ИНФОРМАЦИОННЫЕ СИСТЕМЫ. СРЕДСТВА, ТЕХНОЛОГИИ, БЕЗОПАСНОСТЬ

СИСТЕМА АНАЛИЗА HTML ПОТОКОВ

А.А.Борисов, Л.Ю.Ротков

Нижегородский госуниверситет

Активное внедрение сетевых технологий привело к ускоренному росту нагрузок на локальные сети. В новых условиях к сетям предъявляются более жесткие требования относительно качества предоставляемого сервиса. Типична ситуация когда на созданную ранее сеть, после некоторой ее доработки возлагают новые задачи, с которыми она плохо справляется. В связи с этим становится весьма актуальной задача диагностики сети.

Первые сетевые анализаторы при диагностике сети считывали заголовки пакетов данных, предоставляя администраторам информацию о сетевых адресах, размере пакетов, с предоставлением результатов в виде графов и текстовых описаний. Анализаторы помогали сетевым администраторам провести диагностику серверов, сетевых каналов, концентраторов и коммутаторов, а также приложений.

Современное множество анализаторов можно подразделить на два вида. К первому относятся автономные программные продукты, устанавливаемые на мобильном компьютере. Эти продукты способны провести диагностику состояния сетевых каналов и сетевого оборудования. Второй вид анализаторов предназначен для мониторинга и управления сетью, входит составной частью во многое коммуникационное оборудование и позволяет контролировать локальные и глобальные сетевые службы. Эти модули дают администраторам целостное представление о состоянии сети. Например, с помощью таких продуктов можно определить, какие из приложений выполняются в данный момент, какие пользователи зарегистрированы в сети и кто из них генерирует основной объем трафика.

Современные сетевые анализаторы решают следующие задачи:

- протоколирование работы сетевого сегмента;
- определение стека используемых протоколов;
- расчет загруженности канала;
- расчет распределения трафика по протоколам;
- определение характера проблем, возникающих в сети.

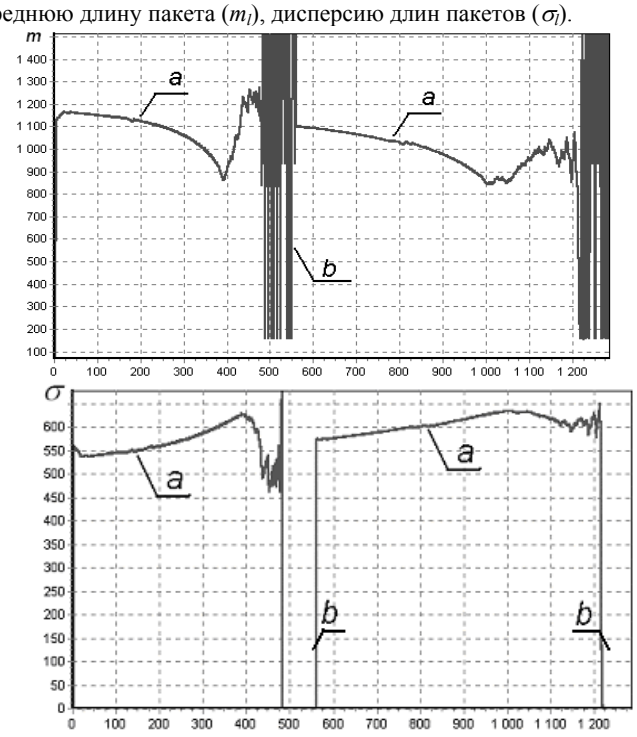
Для решения всех вышеперечисленных задач в совокупности в режиме реального времени требуются большие вычислительные ресурсы.

Для минимизации необходимых вычислительных ресурсов поставим задачу диагностики состояния сети с использованием двухпараметрической модели трафика.

Будем считать, что передаваемый поток является случайным процессом, но при обмене данными между рабочими станциями существуют серии запросов и ответов. Под серией понимается запрос рабочей станции к серверу, на который сервер генерирует некоторое число пакетов.

В качестве параметров модели используем длину пакета (l) и время отправки пакета в сеть (t).

Применим метод скользящего окна для исследования двухпараметрической модели $A\{l_i,t_i\}$. В качестве выходных параметров определим статистические характеристики трафика: среднюю длину пакета (m_l), дисперсию длин пакетов (σ_l).



На рисунках показано изменение m_l и σ_l при передаче по сети файлов. Участки «*a*» рисунков соответствуют устойчивому режиму обмена, участки «*b*» отвечают передаче служебной информации, т.е. с точки зрения передачи данных канал используется неэффективно.

Таким образом уже по двум параметрам в режиме реального времени и при незначительных вычислительных затратах можно судить о состоянии канала передачи данных.

[1] Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. –СПб: Издательство «Питер», 2000. 672с.

[2] Рытов С.М. Введение в статистическую радиофизику. –М.: Наука, 1976. 496с.

УЛУЧШЕНИЕ СВОЙСТВ СИНХРОНИЗАЦИИ В СИСТЕМЕ МОРСКОЙ ПОДВИЖНОЙ СЛУЖБЫ NAVTEX СТАНДАРТА “B”

М.В.Докукин

Балтийская государственная академия рыбопромыслового флота

Согласно [1], для синхронизации в системе морской подвижной службы NAVTEX стандарта “B” передаются два фазирующих сигнала *phasing signal 1 (1111000)* и *phasing signal 2 (0110011)* четыре раза подряд. Условием синхронизации является прием последовательности *phasing signal 1 (PS1)* и *phasing signal 2 (PS2)* или *phasing signal 2* и *phasing signal 1* и последующий прием любого из этих фазирующих сигналов. Таким образом, для синхронизации в системе “NAVTEX” необходимо принять без ошибки три фазирующих сигнала с выше описанным условием.

Для повышения надежности синхронизации приема предлагается провести некоторые преобразования синхропоследовательности.

1) Принятая синхропоследовательность;
2) Разбиение на блоки (рис. 1);
3) Циклический сдвиг влево на один бит каждый блок (рис. 2);
4) Перемежение битов блока и разбиение их на две последовательности произведем согласно следующим формулам:

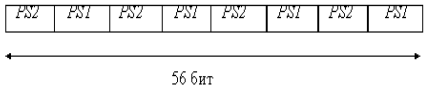


Рис. 1

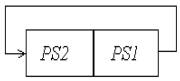


Рис. 2

$$K1(2 \cdot i + 1 - i) = N(2 \cdot i + 1),$$

$$K2(i + 1) = N(2 \cdot i + 2)$$

для $i = 0,1,...,6$, где $N(i)$ – бит i-ой позиции в синхропоследовательности N, а $K1(i)$ и $K2(i)$ – биты i-ой позиции в последовательностях $K1$ и $K2$. Знак равно означает знак соответствия номеров битовых позиций;

5) Формирование двух последовательностей L и M (рис. 3).

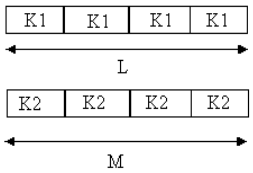


Рис. 3

По окончанию процедуры перемежения последовательность $K1$ будет соответствовать $K2$, поэтому L и M последовательности будут идентичны.

В связи с тем, что для приемного устройства иногда возникает задача определить начало передачи сообщения при пропуске части синхропоследовательности (например, из-за воздействия помех в канале передачи) необходимо представить последовательность $K1$ как слово циклического кода. При этом любые подряд идущие семь бит последовательностей L и M можно представить как кодовое слово циклического кода, так как L и M состоят из четырех последовательностей $K1$. По информационной части данного слова возможно однозначно определить начала пе-

редачи. Можно доказать, что последовательность $K1$ является кодовым словом полиномиального несистематического циклического кода (7,3) с порождающим полиномом [2]:

$$g(x) = x^4 + x^2 + x + 1.$$

Данный код ортогонален коду Хемминга и способен исправлять все однократные и часть двукратных ошибок. В случае пропуска четного количества бит начала синхропоследовательности и при использование алгоритма преобразования первые семь бит последовательностей L и M преобразуются в кодовые слова кода (7,3). Информационная часть данных слов будет соответствовать количеству пропущенных бит. В случаях нечетных пропусков начала синхропоследовательности вся последовательность должна быть сдвинута на один бит влево (то есть, удалится первый бит), чтобы первые семь бит последовательности L соответствовали первым семи битам последовательности M.

На рис.4 представлены результаты сравнения вероятности отказа от синхронизации алгоритма разработанного на базе декодирования каскадно-кодовых конструкций (Рот.синхр.А) и существующего в NAVTEX (Рот.синхр.N) в зависимости от вероятности ошибки x. Расчет проводился для модели симметричного гауссовского канала.

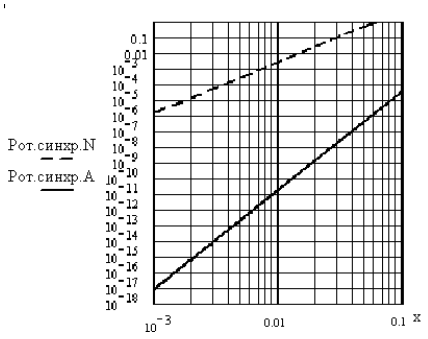


Рис. 4

Из представленных графиков следует, что предложенный способ синхронизации (Рот.синхр.А) позволяет на несколько порядков уменьшить вероятность отсутствия синхронизации при приеме сообщения системы NAVTEX.

[1] GMDSS Handbook, IMO London, 1992.
[2] Мак-Вильямс Ф.Дж., Слоэн Н.Дж. Теория кодов, исправляющих ошибки/ Пер. с англ. И.И.Грушко и В.А.Зиновьева. Под ред. Л.А.Бассалыго. –М.: Связь, 1979, с.157.

СРАВНЕНИЕ СПЕКТРОВ РАЗНОСТЕЙ ЧАСТОТ ВОДОРОДНЫХ СТАНДАРТОВ ПРИ РАЗЛИЧНЫХ ИНТЕРВАЛАХ ИХ ВЫБОРКИ

А.В.Пастухов[1)], К.Г.Кирьянов[1)], Р.Ф.Фахруллин[2)], А.В.Шабельников[3)]

[1)]ФГУП ННИПИ «Кварц», [2)]НГТУ, [3)]ИРЭ АН РФ

Достигнутый уровень стабильности частоты $\Delta f / f$ промышленного водородного стандарта Ч1-75 составляет примерно 10^{-14} [1]. Это на 1,5–2 десятичных порядка хуже, чем зависящий только от атомных констант и температуры теоретический предел, полученный при определённых идеализациях на основе представлений о физической модели используемого в стандарте квантового перехода. Предельные возможности по стабильности частоты в атомных стандартах полностью, по-видимому, пока не реализованы. Косвенно этот факт можно наблюдать в экспериментах, на устройствах сличения, по зависимости от времени выборок разности частот двух стандартов и относить его к "техническим" уходам частоты. Поэтому одной из важных задач метрологии и измерительной техники в области частотных и временных измерений (госслужба частоты и времени, радионавигация, радиоастрономия, геодезия, научные исследования и т.д.) является учёт у существующих водородных стандартов частоты и времени *прогнозируемой и аттестуемой систематической составляющих "медленных" и "быстрых" уходов частоты*.

