

ВЫВОДЫ

Расширение номенклатуры средств описания формализуемых в моделях представления экспертных проблемных знаний, за счет предлогов выделенных классов, позволило расширить применение этих моделей для формального описания простых и сложных предложений русского языка со свободным порядком слов в предложении. Приведенные в данной статье, а также в работе [1] результаты позволяют формализовать полученные от экспертов проблемные знания, что делает их пригодными для дальнейшей машинной обработки с целью автоматизации процесса приобретения знаний из экспертов в предметных областях.

ПЕРЕЧЕНЬ ССЫЛОК

1. Сердюк С. М. Моделі представлення експертних проблемних знань. // Вісник ДУ «Львівська політехніка». Радіоелектроніка та телекомунікації. – 2000. – № 387. – С. 387–391.

2. Гужва Ф. К. Морфология русского языка: пособие для преподавателя. – К.: Рад. шк., 1987. – 223 с.

Надійшла 17.09.05
Після доробки 19.05.06

Проведено дослідження семантико-синтаксичної структури речень російської мови з метою удосконалення моделей представлення експертних проблемних знань, що дозволяють автоматизувати процес здобуття знань про більш вузьку область – із експерта по предметним знанням. Процес формалізації проблемних знань поданий двома взаємозалежними моделями: прикладною моделлю представлення знань експертів і моделлю представлення цих знань за допомогою спрямованих графів.

The analysis of Russian language sentence for semantic and syntactic structure is undertaken targeting to enhance models of expert objective knowledge. The models are engaged to automate the retrieval process to extract knowledge from experts specifically objected to a certain knowledge area. The knowledge subject's formalisation process is driven by two interconnected models: one is the application expert knowledge retrieval model and the model for that knowledge itself, presented by directional graphs.

УДК 681.324.01

А. Я. Склярів, І. А. Макрушан

МОДЕЛИРОВАНИЕ ТРАНСПОРТНЫХ КОММУНИКАЦИЙ КОРПОРАТИВНЫХ КОМПЬЮТЕРНЫХ СЕТЕЙ: РАСПРЕДЕЛЕНИЕ ТРАНЗАКЦИЙ ПО МАРШРУТАМ

В данной статье рассмотрены вопросы моделирования транспортных коммуникаций корпоративных компьютерных сетей. Предлагается два подхода решения задачи: построение модели распределения потоков данных на основе методов максимизации энтропии, а также использование гравитационной модели распределения транзакций. В соответствии с предлагаемыми подходами представлены гравитационная и энтропийная модели распределения потоков данных. Показана практическая значимость предложенных подходов.

ВВЕДЕНИЕ

При проектировании систем управления распределенными корпоративными информационными технологиями, для того чтобы оценить интенсивность потоков данных (порождаемых задачами управления) между различными хостами сети, а также для моделирования распределения потоков данных по участкам транспортной сети, обычно используется некоторая математическая модель [1]. Эта модель, как правило, должна состоять из четырех подмоделей, которые связаны с генерацией трафика, распределением запросов по комму-

никациям, расщеплением типов коммуникаций (маршрутов) и размещением общих сетевых ресурсов сети. В этой статье вопросы, связанные с назначениями транзакций, не рассматриваются; здесь предполагается, что некоторые оценки транзакций в конечных пунктах заранее известны. Некоторые результаты статистического моделирования позволяют делать определенные выводы по размещению информационных ресурсов, но нас в основном будут интересовать распределение транзакций и расщепление типов коммуникаций [2]. Для решения этих задач будут применяться методы максимизации энтропии [3].

Цель статьи заключается в том, чтобы показать главное преимущество методов максимизации энтропии в области транспортных задач, которое состоит в том, что они помогают строить модели, описывающие довольно сложные явления, а также в том, что они допускают прямую интерпретацию полученных уравнений.

В связи с этим возникают следующие задачи:

- 1) подробно рассмотреть вопросы, связанные с гравитационной моделью распределения транзакций [2];

2) показать, как необходимо строить модели распределения потоков данных на основе методов максимизации энтропии;

3) привести ряд выводов, существенных для интерпретации полученных результатов моделирования.

1 ГРАВИТАЦИОННАЯ МОДЕЛЬ РАСПРЕДЕЛЕНИЯ ПОТОКОВ ДАННЫХ

Задача, которая ставится перед любой моделью распределения транзакций, заключается в оценке числа транзакций между каждой парой «точек» [4, 5]. Под «точкой» в данном случае понимается некоторый общий сетевой ресурс компьютерной сети. Предполагается, что рассматриваемая корпоративная сеть некоторым подходящим образом разбита на подсети [6].

