемий.

С помощью алгоритма сбора информации о численности населения обработчик карт MapManager конвертирует карты Google Maps в их отображение цилиндрической проекцией Меркатора с дальнейшей генерацией перколяционной решетки и «привязкой» географических координат к городам, как основным концентраторам плотности населения. Карта вводится в функцию посредством множества png-файлов, которые представляют разбику карты на квадранты. Получены данные о 56 976 городах мира с общей численностью населения 6 114 628 510 человек и

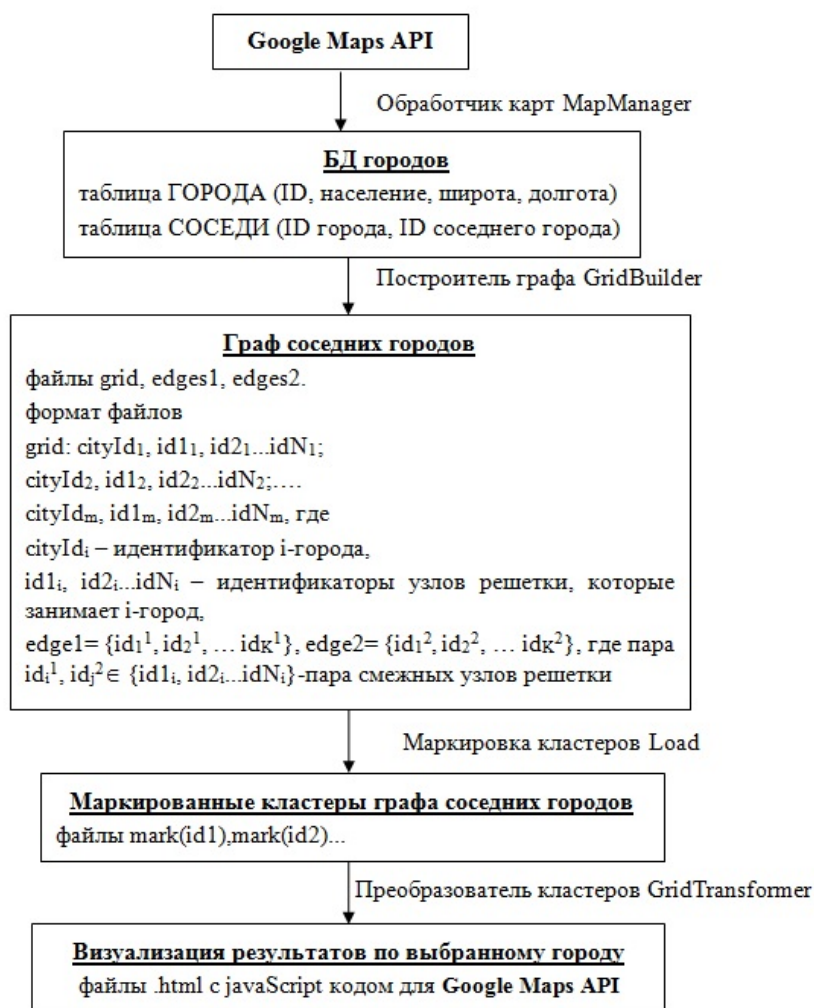


Рис. 1. Схема действующего прототипа

они сохранены в специализированной таблице городов (номер города, численность населения, широта, долгота). На этом же этапе производится создание и заполнение таблицы городов-соседей (идентификатор города — идентификатор соседа), которая состоит примерно из 50 млн. записей.

Алгоритм построения графа взаимодействия представителей популяции GridBuilder создает исходную решетку и сохраняет ее в трех файлах специального формата. Множество узлов, приходящихся на определенный город, сохраняется в файле *grid*, пары ребер-связей между узлами - в файлы *edges1* и *edges2*. В ходе эксперимента была создана исходная решетка и сохранена в трех файлах размеров: решетка *grid*, размер 475 Мб, файлы *edges1* и *edges2* – ребра, общим размером 429 Мб. Для каждого значения входного параметра p - вероятности заражения при контакте с инфицированным – для исходной решетки в оперативной памяти формируется решетка анализируемая.

Граф соседних городов разбивается на связанные подмножества алгоритмом Хошена-Копельмана (маркировка кластеров Load). Для каждой анализируемой решетки запускается алгоритм многократной маркировки перколяционных кластеров. В качестве результата формируется массив соответствующих кластерных меток (файлы *marks_x*, где $x=0..99$). Размер каждого сохраненного массива составляет примерно 24 Мб.

Программа Load запускалась в Межведомственном суперкомпьютерном центре на суперкомпьютере МВС-100К с общим количеством процессоров 11 680 на 48-264 процессорах. Среднее время выполнения программы с входным параметром p от 0,01 до 1 с шагом в 0,01 при постоянных значениях $t = 1, 3, 5, 10, 15, 20, 25, 30$ дней составило около 10 минут для каждого значения t .

На суперкомпьютере МВС-10П с общим количеством процессоров 28 704 среднее время выполнения этой программы составило около 5 минут.