Знание прогнозируемых составляющих точности у стандартов Ч1-75 и др. типов и правил (методик, алгоритмов) их учёта позволяет улучшить метрологические характеристики стандартов наряду с традиционно и постоянно ведущимся технологическим совершенствованием квантовой (вакуумная колба и др.) и радиотехнической (система преобразователей частоты и т.п.) частей прибора. В сложившейся ситуации представляет интерес и важны любые возможности, которыми можно воспользоваться при проектировании новых типов стандартов частоты с улучшенными характеристиками по точности.

На сегодняшний день известны следующие обстоятельства.

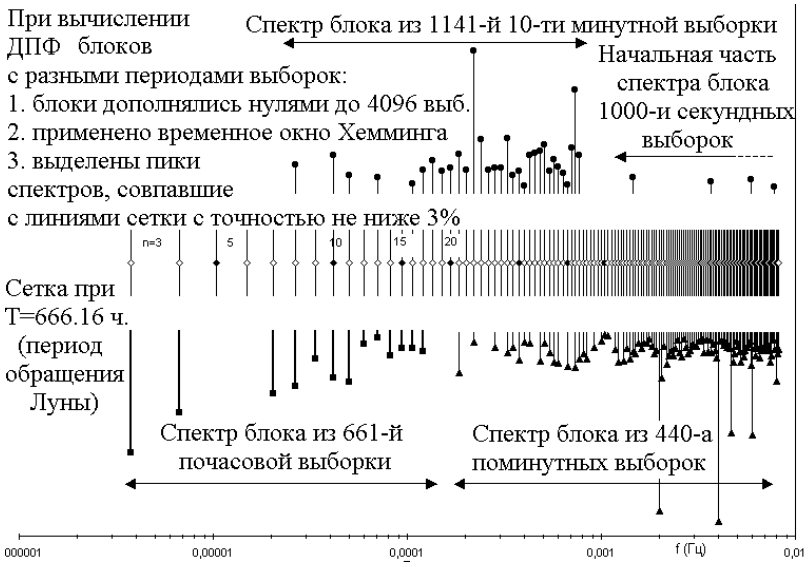
1. Существует физическая гипотеза о гравитационной природе влияния глобальной синхронизации вращающихся масс на спектры процессов различной физико-химической природы, протекающих в земных условиях [2,3];

2. Согласно этой гипотезе в мало зашумлённых земных процессах должны наблюдаться специфические "сетки частот", связанные жёстко с периодами вращения упомянутых масс (Солнца, Луны, планет и т.д.) [4,5]. Эти закономерности довольно хорошо [6-9] наблюдались и описаны в работах различных организаций (ИРЭ РАН, МГУ, ФГУП ННИПИ "Кварц", ННГУ, ИФЗ РАН и др.);

3. Существуют публикации, из которых следует, что наличие сетки частот в спектрах сигналов на выходе частотного компаратора позволяет учесть параметры сетки в прогнозирующих моделях ухода частоты [7,8].

В цитированных работах рассматривались, насколько известно, сетки частот в природных процессах с постоянным периодом выборки. В настоящей работе наличие спектральной сетки устанавливалось *для различных интервалов выборки.*

(1 час, 10 минут, 1 минута, 1 секунда) разности частот водородных стандартов (см. совмещенные графики спектров; подробные таблицы полученных спектров не приводятся ввиду их большого объёма). Совпадения частот в спектрах блоков данных с разными периодами выборки указывает на объективность искомой закономерности и возможность её учёта при прогнозировании.



Работа выполнена при частичной поддержке РФФИ (грант 02-02-17573).

[1] Стандарт частоты и времени Ч1-75. Рекл. Проспект ННИПИ “Кварц”, 2001.
[2] Вернадский В.И. Размышления натуралиста –М.: Наука, 1975.
[3] Чижевский А.Л. Земное эхо солнечных бурь. Изд. 2-е. –М.: Мысль, 1976.
[4] Кирьянов К.Г., Шабельников А.В. //Вестник ВВО АТН РФ, Серия: высокие технологии в радиоэлектронике, 2(4), 1997, с30.
[5] Шноль С.Э. //Биофизика. 1995. Т. 40, вып. 4, с725.
[6] Кирьянов К.Г., Шабельников А.В. //Биофизика. 1998, т.43, №5, с874.
[7] Ефременко В.В., Кирьянов К.Г., Шабельников А.В. и др. Исследование возможностей и разработка методов прогноза временных изменений параметров некоторых природных процессов на глобальном и региональном уровнях. Инициативный отчет по НИР «Прогноз». ИРЭ РАН, 1993.
[8] Кирьянов К.Г., Шабельников А.В. //Радиотехника и электроника. 95, в. 5, с.753.
[9] Кирьянов К.Г., Пастухов А.В., Фахруллин Р.Ф., Шабельников А.В. //В кн.: Тр. (пятой) научной конференции по радиофизике 7 мая 2001г./Ред. А.В. Якимов. –Нижний Новгород: ТАЛАМ, 2001, с353.

О РАЗРАБОТКЕ АЛГОРИТМОВ ВЫЧИСЛЕНИЯ МАТЕМАТИЧЕСКИХ ФУНКЦИЙ В СТАНДАРТЕ IEEE-754

Е.Н.Гладков

Нижегородский госуниверситет

Потребность в разработке новых алгоритмов для вычисления известных математических функций может быть связана как с появлением новых процессоров, так и с повышенными требованиями к точности и скорости вычислений.

В статье излагается опыт разработки новых алгоритмов, удовлетворяющих стандарту IEEE-754 [1]. Этот опыт может быть также использован при разработке алгоритмов и в других форматах представления вещественных чисел с плавающей запятой.

Известно, что стандарт обычно задает ограничения не только на точность представления исходных данных, но и на точность результатов вычислений. Так в IEEE-754, допустимая ошибка для арифметических операций должна быть не более 0.5 единицы младшего разряда мантиссы (0.5 ulp – units in the last place). Но при вычислении трансцендентных функций, невозможно написать алгоритм в арифметике с ограниченной точностью, который в любой заданной точке гарантировал бы ошибку не более 0.5 ulp (Table Maker's Dilemma [2]). Поэтому обычно при написании библиотек математических функций ставится целью по возможности минимально отступать от «идеальной» ошибки 0.5 ulp.

Алгоритм вычисления математической функции $F: X \to Y$, где $X \subseteq R^1$, $Y \subseteq R^1$ обычно состоит из трех стадий: *приведение аргумента, вычисление, и реконструкция (восстановнение) результата.*

1) *Приведение аргумента* заключается либо в простой перекодировке аргумента, либо замене его на новый аргумент то есть смещении аргумента в область, удобную для вычислений при условии, чтобы искомое значение легко восстанавливалось. Здесь существенную роль играют и свойства вычисляемой функции, и формат представления чисел, и специфика процессора, на котором предполагается производить вычисления.
2) *Вычисление* чаще всего реализуется либо ньютоновскими итерациями либо с помощью отрезка ряда Тейлора или полинома наилучшего равномерного приближения.
3) *Реконструкция (восстановление)* часто сводится к простым (надежным) арифметическим операциям и подгонке под стандарт представления результата.

На каждой стадии вычисления возможны как погрешности метода, так и погрешности вычисления. Поэтому планировать их надо так, чтобы величина результирующей погрешности лежала в допустимых пределах (например, меньше 0.6 ulp, что достаточно близко к «идеальным» 0.5 ulp).

Например, вычисление e^x в формате IEEE-754 Single Precision можно реализовать в следующие 3 этапа.

1) *Приведение аргумента.* По данному аргументу x вычисляем две переменные: $n=[x*\log_2(e)]$, $r=x*\log_2(e)-n$, где за $[x*\log_2(e)]$ обозначена целая часть числа $x*\log_2(e)$. В данном случае приведенный аргумент – это набор величин (n,r).
2) *Вычисление.* На интервале [–0.5; 0.5) вычисляем e^r при помощи полинома наилучшего равномерного приближения P(r) степени 7.
3) *Реконструкция.* $e^x = 2^{x*\log2(e)} = 2^n * 2^r = 2^n * P(r)$.

Но этот простой способ вычисления e^x, не позволяет достичь точности 0.6 ulp, если бы все промежуточные вычисления выполнялись с той же точностью, в которой задан аргумент x. Чтобы достичь 0.6 ulp, этот алгоритм требует усовершенствования.

В настоящее время существуют программные системы, позволяющие вычислять математические функции в арифметике «неограниченной точности». Одной из таких систем является Maple V.

Автором статьи написана библиотека функций на языке системы Maple V, помогающая разрабатывать новые эффективные алгоритмы вычисления математических функций на компьютерных архитектурах с вещественной арифметикой, удовлетворяющей стандарту IEEE-754. Эта библиотека позволяет также эмулировать вычислительную работу блока FPU (Floating Point Unit) компьютера (с промежуточными погрешностями и их накапливанием). Разработанная библиотека содержит следующие группы процедур-функций:

- функции преобразования вещественных чисел между различными форматами представления (научный, IEEE-754 или C99). Поддерживаются 4 режима округления: к ближайшему, к нулю, к плюс (минус) бесконечности;
- функции определения погрешности вычислений (в единицах ulp, относительная, абсолютная);
- функции генерации коэффициентов минимаксных полиномов, разбиения констант и др.;
- арифметические и логические операции над шестнадцатеричными целыми числами.

С помощью этой библиотеки разработано несколько алгоритмов для библиотеки LIBM для архитектуры IA-64 [3], например, гиперболические функции (tanh, atanh, asinh, acosh), специальные функции (erf, erfc, lgamma, tgamma). Система Maple V применялась для вычисления вспомогательных констант, коэффициентов аппроксимационных полиномов и для анализа вычислительной погрешности.

[1] ANSI/IEEE 754-1985, American National Standard for Binary Floating-Point Arithmetic.
[2] Goldberg D. What Every Computer Scientist Should Know About Floating-Point Arithmetic. Online. http://www.validgh.com/
[3] Open source LIBM for IA-64. Online. http://developer.intel.com/software/products/opensource/libraries/num.htm.

ПОДХОДЫ К ПОСТРОЕНИЮ ЗАЩИТЫ ПРОГРАММ ОТ ИССЛЕДОВАНИЯ ИХ КОДА ЗЛОУМЫШЛЕННИКОМ

С.В.Корелов[1)], Д.Л.Туренко[2)], С.В.Калинин[2)]

[1)]*ФАПСИ,* [2)]*Нижегородский госуниверситет*

С развитием отечественного рынка программ возрастает значение их защиты от несанкционированного копирования. Для большинства фирм и организаций, занимающихся разработкой программного обеспечения (ПО), продажа копий – основное средство существования. Поэтому каждый разработчик, несомненно, рано или поздно сталкивается с проблемой защиты ПО.