Для простоты рассмотрим в качестве примера формирование и передачу запросов к файл-серверам корпоративной сети. Пусть T_{ij} – число транзакций (обращений к сетевому ресурсу), а C_{ij} – затраты на передачу из сети i в сеть j ; пусть Q_i – полное число отправленных запросов из сети i к файл-серверам других сетей, а D_j – полное число принятых запросов файл-сервером сети j . Модель распределения транзакций оценивает T_{ij} как функцию от Q_i , D_j , C_{ij} . Эти величины сами также могут зависеть от других независимых переменных.

Наиболее простая модель такого типа – это так называемая гравитационная модель, разработанная по аналогии с ньютоновским законом, связывающим силу притяжения F_{ij} между двумя массами m_i и m_j , расположенными на расстоянии d_{ij} друг от друга [2]

$$F_{ij} = \gamma \frac{m_i m_j}{d_{ij}^2}, \quad (1)$$

где γ – некоторая константа. По аналогии транспортная гравитационная модель может быть представлена в виде:

$$T_{ij} = k \frac{Q_i D_j}{c_{ij}^2}, \quad (2)$$

где k – некоторая константа, а затраты на передачу c_{ij} выступают в качестве «расстояния». Согласно уравнению (2) величина T_{ij} пропорциональна Q_i и D_j и обратно пропорциональна квадрату «расстояния» между ними. Но у этого уравнения имеется, по меньшей мере, один очевидный недостаток: если удвоить заданные значения Q_i и D_j , то число транзакций между этими сетями в соответствии с (2) учетверится, а естественно ожидать, что оно лишь удвоится. Или более точно: величины T_{ij} всегда должны удовлетворять следующим

ограничениям (число отправленных запросов из всех сетей в сеть j должно быть равно количеству прибывших запросов в сеть j):

$$\sum_i T_{ij} = D_j, \quad (3)$$

$$\sum_j T_{ij} = Q_i, \quad (4)$$

а уравнение (1) этого не обеспечивает. Это означает, что суммы по строкам и столбцам матрицы транзакций должны совпадать с числом запросов из каждой сети, и с числом полученных запросов в каждой сети соответственно. Этим ограничениям можно удовлетворить, если ввести наборы констант A_i и B_j , связанные соответственно с сетями источниками и сетями назначения. Иногда их называют балансирующими множителями. Кроме того, нет оснований считать, что расстояние играет в уравнении (2) такую же роль, что и в ньютоновской физике, поэтому используем более общую форму функции расстояния $f(c_{ij})$. Модифицированная гравитационная модель, при этом, будет иметь следующий вид:

$$T_{ij} = A_i B_j Q_i D_j f(c_{ij}), \quad (5)$$

где

$$A_i = \left[\sum_j B_j D_j f(c_{ij}) \right]^{-1}, \quad (6)$$

$$B_j = \left[\sum_i A_i Q_i f(c_{ij}) \right]^{-1}. \quad (7)$$

Уравнения для A_i и B_j решаются итерационными методами, и можно утверждать, что они гарантируют, что величины T_{ij} из уравнения (5) удовлетворяют ограничениям (3) и (4). Отметим также, что величины c_{ij} в такой модели могут служить общей мерой «сопротивления» передач данных между сетями i и j , в качестве которой могут выступать: географическое расстояние, время передач, затраты на передачи или, что часто бывает более эффективным, взвешенная комбинация этих факторов, которую иногда называют «обобщенными затратами».

Рассмотрим другой способ получения гравитационной модели транспортной системы с помощью максимизации энтропии [3], используя вначале другую физическую аналогию. Для этого введем дополнительное к (3) и (4) ограничение на T_{ij} . Это ограничение имеет вид:

$$\sum_i \sum_j T_{ij} c_{ij} = C. \quad (8)$$

Тогда наиболее вероятным распределением транзакций будет матрица $\{T_{ij}\}$, максимизирующая энтропию

$$\ln W(\{T_{ij}\}) = \ln T! - \sum_i \sum_j \ln T_{ij}!, \quad (9)$$

где T – полное число транзакций при ограничениях (3), (4), (8). Если учесть, что $W(\{T_{ij}\})$ – полное число состояний системы, соответствующих распределению $\{T_{ij}\}$, то число всех состояний системы можно оценить по формуле [3]

$$w = \sum W(\{T_{ij}\}), \quad (10)$$

где суммирование проводится по всем T_{ij} , удовлетворяющим (3), (4) и (8). При этом оказывается, что максимальное значение $W(\{T_{ij}\})$ настолько доминирует над всеми остальными членами суммы, что распределение $\{T_{ij}\}$, соответствующее максимуму энтропии, является существенно наиболее вероятным распределением.