Оптимальное количество процессоров, на которых имеет смысл вести модельный расчет программой Load для каждого значения t , равно 128. Это объясняется большим размером анализируемых решеток. С одной стороны их нужно разбить на достаточное количество частей, чтобы каждый процессор мог загрузить их в свою оперативную память и провести обработку, а с другой стороны этих частей не должно быть слишком много, т.к. обмен приграничными метками между процессорами достаточно долгий и ресурсоемкий процесс.

Однако программа Load допускает и большую масштабируемость, если в параллельном режиме запустить расчет для разных временных интервалов развития эпидемии и специальным образом организовать массив соответствующих кластерных меток. В Межведомственном суперкомпьютерном центре подобная имитация запускалась на суперкомпьютере МВС-10П при переменных значениях $t = 1-30$ дней с шагом в 1 день на 3840 процессорах. Среднее время выполнения этой программы с входным параметром p от 0,01 до 1 с шагом в 0,01 составило около 8 минут.

Основная особенность программы Load - работа с целочисленными данными большого размера с возможными конфликтами по использованию общей памяти в совокупности с хорошей масштабируемостью. Это определило целесообразность ее применения при тестировании новых видов многоядерных процессоров и архитектуры суперкомпьютеров.

После работы программы Load преобразователь кластеров GridTransformer проводит визуализацию полученных результатов при помощи *Google Maps Api*. Выбирается город – очаг заражения, задается количество зараженных узлов в этом городе, загружается вычисленный ранее массив кластерных меток для данных p и t . Определяются метки зараженных узлов данного города. Производится последовательный обход меток всех узлов из загруженного массива. Если метка являлась “потенциально зараженной”, то город, в который входил узел с данной меткой объявляется потенциально зараженным. Таким образом, вычисляется список потенциально зараженных городов. Этот список сохраняется в виде *javascript* команд *Google Maps API* в *html* файл. В зависимости от численности населения зараженный город помечается красным кругом различных размеров. Для просмотра результатов *html* файл запускается в браузере.

Автор считает, что в данной работе новыми являются следующие положения и результаты:

- Показана схема действующего прототипа мультиагентной модели распространения эпидемий на основе теории перколяции которая является существенно новым подходом как в имитационном моделировании, так и в направлении моделирования процессов распространения заболеваний;

- Предложенная программная реализация алгоритма является оригинальной в области распараллеливания алгоритма многократной маркировки перколяционных кластеров и позволяет эффективно проводить имитационные эксперименты в многопроцессорной вычислительной среде.

Литература

1. *Тарасевич Ю.Ю.* Перколяция: теория, приложения, алгоритмы: учебное пособие. – М.: УРСС, 2002. – 64 с.

2. *Лапишина С.Ю.* Высокопроизводительные вычисления в практике моделирования роста перколяционных кластеров // Программные продукты и системы. 2011. № 4 (96). С. 75-79.

3. *Лапишина С.Ю.* Мультиагентная модель распространения массовых эпидемий на основе теории роста перколяционных кластеров // Информационные технологии в науке, социологии и бизнесе: материалы XL-ой Международной конференции и X Международной конференции молодых ученых IT+SE'12. Осенняя сессия. Украина. Крым. Ялта-Гурзуф. – приложение к журналу «Открытое образование». С 47-49.

4. *Лапишина С.Ю.* Параллельный алгоритм многократной маркировки кластеров для задач теории перколяции // Информационные технологии в науке, социологии и бизнесе: материалы XLI-ой Международной конференции и XI Международной конференции молодых ученых IT+SE'13, Осенняя сессия, Украина, Крым, Ялта-Гурзуф. – приложение к журналу «Открытое образование». С. 45-47.

Scheme of the prototype Mass epidemics' multi-agent model expansion

Svetlana Yurevna Lapshina, Research associate

Scientific Federal State Budgetary Institution Interdepartmental Supercomputer Center of Russian Academy of Sciences

This paper describes a prototype implementation of mass epidemics multi-agent model with parallel multiprocessing cluster multiple labeling technique used in the software and hardware complex BIO-CLUST JSCC RAS.

Key words: percolation's cluster, cluster multiple labeling technique, simulation.

УДК 658.314.7:330.115

ОЦЕНКА ПОКАЗАТЕЛЕЙ СЛОЖНОСТИ УПРАВЛЕНИЯ ДВИЖЕНИЕМ

Олег Викторович Белый, д-р. техн.наук, проф., директор

Тел.: 812 321 97 42, e-mail: belyi@iptran.ru

Анвер Касимович Еналеев, канд. техн.наук, ведущий научный сотрудник

Тел.: 812 321 97 42, e-mail: belyi@iptran.ru

Владимир Викторович Цыганов, д-р.техн.наук, проф., зав. отделом

Тел.: 812 321 97 42, e-mail: belyi@iptran.ru

Институт проблем транспорта РАН им. Н.С. Соломенко

<http://wikimapia.org>

Предложены процедуры оценки сложности управления движением для решения задачи определения границ полигонов управления движением транспорта на сети (на примере желез-