Чтобы модернизировать программу, надо понять принцип работы, определить основные и вспомогательные части. Для программ существуют специальные инструменты, позволяющие разбирать их «до винтика».

Все средства исследования ПО можно разбить на несколько категорий [1]:

- статические;
- динамические;
- средства мониторинга и другие средства.

Первые оперируют исходным кодом программы как данными и строят ее алгоритм без исполнения. Из всех известных дизассемблеров самым мощным признается IDA (Interactive DisAssembler).

Динамические средства изучают программу, интерпретируя ее в реальной или виртуальной вычислительной среде. Самым лучшим признается SoftIce от NuMega, предназначенный для поиска ошибок в программах и их отладки и используемый злоумышленником, как средство исследования программ.

К средствам мониторинга относятся такие утилиты, как FileMon, RegMon и (возможно) VxDMon, написанные Марком Руссиновичем, известным экспертом в области низкоуровневого программирования для Windows. FileMon показывает детальную информацию обо всех обращениях к файловой системе; имеется возможность устанавливать различные фильтры и даже отслеживать работу со swap-файлом. RegMon делает примерно то же самое, но не с файловой системой, а с реестром Windows. VxDMon, соответственно, осуществляет мониторинг вызовов различных сервисов VxD, что в некоторых случаях тоже может оказаться полезным.

Абсолютную защиту от всех видов атак создать невозможно. Критерием надежности защиты является соизмеримость времени, затраченного на ее написание, со временем, затраченным на ее снятие. Программа-минимум – это обеспечить невозможность подбора пароля (регистрационного ключа). Если у злоумышленников не будет другого выхода, кроме как написать crack (patch), то это уже весьма неплохо: с выпуском очередной версии им снова придется напрягаться.

Эффективными являются следующие подходы к созданию защиты:

1) Шифрование.
2) Предварительное архивирование.
3) Метод самогенерируемых кодов.

4) Метод самомодификации кода.

Шифрование файлов – наиболее простое средство для реализации. Достаточно, например, к каждому байту модуля добавить некоторую константу, чтобы дизассемблер ничего “не понял”. Усиление защитного эффекта придается с использованием поэтапной дешифрации, в различных местах программы и в разные моменты ее работы.

Предварительное архивирование (упаковка) также является эффективным методом. Использование интеллектуальных упаковщиков затрудняет помимо исследования дизассемблером еще и отладку. Упакованные такими средствами исполняемые файлы распаковываются динамически небольшими блоками.

Следующий подход к созданию защиты – использование метода самогенерируемых кодов. Суть его такова. В исполняемый модуль записывается массив данных, являющийся исполняемым кодом, который реально получает управление в ходе выполнения программы.

Эффективным способом защиты ПО является метод самомодификации кода.

В тексте программы организовывается участок, где будут храниться цепочки команд, с указанием адресов эквивалентных им участков. При очередной работе программа случайным образом меняет местами отдельные части из собственного тела и этого массива. Таким образом, после очередного запуска программы ее код будет отличаться от последовательностей команд предыдущего запуска.

Более сложным способом является модификация кодов команд с изменением характера выполняемых операций. Осуществляется это следующим образом. Отлаженный модуль транслируется в объектный код с получением листинга. Не обращая внимания на мнемонику, ищем участки кода с похожими закономерностями изменения величин кода программ. Затем выделяем эти участки в отдельную подпрограмму и составляем алгоритм ее преобразования в коды соответствующих участков.

Возможно также использование какого-либо осмысленного участка текста как для данных, так и для загрузки нужных регистров, то есть одновременное использование некоторых байт данных в качестве операторов и операндов.

Полезным будет создание в исполняемом модуле участка кода, распознающего активность отладчика и средств мониторинга, после чего пустить программу по «ложному пути», совершая действия, отличные от тех, которые выполняются при нормальной работе.

Можно запрограммировать защиту в виртуальном драйвере.

В данной статье рассмотрены далеко не все методы и приемы защиты программ от исследования. Очень много советов по созданию защищенного кода можно найти на сайтах и форумах в сети Internet.

[1] Каталов В. Защита shareware-программ // Компьютерра, №12 (240), 1998 г.

ФОРМАЛИЗАЦИЯ ПРЕДСТАВЛЕНИЯ АПРИОРНОЙ ИНФОРМАЦИИ В ОБУЧАЮЩЕЙСЯ ЭКСПЕРТНОЙ КРИМИНАЛИСТИЧЕСКОЙ СИСТЕМЕ

А.Т.Надеев, О.С.Данилова, С.В.Кошелев

Волго-Вятская академия государственной службы

На предыдущей научной конференции говорилось о том, что на основе априорной и оперативной (апостериорной) информации, получаемой правоохранительной системой возможно использование формальных моделей уголовного процесса для принятия статистически обоснованных решений на всех стадиях расследования и судебного разбирательства. Было также показано, что с этой точки зрения уголовный процесс поддается достаточно адекватной формализации.

В настоящее время нами завершена разработка эскизного проекта соответствующей экспертной системы. Ее работа в оперативном режиме (т.е. в режиме расследования) состоит из следующих основных стадий:

- определение на основе первичной оперативной информации множества признаков, описывающих конкретное совершенное преступление;
- обобщение указанных признаков;
- определение на их основе множества обобщенных версий;
- конкретизация следователем обобщенных версий и формирование списка (множества) конкретных версий данного преступления;
- ранжирование указанных версий на основе имеющейся статистики по соответствующим условным априорным вероятностям и продуктивностям;
- формирование множества конкретных доказательных признаков версий, получаемых в ходе следственных действий и розыскных мероприятий по принципу «за» и «против», их агрегирование в эквивалентные комплексные, выбор схем, описывающих надежность указанных агрегированных признаков;
- расчет апостериорных вероятностей истинности прорабатываемых версий и уточнение значений продуктивности;
- исключение версий принципу «полного алиби»;
- расчет отношений правдоподобия, характеризующих доказательную силу (надежность доказательной базы) прорабатываемых версий;
- принятие статистически обоснованных решений, касающихся либо продолжения расследования, либо его прекращения за не доказанностью или по причине отсутствия продуктивных версий или по причине исчерпания лимита времени, либо передачи материалов дела в суд.

Вся работа экспертной системы строится вокруг так называемой матрицы версий. Кроме того, в ней широко используются различного рода классификаторы и матрицы обобщенных статистических априорных данных.

Для решения указанных выше задач необходимо формализовать понятие версии в уголовном деле. Для этого выделяются следующие факторы преступления: объект, субъект, условия, технология и мотив. Далее проводится классификатор каждого из факторов по определенным группам признаков. Так, например, субъект

может быть классифицирован по возрасту, социальному положению, судимости, образованию и т.д. После этого составляется матрица встречаемости версий, каждая из строк которой соответствует заданному признаку. Столбцы матрицы соответствуют раскрытым уголовным делам. В столбце ставим 1 если соответствующий признак реализуется и ноль, если этот признак отсутствует. Таким образом, формально обобщенная версия преступления $\mathbf{A}_i$ представляет собой n-компонентный вектор столбец, где n общее число признаков. Представим вектор $\mathbf{A}_i$ в виде совокупности двух подвекторов $\mathbf{A}_i=(\mathbf{r}, \mathbf{s})$, где $\mathbf{r}$ состоит из признаков субъекта и мотива, а $\mathbf{s}$ из признаков объекта, технологии и условий. Таким образом, при расследовании конкретного уголовного дела следователя формально интересует вектор $\mathbf{r}$, соответствующий заданному $\mathbf{s}$. Причем заданному вектору $\mathbf{s}_k$ может соответствовать несколько векторов $\mathbf{r}_j$. Как уже было сказано выше, экспертная система проводит ранжирование версий в соответствии с условными вероятностями $P(\mathbf{r}|\mathbf{s})$. Однако, при частотном определении значения вероятности возникает проблема размерности, связанная с большим числом признаков внутри каждой из групп, а также с большим числом самих групп признаков. Действительно, если вектору $\mathbf{r}$ соответствует, например, 7 групп признаков в среднем по 10 компонент в каждой, то число векторов $\mathbf{r}$, соответствующих заданному $\mathbf{s}$ будет равно 10^7. Число векторов $\mathbf{s}$ при этом оказывается еще больше. Чтобы решить эту проблему нужно создавать криминалистические системы не только для определенного типа преступления, но и для конкретного объекта, сократить число признаков в каждой из групп, за счет выведения ненужных, указывающих на одного и того же субъекта. Тогда априорные вероятности отрабатываемых версий вычисляются по формуле

$$P(\mathbf{r}_j|\mathbf{s}_k) = \frac{n_{jk}}{\sum_k n_{jk}}$$

где n_{jk} число векторов в матрице встречаемости версий, соответствующих данному вектору $\mathbf{s}_k$.

На основе полученных априорных вероятностей происходит ранжирование версий и планирование процесса расследования до поступления дополнительной (доказательной) информации. Следует отметить, что в результате описанной процедуры система формирует множество обобщенных версий. Далее, по мере поступления доказательных признаков в систему должна быть введена информация о конкретных доказательных признаках основных составляющих версии (субъекта и мотива преступления). Пользователь системы заполняет предлагаемую ему таблицу признаков преступления, на основе которой будет происходить уточнение множества обобщенных версий. С помощью процедуры Байеса, а также на основе схем надежности для комплексов доказательств и отдельных доказательных признаков, а также соответствующих классификаторов доказательных признаков осуществляется оценка надежности для каждой из множества обобщенных версий и ранжирование версий в соответствии с показателем надежности. Процедура повторяется для всего перечня доказательных признаков.

ОПТИМИЗАЦИЯ СТРУКТУРЫ ИНТЕРНЕТ- ДОКУМЕНТОВ ДЛЯ СОКРАЩЕНИЯ ОБЪЕМА ПЕРЕДАВАЕМОЙ ИНФОРМАЦИИ

С.Л.Моругин, А.В.Потехин, М.В.Ширяев

Нижегородский государственный технический университет

Рассматривается выбор структуры Интернет-документов и способов навигации по таким документам при совместной работе клиента и Интернет-сервера в целях минимизации объема передаваемой информации при многократной передаче изменений в составе документа.

Будем представлять электронный документ как источник сообщений. В качестве элементарного сообщения возьмем элемент структуры документа. Любой документ, несущий в себе информацию, изначально можно представить в виде линейной последовательности составляющих его элементов. Последовательность элементов, составляющих документ, обозначим $u_1, u_2,\ldots,u_i,\ldots,u_N$, где N - количество структурных элементов документа (объем алфавита источника, в нашем случае электронного документа). Чем больше значение N, тем больше энтропия источника $H(U)$ и тем больше информации может нести одно сообщение. Соответственно тем более необходимым является рассмотрение способов оптимальной передачи информации с целью уменьшения загрузки канала.