Для получения набора T_{ij} , максимизирующего $\ln W(\{T_{ij}\})$ из уравнения (9) при ограничениях (3), (4) и (8), следует максимизировать лагранжиан L , равный

$$L = \ln W + \sum_i \lambda_i^{(1)} \left(Q_i - \sum_j T_{ij} \right) + \sum_j \lambda_j^{(2)} \left(D_j - \sum_i T_{ij} \right) + \beta \left(C - \sum_i \sum_j T_{ij} c_{ij} \right), \quad (11)$$

где $\lambda_i^{(1)}$, $\lambda_j^{(2)}$ и β – множители Лагранжа. Поскольку предполагается, что количество транзакций T_{ij} достаточно велико, то можно воспользоваться формулой Стирлинга [3] ($\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij}$ – эта замена справедлива для больших значений T_{ij}), согласно которой из (9) получим

$$\ln W = - \sum_i \sum_j T_{ij} \ln T_{ij}.$$

Значения T_{ij} , которые доставляют максимум L и, следовательно, являются наиболее вероятным распределением транзакций, представляют собой решение системы уравнений $\frac{\partial L}{\partial T_{ij}} = 0$ совместно с ограничениями (3), (4) и (8).

Будем иметь

$$\frac{\partial L}{\partial T_{ij}} = \ln T_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}.$$

Это выражение будет равно нулю, когда

$$T_{ij} = \exp(-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}). \quad (12)$$

Подставляя (12) в (3) и (4), получим выражения, используя которые можно определить $\lambda_i^{(1)}$ и $\lambda_j^{(2)}$:

$$\exp(-\lambda_i^{(1)}) = Q_i \left[\sum_j \exp(-\lambda_j^{(2)} - \beta c_{ij}) \right]^{-1}, \quad (13)$$

$$\exp(-\lambda_j^{(2)}) = D_j \left[\sum_i \exp(-\lambda_i^{(1)} - \beta c_{ij}) \right]^{-1}. \quad (14)$$

Чтобы представить окончательный результат в более привычном виде, запишем

$$A_i = \frac{\exp(-\lambda_i^{(1)})}{Q_i}, \quad (15)$$

$$B_j = \frac{\exp(-\lambda_j^{(2)})}{D_j}. \quad (16)$$

Отсюда

$$T_{ij} = A_i B_j Q_i D_j \exp(-\beta c_{ij}), \quad (17)$$

и в соответствии с уравнениями (13)–(16) имеем

$$A_i = \left[\sum_j B_j D_j \exp(-\beta c_{ij}) \right]^{-1}, \quad (18)$$

$$B_j = \left[\sum_i A_i Q_i \exp(-\beta c_{ij}) \right]^{-1}. \quad (19)$$

Таким образом, наиболее вероятное распределение транзакций является таким же, как в рассмотренной ранее гравитационной модели, которая определена уравнениями (5)–(7). Причем функция f была заменена на экспоненту с отрицательным показателем степени. Таким образом, статистическая теория утверждает [5, 7–9], что при заданных величинах запросов данных в удаленную сеть и приема их в каждой сети назначения и однородной цели запросов, при заданных затратах на передачу данных между сетями и фиксированных полных затратах на транспортировку, в корпоративной сети существует наиболее вероятное распределение транзакций между локальными сетями, и это распределение совпадает с тем, которое задается гравитационной моделью.

2 ИНТЕРПРЕТАЦИЯ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Особенность статистической механики состоит в том, что параметры, появляющиеся в уравнении, которое описывает наиболее вероятное распределение, имеют определенный физический смысл. То же самое справедливо и здесь в нашем случае. Величины Q_i , D_j и c_{ij} были определены ранее, а выражение $\exp(-\beta c_{ij})$ появляется в этой формулировке как явная форма зависимости от расстояния, причем параметр β определяется теоретически через уравнение (11). Но он имеет свою обычную интерпретацию. Он тесно связан со средним расстоянием между парами соединений: чем больше β , тем меньше среднее расстояние. Этот факт очевидным образом связан с величиной C в уравнении (8). Если увеличить C , то увеличатся затраты на передачу, и среднее расстояние между передачами возрастает; анализ левой части (8) показывает, что β при этом уменьшится.