Из практики работы с документами можно заметить, что различные элементы документа обновляются с разной частотой. Элементы описания, как правило, вообще не изменяются в процессе работы с документом и, наоборот, динамические данные меняются часто. Соответственно логично будет использовать такой критерий для структуризации, как вероятность обновления данного элемента в документе - $P(u_i)$. Вероятность обращения к элементам документа описывается некоторым законом распределения.

Рассмотрим процесс обращения к элементам документа с целью их просмотра и изменения, характеризуемый вероятностью обращения к элементу в течение некоторого продолжительного промежутка времени T.

Для документов с динамически обновляемой информацией обработка значений вероятности обращения к структурному элементу за продолжительный промежуток времени дает сокращение объема пересылаемых данных по сравнению с использованием вероятности появления элемента в самом документе (обычное статистическое кодирование), так как с точки зрения многократной передачи оптимальным будет код, который приводит к минимизации передаваемых данных в среднем на всем промежутке времени работы с документом.

$$\sum_{t=0}^{T} D\ (P_t) < \sum_{t=0}^{T} D\ (P_v)\ ,$$

где $D(P_t)$ - объем данных, пересылаемых при использовании вероятности, учитывающей частоту обращения к элементу за время работы с документом, $D(P_v)$ - объ-

ем данных, пересылаемых при использовании вероятности учитывающей частоту появления элемента в структуре документа, T – время работы с документом.

Последовательность вероятностей, описывающих изменения i-го элемента из всей совокупности структурных элементов, обозначим $P(u_1)$, $P(u_2)$,…, $P(u_i)$,…,$P(u_N)$. Вероятность $P(u_i)$ задается для i-го элемента на основании статистики его изменения. Вид зависимости величины вероятности от местоположения элемента в исходной структуре документа может принимать любую форму. Определенная закономерность может исходить из логической и функциональной зависимости расположения элементов u_i при условии их наличия (что зависит от типа документа и формата данных).

Исходя из вышесказанного, структуризацию документа следует производить в соответствии с величиной вероятности обновления определенного элемента.

Чем больше вероятность изменения элемента $P(u_i)$, тем быстрее должен производиться доступ к нему, соответственно данный элемент в структуре должен находиться ближе к ее началу для ускорения его поиска (сокращение затрат на навигацию).

Исходя из этого критерия, производится построение структуры документа в виде дерева, в котором префиксные коды соответствуют путям к элементам. Соответственно путь к элементу, вероятность изменения которого велика, будет кодироваться более короткой бинарной последовательностью.

Одним из возникающих на данном этапе вопросов, является вопрос о степени зависимости структуризации на основе вероятностей от логической (смысловой) и функциональной структуризации документа. Существуют два крайних случая:

1. Все элементы имеют сильную семантическую зависимость друг от друга. Тогда вероятности изменения элементов, стоящих рядом друг с другом, равны по своей величине, и, соответственно, структура, построенная по вероятностному критерию, будет очень похожа на семантическую (логическую) или функциональную, по крайней мере, структурой поддеревьев. Данные структуры имеют большую величину корреляции.

2. Если элементы независимы друг от друга, то структура, построенная на оценке вероятности изменения элемента, может иметь совершенно другой вид, нежели логическая или функциональная. Данные в структуре будут слабо коррелированны.

Выбор размера структурного элемента производится с учетом доли информации на навигацию, общего количества элементов и их взаимосвязи. Главным критерием является минимум передаваемой информации. В первую очередь необходимо обратиться к таблице состояний и, анализируя слова и их состав, можно сделать приближенное заключение о размере структурной единицы, так как именно в данной таблице изначально фиксируются только элементарные изменения.

Удобными программными средствами для реализации идеи структурирования изменений в документах являются Java-модули и представление документа в формате XML.

СЖАТИЕ ГЛАДКИХ СИГНАЛОВ ПРИ ПЕРЕДАЧЕ ПО КАНАЛУ СВЯЗИ С ОГРАНИЧЕННОЙ ПРОПУСКНОЙ СПОСОБНОСТЬЮ

С.Л.Моругин, Т.В.Моругина

Нижегородский государственный технический университет

Сжатие непрерывных сигналов может быть выполнено на основе использования априорной информации об общих свойствах классов сигналов, например, об ограниченности производных (конечных разностей) до определенного порядка.

Рассматривается применение нового метода сжатия непрерывных сигналов класса $W^r_\infty([a,b],C^r)$ (сигналов, заданных на интервале $[a,b]$, имеющих ограниченные величиной C^r производные порядка r), названного *методом многослойной аппроксимации производных* (*МАП*). В качестве исходных данных задана функция $x(t) \subset W^r_\infty([a,b],C^r)$, длина Δ интервала $[a,b]$, допустимая погрешность аппроксимации ε_0. В методе МАП для аппроксимации сигнала применяется кусочно-непрерывная последовательность полиномов степени r, у которых производная $x^{(r)}(t)$ изменяется с шагом дискретизации h и после нормировки принимает одно из значений {-1, 0, +1}. Величина $x^{(r)}$ кодируется символами, передаваемыми по каналу связи. Метод МАП по сравнению с известными методами аппроксимации характеризуется тем, что при незначительном усложнении алгоритма может реализовать высокий порядок аппроксимации r, причем последний выбирается в соответствии с допустимой погрешностью задания исходных данных. Метод может использоваться для обработки сигналов в системах реального времени и не требует обработки будущих значений сигнала, используя для обработки всего r прошлых значений сигнала; является экстремальным и реализует ε-энтропию класса сигналов [1].

Для канала с ограниченной пропускной способностью рассмотрим задачу сжатия и восстановления сигнала с требуемой точностью, исходящей из допустимого объема передаваемой информации.

Постановка задачи. Для канала связи с ограниченной пропускной способностью надо передать непрерывный сигнал с требуемой погрешностью ε_0 при условии, что за интервал Δ можно передать не более H единиц информации. Запаздывание восстановленного на приемном конце сигнала не должно превышать D единиц времени, если требования к величинам ε_0, H и D не противоречивы.

При решении такой задачи следует использовать принцип достижения максимальной точности при заданном потоке информации.

Рассмотрим сигналы с медленно изменяющимися характеристиками класса $W^r_\infty([a,b],C^r)$, в частности, пусть медленно меняется величина C^r.

Для класса сигналов введем величину $L_n^{\mathrm{T}}(t)$ - константу Липшица для n-ой производной, максимальную на интервале некоторой длины T ($T << \Delta$):

$L_n^{\mathrm{T}}(t) = sup\, L_n(t+\xi)$ при $\xi \in [0, \Delta]$.

Обозначим L^{T}_{min} – минимальное значение константы $L_n^T(t)$ на всем входного сигнала интервале длиной Δ, L^{T}_{max} – максимальное значение константы $L_n^T(t)$ на том же интервале.

Оценим величину γ относительных потерь метода кодирования как относительную разность между объемом данных $I(u)$ кодированного сообщения $u(t)$ и величиной ε-энтропии сообщения $H(u)$:

$\gamma = (I(u) - H(u)) / H(u)$.

Пусть функция $L^T_n(t)$ медленно меняется по мере изменения аргумента t, удовлетворяя условию Липшица в виде

$|L^T_n(t_1) - L^T_n(t_2)| < \alpha\ |t_1 - t_2|$.

с константой $\alpha > 0$.

Тогда справедливо следующее утверждение. Величина относительных потерь метода МАП стремится к нулю при $\alpha \to 0$.

Считаем время запаздывания D конечным, но значительно превосходящим длительность одного шага h.

Метод МАП является локально экстремальным на каждом шаге h. Если пропускная способность канала ограничена величиной H, то при плотности потока информации $I(\varepsilon_0)$, меньшей пропускной способности канала передачи, канал используется не полностью. Разность между H и $I(\varepsilon_0)$ за интервал T составляет величину $\Delta I = (H - I(\varepsilon_0))T$, которую можно рассматривать как резерв для передачи фрагмента сообщения при локальной производительности источника, большей H.

В динамике канал передачи не будет перегружен, если выполняется соотношение

$$\int_t^{t+T} I(\varepsilon_0, t)\, dt \le H,$$

где $I(\varepsilon_0, t)$ – мгновенная производительность источника информации с кодированным сообщением, усредненная по длине шага.

Оптимальным сжатием сообщения с переменными параметрами класса будет такое, для которого при обработке участка сообщения с $I(\varepsilon_0, t) > H$ передача сообщения с погрешностью не более ε_0 продолжается как можно дольше

$$t_H \to max, \qquad (1)$$

где t_H – момент времени, при котором запас пропускной способности канала связи исчерпан.

Из (1) следует стратегия поведения метода сжатия: для того, чтобы на интервале, где $I(\varepsilon_0, t) > H$, перегрузка канала наступила как можно позже (или, может быть, не наступила совсем), надо передавать на каждом шаге минимальный объем информации, необходимый для восстановления сигнала с погрешностью не более ε_0.

Итак, на вопрос, как оптимально сжать и передать сигнал класса $W^r_\infty([a,b], C^r)$ длиной Δ, метод МАП дает такой ответ: надо сжимать его с погрешностью ε_0 на каждом шаге h и передавать необходимый минимум информации, тогда получим сжатие, оптимальное в смысле (1).

[1] Моругина Т.В. Компьютерные технологии в науке, проектировании и производстве. Тезисы докладов I Всероссийской научно–технической конференции, Н.Новгород, 1999, с.12.

МОДЕЛИРОВАНИЕ ВОЗРАСТНОЙ ДИНАМИКИ НАСЕЛЕНИЯ

А.Т.Надеев

Волго–Вятская академия государственной службы

Если на момент рождения некоторого поколения известно возрастное распределение населения, то начальную численность указанного поколения можно выразить следующим образом

$$\rho(t-\tau,0)=\int_0^{\infty}\alpha\left(t-\tau,\tau'\right)\rho\left(t-\tau,\tau'\right)\tau', \tag{1}$$

где $\rho(t-\tau,0)d\tau$ – начальная численность поколения, имеющего в момент времени t возраст τ, $\alpha(t-\tau,\tau')$ – средний темп рождаемости в момент времени $t-\tau$ в расчете на одного жителя, имеющего возраст τ, $\rho(t-\tau,\tau')d\tau$ – численность поколения, имеющего в момент времени $t-\tau$ возраст τ'.

Учитывая, что скорость сокращения населения в связи со смертностью пропорциональна его численности, получим

$$\rho(t,\tau)=(t-\tau,0)\exp\left[-\int_{t-\tau}^{t}\mu\left(t',t'+\tau-t\right)dt'\right], \tag{2}$$

где $\mu(p,q)$ – средний темп смертности в возрастной группе, имеющий в момент времени p возраст q. Таким образом, система интегральных уравнений возрастной динамики однородных (т.е. без стратификации по половым, социальным и иным признакам) поколений будет выглядеть так

$$\left.\begin{aligned}\rho(t,\tau)&=\rho(t-\tau,0)\exp\left[-\int_{t-\tau}^{t}\mu\left(t',t'+\tau-t\right)dt'\right]\\ \rho(t-\tau,0)&=\int_0^{\infty}\alpha\left(t-\tau,\tau'\right)\rho\left(t-\tau,\tau'\right)d\tau'\end{aligned}\right\}. \tag{3}$$

Первое уравнение системы (3) при известных функциях α и μ и заданных начальных значениях плотности возрастного распределения $\rho(t-\tau,0)$ позволяет прослеживать эволюцию возрастной плотности поколения с течением исторического времени.

Определим теперь взаимосвязь между интегральной и дифференциальной формами уравнений демографической динамики. Для этого определим производные по времени и по возрасту от плотности возрастного распределения. В результате получим

$$\frac{d\rho(t,\tau)}{dt}+\frac{\partial\rho(t,\tau)}{\partial\tau}=-\mu(t,\tau)\rho(t,\tau).$$

Учитывая начальные и граничные условия, приходим к следующей дифференциальной форме уравнений демографической динамики

$$\left.\begin{aligned} &\frac{d\rho(t,\tau)}{dt}+\frac{\partial\rho(t,\tau)}{\partial\tau}=-\mu(t,\tau)\rho(t,\tau)\\ &\rho(t,0)=\int_0^\infty \alpha(t,\tau')\rho(t,\tau',)d\tau',\\ &\rho(0,\tau)=\rho_0(\tau) \end{aligned}\right\}. \qquad (4)$$

Задавая начальные и краевые условия при известных функциях $\alpha(t,\tau)$ и $\mu(t,\tau)$ можно определить возрастные плотности населения $\rho(t,\tau)$ для заданных моментов времени t и возрастов τ.

$$\rho(t,\tau)=\rho(0,\tau-t)\exp\left[-\int_{\tau-t}^{\tau}\mu(t-\tau+\tau',\tau',)d\tau\right],\tau\geq t \qquad (5)$$

Величины $\rho(0,\tau-t)$ являются начальными условиями и поэтому, согласно формулировке (4) они предполагаются известными.

Для $\tau<t$ решение находится через краевые условия

$$\rho(t,\tau)=\left[\int_0^\infty \alpha(t-\tau,\tau'')\rho(t-\tau,\tau'',)d\tau',\right]\exp\left[-\int_0^{\tau}\mu(t-\tau+\tau',\tau',)d\tau',\right],\tau<t\ . \qquad (6)$$

Системы (3) и (4) могут быть легко распространены и на случай учета миграции населения. При наличии относительно слабых возмущений величины $\rho(t,\tau)$ (например, за счет незначительных колебаний рождаемости) возможно возникновение, как правило, затухающих демографических волн.

Проходя через интервал репродуктивных возрастов, исходная демографическая волна за счет колебания уровня рождаемости порождает вторичные (а далее третичные и т.д.) волны. Однако последующие волны в силу размытости функции $\alpha(t,\tau)$ по возрастам будут иметь все меньшую и меньшую амплитуду. Волновой пакет при этом будет все более и более расплываться по координате τ. Эта эргодичность связана не столько со случайными колебаниями темпов рождаемости и смертности, сколько с вполне детерминированным преобразованием демографической волны в поток рождений. Такое преобразование за счет значительного интервала репродуктивных возрастов растянуто во времени, что и приводит к постепенному гашению демографических возмущений. Более существенными с точки зрения структурной являются изменения, возникающие в возрастных распределениях, обусловленные периодическими возмущениями. Очевидно, что периодические возмущения относительно малых периодов (сезонные) практически в силу сравнительно большого интервала репродуктивных возрастов (15 – 20 лет) гасят друг друга, вызывая на демографической кривой лишь некоторую «рябь». Периодические возмущения, близкие по своим временным параметрам к указанному интервалу, могут привести к возникновению резонансных явлений и даже породить хаотические движения. Подобные движения для популяционных моделей, оперирующих интегральными численностями, уже давно и хорошо изучены.

СТАТИСТИЧЕСКИЙ АНАЛИЗ СЕТЕВЫХ ПОТОКОВ

Г.А.Салагацкий, А.А.Рябов, С.Н.Иванов

Нижегородский госуниверситет

Развитие информационных технологий ведет к появлению новых областей человеческой деятельности связанных с информационными сетями и проникновению сетевых технологий в ранее не связанные с ними сферы. Все более широкое распространение получают системы электронного документооборота, сетевого управления и контроля технологическими процессами, сетевая телефония, сетевое теле- и радиовещание, видеоконференции.

Для обеспечения рабочих мест персонала современными информационными технологиями создаются развитые инфраструктуры корпоративных сетей. Многие современных сетевых приложений требуют передачи большого объема трафика и чувствительны к временной задержке, а мультимедийные приложения и к дисперсии задержки. Далеко не каждая сеть удовлетворяет совокупности перечисленных требований, особенно если была создана несколько лет назад. В этом случае возникает необходимость в методике диагностики сети, позволяющей эффективно обнаруживать причины ее низкой работоспособности. Примерная методика диагностики и возможные причины неэффективной работы сети приведены в [1],[2].

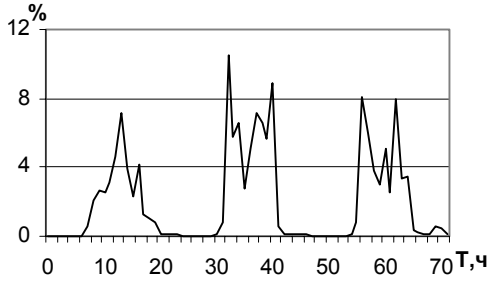


Рис. 1

Большинство средств мониторинга и диагностики состояния сети, встроенных в современные сетевые устройства (концентраторы, коммутаторы, маршрутизаторы), а также распространенные программы мониторинга и анализаторы сетевого трафика позволяют получить интегральные характеристики сети, такие как общий объем трафика и количество пакетов за интервал времени, распределение трафика по протоколам.

Исследование работы множества локальных сетей показало - интегральные характеристики не всегда адекватно отражают реальную ситуацию, сложившуюся в сети, что может привести к неверной оценке работоспособности сети.

Чтобы оценить эффективность работы сети и определить причины низкой работоспособности можно использовать методы статистического анализа сетевых потоков, собранных при помощи программ анализаторов сетевого трафика.

В ходе исследований проведен поиск причин большой задержки доставки данных в реальной сети, образующей один коммутируемый сегмент Ethernet более чем из трехсот рабочих станций.

Для анализа работы сети были построены распределения длин кадров и интервалов между кадрами.

На рис.1 приведена кривая общей загруженности сети за трое суток имеющая ярко выраженный циклический характер. Каждая точка соответствует средней загруженности канала сети за час работы в процентах. Из нее видно, что загрузка сети не превышает 10 процентов. Однако время передачи сообщений в сети оказывается аномально большим, то есть сеть не справляется со своими функциями. На рис.2 приведено распределение длин кадров в сети. Также было получено распределение длительности интервалов между кадрами.

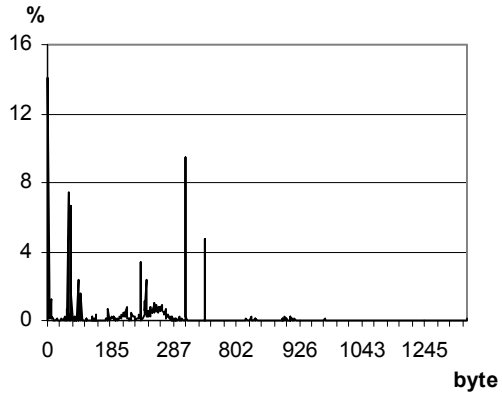


Рис. 2

Из распределения длин кадров видно, что большая часть информации по сети передается короткими кадрами, что ведет к резкому снижению эффективной полосы пропускания канала сети Ethernet. Росту числа коротких кадров способствовали некорректные правила сетевой коммутации.

Распределение длин интервалов между кадрами показывает, что значительная часть кадров имеет задержку в 50 и более раз превышающую минимально допустимую. Что говорит о низкой эффективности работы сети при низкой загруженности канала передачи данных.

Для исследования эффективности работы сетей Ethernet создано программное обеспечение, позволяющее строить статистические распределения длины кадра и длительности интервала между кадрами на основе протокола работы сети полученного с помощью сетевого анализатора.

[1] Семенов Ю.А. Сети Интернет. Архитектура и протоколы. –М: Сиринъ, 1998, с.424.

[2] Борисенко В., Подлазов О., Юдицкий С. Искусство диагностики локальных сетей //LAN 1998. №07-08.

ПАРЦИАЛЬНАЯ СИНХРОНИЗАЦИЯ СВЯЗАННЫХ ПСП ГЕНЕРАТОРОВ

К.Г.Кирьянов, В.Д.Шалфеев

Нижегородский госуниверситет

Исключительные свойства и экономичность генераторов ПСП определили весьма широкие области их применения к таким задачам как кодирование [1] и криптография [1-4]. Ранее в работах других авторов (см., например, работу [1] и обширную библиографию к ней) свойства синхронизации ГПСП не рассматривались. В [4] предложена математическая модель (ММ) сети связи, приспособленная к некоторым стандартным методам шифрования и кодирования дискретной информации, использующая «паразитные» периодические режимы *цикловой синхронизации* в сетях связанных генераторов сложных псевдослучайных последовательностей (ГПСП) в структуре GF(q). В технике чаще используется q = 2. При этом «ключами», «управляющими таблицами соответствия» между словами – состояниями связанных парциальных генераторов «передатчика» и «приемника» могли быть начальные состояния, параметры и связи между ГПСП. Метод не требовал наличия специальных синхросигналов для формирования слов и расстановки знаков препинания за исключением тактовых.