Осталось интерпретировать A_i и B_j . Пусть одна из величин D_j изменяется, например, D_1 . Тогда величины

$$T_{i1} = A_i B_1 Q_i D_1 \exp(-\beta c_{i1}), \quad (20)$$

характеризующие запросы из каждой сети i в сеть 1, изменяются пропорционально. Величины A_i , как следует из уравнения (18), будут мало изменяться, поскольку выражение, включающее D_1 в каждом A_i – это только один из членов суммы. Величины B_1 возможно изменятся еще меньше, так как любое изменение в них будет вызвано изменениями в A_i .

Таким образом, роль A_i заключается (предполагая, что величина D_1 увеличилась) в небольшом сокращении числа всех запросов, которое компенсирует увеличение числа запросов в сеть 1. Следовательно, под A_i можно понимать некий конкурирующий параметр, сокращающий большинство запросов вследствие роста привлекательности одной сети. Знаменатель A_i также обычно используется как мера доступности, и можно сказать, что увеличение D_1 увеличивает доступность сети 1 для каждого пользователя корпоративной сети, хотя обычно мы будем пользоваться такой интерпретацией относительно изменений c_{ij} . Таким образом, этот анализ позволяет интерпретировать A_i в терминах конкурентной доступности сетевых ресурсов. Аналогичную роль играют величины B_j , которые связаны с основными изменениями в Q_i , а не в D_j . Изменение c_{ij} или несколько одновременных изменений Q_i и D_j приведет к сложному процессу перестройки A_i и B_j .

ВЫВОДЫ

Научная новизна. Проведенный анализ показал, что гравитационная модель имеет прочную основу и может быть применена для статистической оценки потоков

данных в магистральных каналах корпоративных компьютерных сетей. Однако следует заметить, что весь анализ проводился для одноцелевых транзакций и для однородной группы запросов. Транзакции в компьютерных сетях не являются идентичными в том смысле, в котором идентичны частицы в физике, поэтому вряд ли можно ожидать, что для реальной корпоративной сети можно получить достаточно точные результаты моделирования. Тем не менее, рассмотренный в работе подход может быть использован для предварительной оценки загрузки магистральных каналов передачи данных, и даже дать хорошие результаты, если удастся классифицировать транзакции по цели и по типу передач так, чтобы добиться разумной однородности потоков данных.

Практическая значимость. Предложенный подход может быть использован для предварительной статистической оценки потоков данных в корпоративной компьютерной сети. Полученные статистические оценки, в свою очередь, целесообразно использовать при построении общих моделей потоков, предусматривающих различия по множеству типов запросов, а также учитывающих различия в типах коммуникаций, соединяющих различные сети с использованием маршрутизаторов и шлюзов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Вишневский В. М. Теоретические основы проектирования компьютерных сетей. – М.: «Техносфера», 2003. – 512 с.
2. Иммельбаев Ш. С., Шмульян Б. Л. Анализ стохастических коммуникационных систем с применением термодинамического подхода // Автоматика и телемеханика. – 1977. – № 5. – С. 26–31.
3. Хакен Г. Информация и самоорганизация: макроскопический подход к сложным системам: Пер. с англ. – М.: Мир, 1991. – 240 с.
4. Гуржий А. Н., Коряк С. Ф., Самсонов В. В., Скляр А. Я. Контроль и управление корпоративными компьютерными сетями: инструментальные средства и технологии. – Харьков: «Компания СМІТ», 2003. – 664 с.
5. Клейнрок Л. Коммуникационные сети: Перев. с англ. – М.: «Наука», 1975. – 256 с.
6. Скляр А. Я., Макрушан И. А. Математическая модель синтеза иерархических структур систем управления с последовательным применением основных методов декомпозиции // АСУ и приборы автоматики. – 2005. – Вып. 131. – С. 69–73.
7. Советов Б. Я., Яковлев С. А. Построение сетей интегрального обслуживания. – Л.: Машиностроение. Ленинград. отд-ние, 1990. – 332 с.
8. Шеннон К. Работы по теории информации и кибернетике: Пер. с англ. под ред. Р. Л. Добрушина и О. Б. Лупанова, с предисловием А. Н. Колмогорова. – М.: «Изд. иностранной литературы», 1963. – 829 с.
9. Стратонович Р. Л. Теория информации. – М.: Сов. радио, 1985. – 486 с.