В настоящей работе для ММ сети условия парциальной синхронизации ГПСП находятся в обозначениях [2-5] с помощью программных систем в форме понятного для специалиста (пример см. в таблице) «протокола-отчёта»:

1) путём D-разбиения [5] пространство параметров (ПП) P ГПСП на части $\Omega_i(\omega_i^1,\omega_i^2,s)$ с качественно различным поведением графа переходов в дискретном фазовом пространстве (ФП) (числами состояний предельного цикла (ω_i^1) области притяжения (ω_i^2) к нему, символом (s) наличия синхронизации – совпадения соответствующих компонент состояния в цикле);
2) приведением системной матрицы сети M к каноническому виду [1] с последующей факторизацией характеристического уравнения $\chi_M(X; P)$ на множители, соответствующие подматрицам диагональных блоков парциальных ГПСП сети и недиагональным подматрицам Связей матрицы M.

$$M = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

В таблице приведены результаты анализа возможности синхронизации в простой (из-за размера таблиц) сети с системной матрицей M из 2-х ГПСП размера n×n (n=3) в «1-й естественной форме» с вектором перебираемых меченых знаком (∗) элементов {m_{21},m_{54} (параметров парциальных ГПСП), m_{24}, m_{51}(параметров матриц связей)}. Видно, что не тривиальная ($\omega_i^1>1$) синхронизация с периодом $\omega_i^1 = 2^3 - 1 = 7$ имеет место в примере для четырёх вектор-параметров P={m_{21},m_{54},m_{24},m_{51}}: (1100), (0110), (1001) и (0011). Характеристические уравнения $\chi_M(X;1100) = (1 + X + X^3)^2 = (1 + X^2 + X^6)$, $\chi_M(X;0110) = X^3\cdot(1 + X + X^3)$, $\chi_M(X;1001) = X^3\cdot(1 + X + X^3)$, $\chi_M(X;0011) = (1+X^2+X^6)$ относятся, соответственно, к

1) не связанным, 2) связанным «сносовой» (односторонней – со второго на первый), 3) связанным сносовой (односторонней – с первого на второй) связями и 4) двухсторонними связями ГПСП. Случай 1) вырожденный имеет место из-за наличия в ММ тактовой синхронизации. В случаях 2) и 3) цикловая синхронизация происходит за $n \le 3$ такта из-за факторизации $\chi_M(X;0110)=\chi_M(X;1001)= X^3\cdot\chi_{ГПСП}(X)$, а в 4) – при попадании на нужный цикл – за 0 тактов, т.к. $\omega_i^1 \equiv \omega_i^2$.

ФП	D–разбиение ПП (набор меченых параметров P в M)	Число (C) циклов ФП	Структуры несвязных графов ФП: $\Omega_i(\omega_i^1,\omega_i^2,s)$ (i= 1,2,…,c) с метками синхронизации
1	[0000]		1,64,s;
2	[1000]	2	1,8,s;7,56;
3	[0100]	2	1,8,s;7,56;
4	[**1100**]	10	1,1,s;7,7;7,7;**7,7,s**;7,7;7,7;7,7;7,7;7,7;7,7;7,7;
5	[0010]	1	1,64,s;
6	[1010]	2	1,8,s;7,56;
7	[**0110**]	2	1,8,s;**7,56,s**;
8	[1110]	6	1,1,s;7,7;14,14;14,14;14,14;14,14;
9	[0001]	1	1,64,s;
10	[**1001**]	2	1,8,s;**7,56,s**;
11	[0101]	2	1,8,s; 7,56;
12	[1101]	6	1,1,s;14,14;14,14;14,14;7,7;14,14;
13	[**0011**]	6	1,1,s;14,14;14,14;14,14;**7,7,s**;14,14;
14	[1011]	8	1,1,s;15,15;15,15;5,5;15,15;5,5;5,5;3,3;
15	[0111]	8	1,1,s;15,15;15,15;5,5;15,15;5,5;5,5;3,3;
16	[1111]	1	1,64,s;

Работа выполнена при поддержке РФФИ (грант 02-02-17573).

[1] Гилл А. Линейные последовательностные машины. Анализ, синтез и применение. –М.: Наука,1974.

[2] Кирьянов К.Г., Меднов А.С., Акулов В.В. Синхронизация генераторов псевдослучайных последовательностей // Техника средств связи. «ЭКОС» Сер. РИТ. – 1990. – Вып. I. с.56.

[3] Bagrov S.N., Kirjanov K.G., Shalfeev V.D. Complicated Regimes. Synchronization and Structures in Networks of the Pseudo-Random Generators /Dynamic and Stochastic Wave Phenomena. Abstracts of the Second International Scientific School-Seminar. Nizhny Novgorod University Press. N-Novgorod, 1994.

[4] Кирьянов К.Г. Вестник ВВО Академии технологических наук Российской Федерации. Серия: Высокие технологии в радиоэлектронике. –Н.Новгород, 1995. – вып.1, с.95.

[5] Кирьянов К.Г. //Третья научная конференция по радиофизике 7 мая 1999г. – Н.Новгород: Изд-во ННГУ, 1999. с.130.

ИДЕНТИФИКАЦИЯ И ВОССТАНОВЛЕНИЕ АЛГОРИТМОВ ПРОГРАММ

Д.Л.Туренко, К.Г.Кирьянов

Нижегородский госуниверситет

Проблема анализа безопасности программного обеспечения для ЭВМ тесно связана с необходимостью исследования исполняемого кода с целью восстановления алгоритма работы программы в целом или отдельных ее процедур. Классические методы исследования исполняемого кода – дизассемблирование и отладка, требуют большого количества трудозатрат, поэтому не всегда приемлемы из-за больших объемов ПО в современных вычислительных системах.

Предлагаемый подход к автоматизированному анализу основан на построении орграфа вызовов и переходов по последовательности машинных команд исполняемого кода. В качестве исследуемой аппаратно-программной среды выбрана наиболее распространенная платформа ПЭВМ с процессором Intel x86 под управлением операционных систем семейства Microsoft Windows 9x/NT.

Во множестве всех команд процессора *A* выделяется подмножество команд *P*, изменяющих последовательный порядок выполнения кода программы. В подмножество *P* входят команды условного *(je, jng, ...)* и безусловного *(jmp)* перехода, команды вызова *(call)* и возврата из процедур *(ret)*, команды цикла *(loop)*.

Множество вершин орграфа составляют команды подмножества *P* и точки входа в процедуры и функции программы. Вершина орграфа представляется в виде пары: *(addr, info)*, где *addr* – виртуальный адрес в адресном пространства процесса в среде Windows, *info* – информация о данной точке исполняемого кода (инструкция процессора, описание точки входа в процедуру). Ребра орграфа соединяют вершины в соответствии порядком выполнения программы.

Построение орграфа начинается с адреса точки входа исполняемого кода и выполняется в несколько итераций. Сначала строится корневая ветвь – ветвь первого уровня. Во множество вершин (изначально пустое) добавляется точка входа. Далее, просматривается последовательность команд в соответствии с порядком их выполнения. Во множество вершин добавляются команды из подмножества *P* и точки, в которые выполняется переход в соответствии с данными командами, а в набор ребер – соответствующие данным переходам ребра. Ответвления от корня, связанные с вызовами процедур и функций (*call*) во время первой итерации не строятся.

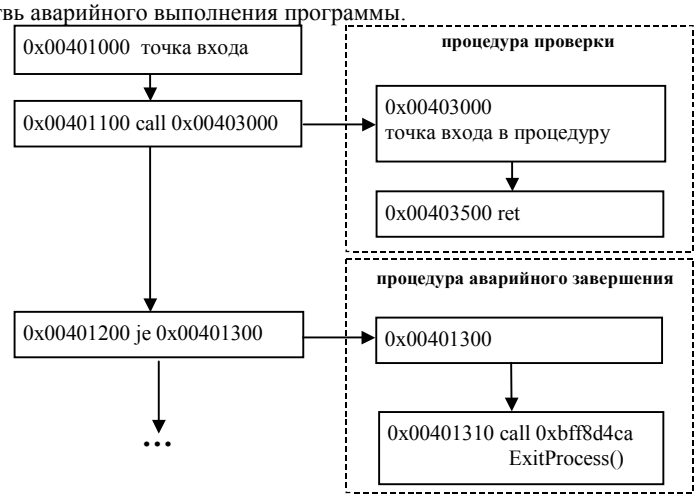
На втором шаге детализируются вызовы процедур из корневой ветви – ветви второго уровня. На третьем шаге – ветви третьего уровня, и так далее. Для построения ветвей орграфа, не связанных с корневой ветвью (функции дополнительных потоков и обработчиков системных событий), требуется найти соответствующие точки входа, которые определяются по специфической последовательности команд. Построение данных ветвей выполняется аналогично корневой ветви.

В итоге, структура программы представляется в виде орграфа, а исследование программы сводится к анализу формального математического объекта с использованием алгоритмов теории графов. Анализ построенного орграфа позволяет полу-

чить компоненты связности, циклы, тупики, степени вершин. Полученные характеристики орграфа используются для формирования критериев, с помощью которых выполняется идентификация алгоритма программы.

В зависимости от цели анализа, может потребоваться подробное построение не всего орграфа, а отдельных его ветвей. Интересующие участки кода исследуются с использованием информации о параметрах функций и структурах данных, полученной из SDK [1].

В качестве иллюстрации описанного подхода на рисунке приведен фрагмент орграфа, в котором выполняется инициализация программы. За счет исключения команд, не входящих в подмножество *P*, листинг значительно сокращается, что существенно облегчает восстановление алгоритма работы программы. В данном примере в зависимости от результатов процедуры проверки параметров программы (или других условий) происходит аварийное завершение (вызов системной функции ExitProcess) или выполняется последовательность действий в нормальном режиме. Таким образом, с помощью описанного метода выявлена процедура проверки и ветвь аварийного выполнения программы.



Построение орграфа и его анализ выполняется по формальному алгоритму, который реализуем в виде программы. Такие программные средства могут существенно сократить трудозатраты на исследование исполняемого кода.

[1] Microsoft Developer Network Library. Win32 Software Developer Kit.

СПЕЦИФИКА ПРИМЕНЕНИЯ СКЗИ ДЛЯ ЗАЩИТЫ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Д.Л.Туренко, Т.А.Николаева, И.А.Кочетков

Нижегородский госуниверситет

Программное обеспечение (ПО) в современных вычислительных системах представляет собой ценный объект интеллектуальной собственности, и поэтому требует эффективных методов защиты. Рассмотрим основные аспекты применения криптографических методов защиты ПО.

Задача 1. Защита от несанкционированной модификации, контроль целостности ПО. Традиционный метод решения данной задачи – вычисление и проверка контрольных сумм и, как их разновидность, ЭЦП (электронная цифровая подпись). Во время начальной инициализации вычислительной системы контролируемые компоненты проверяются по списку верификации, и делается заключение о наличии или отсутствии изменений. Очевидным недостатком данного подхода является невозможность проверки целостности программ непосредственно перед их запуском.

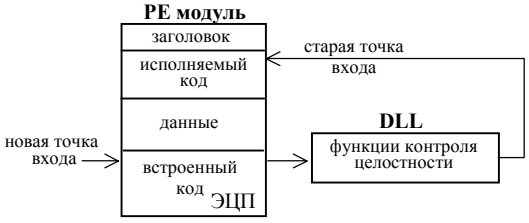


Рис. 1

В рамках исследования специфики применения СКЗИ для защиты программ была реализована следующая система “иммунизации” исполняемых файлов PE программного интерфейса Win32 (см. рис. 1). В исполняемый модуль встраивается ЭЦП и дополнительный код, получающий управление при запуске. Встроенный код обращается к библиотеке динамической компоновки (DLL), реализующей функции проверки ЭЦП. В случае не соответствия ЭЦП выдается сообщение, и запуск программы блокируется. Таким образом, несанкционированные изменения, например, со стороны компьютерных вирусов будут выявлены в момент запуска программы.

Задача 2. Защита ПО от несанкционированного запуска, разграничение доступа. В случае, когда штатные средства операционной системы (ОС) не позволяют организовать разграничение доступа к различным пакетам системного и прикладного ПО (например, в семействе ОС Microsoft Windows 9x), описанная выше система иммунизации позволит решить данную задачу. Для этого исполняемый код или другие секции программы зашифровываются, и встраивается код, осуществляющий запрос ключевой информации, необходимой для преобразования программы в рабочее состояние.

Задача 3. Защита ПО от нелицензионного использования. Программный продукт (ПП) должен предоставлять услуги (функции) пользователю только при вы-

полнении определенных условий. Как правило, это правильно введенный серийный (регистрационный) номер или наличие электронного ключа (HASP).

Использование электронных ключей накладывает очевидные ограничения на область распространения ПП и оправдано в случае защиты сложных и дорогих ПП. Реализация данного метода под силу только крупным фирмам, т.к. требует организации дилерской сети.

Использование регистрационных кодов (даже весьма сложных) не обеспечивает необходимой стойкости защиты от действий хакеров. В связи с этим, для выработки регистрационных кодов предлагается использовать стойкие криптографические алгоритмы, в частности, ЭЦП. С помощью ЭЦП можно привязать ПП к конкретной аппаратно-программной среде, что сделает копирование ПП в другую среду бессмысленным.

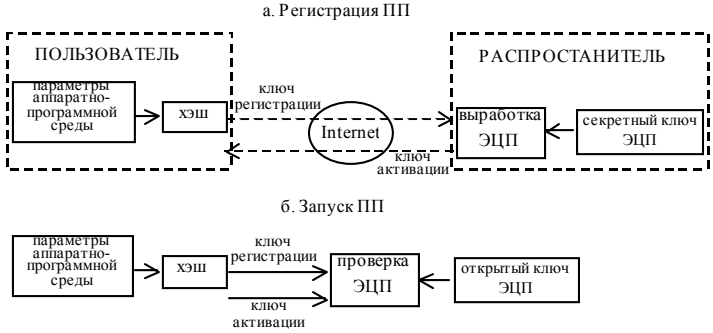


Рис. 2

На основе серийного номера ПП, информации об уникальных параметрах системы и, возможно, другой информации вырабатывается ключ регистрации и отправляется распространителю ПП, который в свою очередь вырабатывает ключ активации (ЭЦП) и отправляет его пользователю (см. Рис. 2а). При запуске ПП проверяется целостность его компонентов, вырабатывается ключ регистрации и проверяется на соответствие ключу активации (см. Рис. 2б). В случае не выполнения проверяемых условий функции ПП блокируются.

Выработка ключа активации выполняется с использованием секретного ключа ЭЦП, которым владеет распространитель ПП. Невозможность подделать ключ активации обеспечивается стойкостью алгоритма ЭЦП. А попытки отключить проверки будут выявлены на этапе контроля целостности ПП. Применение данного подхода наиболее эффективно при распространении ПО через сеть Интернет, особенно, если использование дорогих средств защиты (таких, как электронные ключи и не копируемые машинные носители) не оправдано.

Этот очевидный подход, как ни странно, не получил широкого применения. Лишь недавно, корпорация Майкрософт предложила технологию MPA для защиты своих продуктов серии XP. И тем не менее, появление пиратских копий продуктов XP говорит о том, что реализация данной технологии оставляет желать лучшего.

МОДЕЛЬ МУЛЬТИПЛЕКСИРОВАНИЯ ИСТОЧНИКОВ В ТРАКТАХ АТМ

В.Н.Богданов, П.С.Вихлянцев, М.В.Симонов

Дочернее ГУП НТЦ «Атлас-Северо-Запад»

В настоящее время на мировом рынке телекоммуникационных услуг сформировался сектор интегральных услуг связи, в котором пользователи получают весь комплекс услуг. Расширение этого сектора потребовало создания транспортных сетей с высокой пропускной способностью, использующих технологию пакетной коммутации, которая обеспечивает интеграцию передачи речи, данных, факсимильных сообщений, мультимедиа, то есть комплексного предоставления информации. Применение технологии АТМ, когда каждый источник получает от сети только тот ресурс пропускной способности, который ему нужен, дает возможность использовать выгоды статистического мультиплексирования для повышения эффективности использования цифровых трактов связи. В большинстве случаев функционирование пользователя с изменяющейся скоростью передачи (ИСП) можно аппроксимировать источником типа «вкл-выкл», то есть, с вероятностью $p(k)$ скорость передачи является пиковой $B_p(k)$, а с вероятностью $q(k)=1-p(k)$ скорость источника $B_{min}(k)=0$ [1]. В этом случае среднее значение и дисперсия скорости передачи абонента k-ой службы типа «вкл-выкл» могут быть выражены через параметры трафика:

$B_m^{(k)}=p^{(k)}B_p^{(k)}$, $p^{(k)}=1/k_b^{(k)}$, где $k_b^{(k)}$ – коэффициент пачечности,

$$D\left[b^{(k)}\right]=\begin{cases} p^{(k)}q^{(k)}\left(B_p^{(k)}\right)^2, \text{ для источника с изменяющейся скоростью передачи,} \\ 0, \text{ для источника с постоянной скоростью передачи.} \end{cases} \quad (1$$

Будем оценивать эффективность метода статистического мультиплексирования отношением количества источников N_{BC}, пачечный трафик которых мультиплексируется с заданным качеством $P_{PLR} \le P_{доп}$ в цифровом тракте с пропускной способность $B_{тр}^{(C)}$ к числу каналов N_{KK} , которое должно быть образовано в тракте методом коммутации каналов $G = N_{BC}/N_{KK}$ при $N_{KK} = B_{тр}^{(C)}/B_P^{(C)}$.

В этом случае можно получить искомое выражение для оценки эффективности метода статистического мультиплексирования источников с пачечным трафиком:

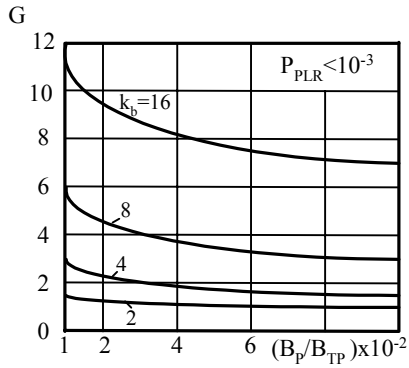
$$G = k_b^{(k)} - \frac{k_b^{(k)}-1}{N_{кк}^{(k)}} \cdot \frac{u^2}{2}\left[\sqrt{1+4\frac{N_{кк}^{(k)}k_b^{(k)}}{u^2(k_b^{(k)}-1)}}-1\right],$$

где u – аргумент, зависящий от допустимого значения вероятности потери пакетов.

Рисунок показывает, что статистическое мультиплексирование дает значительный выигрыш тогда и только тогда, когда скорость любого источника значительно ниже скорости цифрового тракта. При этом, чем выше коэффициент пачечности источника, тем выше эффективность статистического мультиплексирования.

Однако, в реальных концентраторах, мультиплексорах или коммутаторах доступа поступают потоки от разнородных по скорости источников, предъявляющие различные требования к качеству обслуживания, прежде всего, по вероятности потери пакета.

Для источников i–ой службы количество виртуальных соединений, которое можно организовать в тракте с учетом виртуальных соединений для других служб, может быть вычислено по формуле:



$$N_{вс}^{(i)} = \frac{B_{TP}^{(C)} - \sum_{\substack{k=1 \\ k\neq i}}^{K} N_{BC}^{(k)} p^{(k)} B_p^{(k)}}{p^{(i)} B_p^{(i)}} - \frac{\left[u^{(i)}\right]^2}{2} \frac{1-p^{(i)}}{p^{(i)}} \times$$

$$\times \left\{ \sqrt{1 + 4 \frac{B_{TP}^{(C)} - \sum_{\substack{k=1 \\ k\neq i}}^{K} N_{BC}^{(k)} p^{(k)} B_p^{(k)}}{\left[u^{(i)}\right]^2 \left(1-p^{(i)}\right) B_p^{(i)}} + 4 \frac{\sum_{\substack{k=1 \\ k\neq i}}^{K} \left[u^{(k)}\right]^2 p^{(k)} \left(1-p^{(k)}\right) \left[B_p^{(k)}\right]^2}{\left[u^{(i)}\right]^4 \left(1-p^{(i)}\right)^2 \left[B_p^{(i)}\right]^2}} - 1 \right\}$$

Приведенная математическая модель статистического мультиплексирования источников в сетях АТМ и проведенные с ее помощью расчеты убедительно свидетельствуют в пользу метода коммутации пакетов по сравнению с традиционным методом коммутации каналов. Вместо стандартных и многочисленных вторичных сетей телефонной, телеграфной, факсимильной связи, сетей передачи данных и т. п., каждая из которых рассчитана только на обеспечение одного вида связи с тем или иным способом переноса информации, строится единая цифровая сеть на базе широкого использования волоконно-оптических линий связи и единого метода транспортирования всех видов информации с помощью технологии асинхронного режима переноса пакетов фиксированной длины (ячеек).

[1] Назаров А.Н., Симонов М.В. АТМ: Технология высокоскоростных сетей. –М.: Эко-Трендз,1997. 232с.

ИСПОЛЬЗОВАНИЕ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ ДЛЯ ЗАЩИТЫ ДОКУМЕНТОВ ОТ ФАЛЬСИФИКАЦИИ

В.Н. Богданов, П.С. Вихлянцев, М.В.Симонов

Дочернее ГУП НТЦ «Атлас-Северо-Запад»

В целях защиты документов и ценных бумаг от фальсификации в настоящее время используются специальные методы печати, специальные краски с особыми оптическими и магнитными свойствами, голограммы и т.п. Однако сегодня злоумышленник может иметь все возможные компоненты и оборудование для производства поддельных документов.

Широко применяемая система нотариального заверения документов также не обеспечивает абсолютной надежности. Нотариус гарантирует достоверность информации, представленной в заверяемом документе только на момент проверки, и не исключает его фальсификацию в последующем. Кроме того, вне нотариальной конторы верификация нотариально заверенного документа не надежна, так как сводится только к визуальному контролю штампа и/или печати нотариуса, которые также относительно легко могут быть фальсифицированы.

Для защиты информации от преднамеренных или случайных искажений в электронном документообороте находит широкое использование электронная цифровая подпись. Однако, электронная цифровая подпись может также применяться для защиты бумажных документов и товаров от фальсификации.