Надійшла 23.03.06
Після доробки 3.07.06

У даній статті розглянуті питання моделювання транспортних комунікацій корпоративних комп'ютерних мереж. Пропонуються два підходи рішення задачі: побудова моделі розподілу потоків даних на основі методів максимізації ентропії, а також використання гравіта-

ційної моделі розподілу транзакцій. Дано інтерпретацію отриманих результатів.

In the article the problems of modeling of transport communications of corporate network have been considered.

Two approaches to the problems of solution have been proposed: the construction of a model of data flow distribution on the basis of entropy maximization methods, as well as the usage of gravitation model of transactions distribution. The interpretation of received results has been given.

УДК 519.713

Е. Е. Сыревич

ВЕРИФИКАЦИЯ ВЫСОКОУРОВНЕВЫХ МОДЕЛЕЙ ЦИФРОВЫХ УСТРОЙСТВ

Предлагается стратегия верификации моделей цифровых устройств, описанных с помощью языков описания аппаратуры. Основная идея лежит в генерации различных псевдоисчерпывающих тестов для отдельных функциональных элементов, суперпозиции этих тестов и в интерактивном вычислении эталонных реакций.

ВВЕДЕНИЕ

В современной САПР РЭА основным способом описания устройств являются языки описания аппаратуры [1, 2]. С помощью языков описания аппаратуры (ЯОА), например VHDL или Verilog, проще и быстрее ввести и проверить проект в рамках САПР. 10 строками ЯОА можно описать как 1, так и 100000 триггеров. Описание устройств на ЯОА обладает рядом преимуществ. Проект на ЯОА – это объединение структуры цифрового устройства (ЦУ) и алгоритма его функционирования, одновременное описание ЦУ в статике и динамике. ЯОА – модель не требует дополнительного описания на человеческом языке и в виде схем, обладает высокой надежностью. Проект на ЯОА без труда переносится на другую элементную базу, например Altera FPGA или ASIC. Схема под конкретную элементную базу живет 2–4 года, а корректно написанный код на ЯОА может жить десятилетия. С помощью ЯОА можно описывать ЦУ на разных уровнях иерархии: алгоритмическом, структурном, регистровых передач и потоков данных, логическом и аналоговых схем.

Мировые компании по производству цифровых систем вынуждены постоянно сокращать время до попадания их продукции на рынок. А по оценкам производителей верификация, в том числе и функциональная, занимает до 80 % трудоемкости в цикле проектирования. Существует спрос на средства функциональной верификации моделей цифровых устройств на стадии их описания на поведенческом уровне с помощью конструкций языка VHDL.

Несмотря на обилие научных работ, связанных с верификацией и диагностикой цифровых устройств, в на-

стоящее время весьма актуальны и востребованы рынком инструментальные средства автоматической генерации тестов для функциональных моделей сложных ЦУ. Дефицит автоматических генераторов нового поколения испытывают и известные фирмы: Aldec, Altera, Actel, Xilinx, Synopsis.

Таким образом, цель данной работы – разработка стратегии верификации моделей ЦУ на языках описания аппаратуры, которая позволит уменьшить временные затраты на проектирование ЦУ путем использования методов сужения области возможных тестовых значений при автоматической генерации тестов.

1 ПОСТАНОВКА ЗАДАЧИ

Исходя из поставленной цели, для ее достижения решаются следующие задачи:

- 1) разработка внутренней модели цифрового устройства, представленного на ЯОА VHDL в целях последующей верификации;
- 2) разработка метода сужения количества подаваемых тестовых наборов;
- 3) разработка стратегии функциональной верификации моделей цифровых устройств на этапе ввода и моделирования проекта;
- 4) разработка методов оценивания качества верификации;
- 5) разработка информационного и методического обеспечения автоматического генератора тестов, поддерживающего стандарт VHDL.

Объектом исследования являются модели цифровых систем, проектируемые с помощью языков описания аппаратуры высокого уровня.

2 ВНУТРЕННЯЯ МОДЕЛЬ ДЛЯ ВЕРИФИКАЦИИ

Процесс построения тестов основывается на двух аспектах: внутренняя модель, позволяющая выполнять