В НТЦ «Атлас» разработан способ защиты алкогольной продукции от фальсификации [1]. Для защиты от подделки используется аппарат теории информации и шифрования. На каждую единицу алкогольной продукции наносится региональная специальная марка, имеющая поле для записи контрольно-учетной информации в виде двухмерного штрихового кода и в буквенно-цифровой форме. Контрольно-учетная информация может подписываться электронной цифровой подписью (ЭЦП), которая вместе с контрольно-учетной информацией наносится на марку в штриховом коде. Электронная цифровая подпись формируется по стандартному алгоритму с использованием секретного ключа маркировщика, является уникальной и гарантирует целостность контрольно-учетной информации от подделки. Попытки изменить хотя бы один бит в контрольно-учетной информации или в ЭЦП будут выявлены при проверке.

В настоящее время документы исполняются все чаще с использованием компьютеров. Подготовленный в текстовом редакторе документ печатают на принтере. Это дает возможность применить программные средства криптографической защиты информации, в том числе с использованием алгоритма формирования и проверки ЭЦП для защиты информации, наносимой на бумажный носитель [2,3].

Количество выделяемых контрольных фрагментов и их объем (размер) зависят от типа и объема самого документа, а также от важности отображенной в нем информации и требуемой степени ее защиты. Так, в предельном случае при защите всего документа контрольный фрагмент содержит всю информацию.

При многостраничном документе маркироваться может каждая страница объемом до 1500 – 2000 знаков. При более низких требованиях к степени защиты количество и объем каждого из выделяемых контрольных фрагментов может быть сокращено.

Выделенная контрольная информация в цифровой форме подписывается ЭЦП с использованием секретного ключа маркировщика, в качестве которого выступает лицо, уполномоченное подписывать или заверять защищаемый документ. Контрольная информация и соответствующая ей ЭЦП преобразуются в двумерный штриховой код. В частном случае, для упрощения приборной верификации документа, в состав штрихового кода включается открытый ключ маркировщика.

Печать документа осуществляется на принтере с одновременным нанесением защитной маркировки в виде штрихового кода. Защитная маркировка может также наноситься на специальный защитный знак, который размещается на защищаемом документе. При этом специальный защитный знак должен иметь комплексную защиту от подделки, в том числе: голографический защитный элемент; контрольно-учетную информацию в виде штрихового кода и буквенно-цифровой форме; специальную метку, выполненную бесцветной специальной краской, светящуюся при ИК облучении; графические элементы; просечки, служащие для защиты от переклеивания, оригинальную форму.

Промаркированный документ поступает пользователю, вводится в хозяйственное или финансовое обращение и т.п.

В подсистеме верификации с промаркированного документа осуществляется считывание штрихового кода и его преобразование, проверка подлинности ЭЦП и отображение контрольной информации на дисплее или путем печати на принтере для визуального сравнения с информацией, нанесенной на документе в обычной, буквенно-цифровой форме.

Если ЭЦП подлинная, то осуществляется визуальное сравнение выделенной из штрихового кода контрольной информации с информацией, изображенной в документе в обычной буквенно-цифровой форме. Совпадение контрольной информации с информацией, отображенной в документе, гарантирует его подлинность.

Использование секретного ключа при формировании ЭЦП и соответствующего ему открытого ключа при проверке ЭЦП однозначно подтверждает маркировщика, т.е. авторство исполнителя документа или лица подписавшего (заверившего) документ.

[1] Богданов В.Н. и др. Способ идентификации подлинности контролируемого объекта. Патент РФ №2172015. Опубл. 20.08.2001, БИ №22.

[2] Богданов В.Н. и др. Способ подтверждения подлинности информации. Патент РФ №2165643. Опубл. 20.04.2001, БИ №11.

[3] Богданов В.Н. и др. Система защитной маркировки и верификации документов. Свидетельство на полезную модель №19944. Опубл. 10.10.2001, БИ №28.

ОБ ОСОБЕННОСТЯХ ПРОГРАММНОЙ РЕАЛИЗАЦИИ ДИХОТОМИИ ВЕКТОРНЫХ ДАННЫХ

А.Ю.Виценко

Нижегородский госуниверситет

При решении прикладных задач часто возникает подзадача нахождения элемента массива. Традиционно для решения этой задачи используются алгоритмы «быстрого» поиска: дихотомия или бинарный поиск, «Фибоначчиев» поиск, интерполяционный поиск и другие [1]. В данной работе рассматривается бинарный поиск.

Основная сложность применения алгоритмов быстрого поиска к векторным данным или данным более сложной структуры состоит в необходимости упорядочивания исходного массива по возрастанию или убыванию. В тоже время, для рассматриваемых данных операции больше и меньше не определены. Наиболее часто для решения проблемы упорядочивания используется некоторая функция, обеспечивающая однозначное отображение элементов исходного массива в альтернативный массив другого пространства (хэширование), для элементов которого отношения больше и меньше определены. Основной недостаток этого подхода связан с затратами на получение однозначной хэш-функции.

В качестве альтернативного подхода предлагается использовать легко вычисляемую функцию, в общем случае не однозначную, осуществляющую «скаляризацию» данных. Предлагаемый алгоритм поиска также требует предварительного упорядочивания данных и состоит из следующих этапов:

- поиск методом дихотомии в альтернативном массиве;
- поиск методом последовательного перебора в исходном массиве в окрестности элемента, найденного на первом этапе.

Рассмотрим подробнее каждый из перечисленных этапов.

До начала поиска на этапе упорядочивания удобно использовать уже существующие стандартные подпрограммы сортировки, например *qsort*. Функция *qsort* определена в текущем стандарте ANSI C/C++[2,3], что обеспечивает переносимость программной реализации, и позволяет использовать собственную функцию сравнения. Использование внешней функции для сравнения элементов массива несколько снижает быстродействие программной реализации, но в тоже время позволяет легко адаптировать предлагаемый алгоритм к произвольным типам данных.

На этапе дихотомии также удобно использовать уже существующие стандартные подпрограммы бинарного поиска, например *bsearch*. Функция *bsearch* определена в текущем стандарте ANSI C/C++[2,3]. Дополнительным удобством такого подхода является использование единой внешней функции *compar* для функций *qsort* и *bsearch*. В качестве ключевого элемента на вход функции *bsearch* передается скаляризованное значение исходных данных. При удачном завершении этого этапа (совпадение скаляризованной величины найдено) осуществляется переход к сле-

дующему этапу. В противном случае уже на этом этапе можно сделать вывод об отсутствии искомой величины в массиве.

Для реализации последнего этапа можно использовать простую функцию, реализующую последовательный поиск на интервале совпадения «скаляризованной» величины. Ниже приведен один из возможных алгоритмов реализации такой функции:

- от точки совпадения скаляризованной величины осуществляется движение в обе стороны по исходному массиву;
- условием удачного завершения поиска является совпадение исходных данных с элементом массива;
- условием неудачного завершения поиска является достижение границ массива или отклонение скаляризованного значения элемента массива от найденного на предыдущем этапе.

Для реализации предложенного алгоритма разработана функция *seq_int_search*, максимально совместимая по параметрам и возвращаемому значению с библиотечной функцией *bsearch*.

Ниже приведена структура процедуры.

```
qsort(,,,compar);
if ( bsearch(F_Scalar(key),,,compar) )
    if ( seq_int_search(key,F_Scalar(key),,,compar_other)){
        Действия при нахождении искомого элемента .}
    else{
        Действия при отсутствии искомого элемента .... }
```

Шрифтом с нажимом выделены изменения которые следует внести в уже существующий поисковый алгоритм для адаптации его к векторным данным или данным сложной структуры.

Среднюю эффективность предлагаемого алгоритма по отношению к последовательному поиску упрощенно можно оценить следующим образом

$$K = (N / 2)/(C_1 \log_2(N) - 1 + C_2),$$

где: N – число элементов массива, C_1 – коэффициент, зависящий от выбранной функции скаляризации, C_2 - коэффициент зависящий от расположения элементов в массиве.

Экспериментальная проверка предложенного алгоритма показала что на различных, достаточно равномерно распределенных данных, при числе элементов массива 1 000 000 и трехмерных векторах эффективность примерно равна 500.

Таким образом, предлагаемый алгоритм может быть использован для решения задачи нахождения элемента массива сложной структуры при минимальных изменениях в уже существующем программном обеспечении.

[1] Кнут Д. Искусство программирования для ЭВМ. Т.3 Сортировка и поиск. –М: Мир, 1978, 848с.
[2] ISO/IEC 9899 Programming languages – C, ISO/IEC 9899:1999(E), с.317.

[3] ISO/IEC 14882 Programming languages – C++, ISO/IEC 14882:1998(E), c.563.

К ПОИСКУ ДИАГНОСТИЧЕСКИХ КРИТЕРИЕВ РЕАКТИВНОГО АРТРИТА

Ю.Ю.Альтерман

Первая градская клиническая больница, Нижний Новгород

Реактивный артрит – это вызванное инфекцией заболевание, характеризующееся, прежде всего воспалением синовиальной оболочки, посев с которой не выявляет наличия жизнеспособных микроорганизмов. Предрасположенность к реактивному артриту передается по наследству. Предрасположенность к реактивному артриту кодируется HLA-B27Ag. Но развитие данного заболевания четко связано с инфекцией, определенными возбудителями, поражающими чаще всего мочеполовую систему (урогенитальные инфекции) [1,2].

В последние годы проблема реактивных урогенных артритов приобретает все большую актуальность. Они имеют высокую распространенность среди реактивных артритов, а также среди всех артритов в целом. По данным венского университета, при обследовании 234 больных с олигоартритами, их урогенная природа доказана в 44% случаев. Из них 15% – хламидии, 14% – микоплазмы, 28% – уреаплазмы.

По данным ревматологического отделения 1 ГКБ Н. Новгорода за 1997 г., болезнь Рейтера (ХА) составляет 7,9%, микоплазменные артриты – 0,2%, уреаплазменные артриты – 2,9%, артриты, вызванные микст-инфекцией (хламидии + уреаплазмы) – 0,7%.

Опыт наблюдения за больными реактивным урогенным артритом показал наличие различий в клинической картине уреаплазменного (УА) и хламидийного артритов (ХА), чему и посвящена настоящая работа.

В исследование вошли больные реактивным уроненным артритом хламидийной и уреаплазменной этиологии. Диагноз базировался на наличии у больных клинических симптомов серонегативного артрита в сочетании с выявленными возбудителями (хламидии и уреаплазмы). При постановке диагноза использованы критерии Немецкого ревматологического общества 1995 года. Обследование больных на хламидии и уреаплазмы проводилось методом прямой иммунофлуоресценции.

Полученные данные обработаны способом подсчета процентного соотношения признаков, (диагностических) коэффициентов и информативности. В исследование вошли 81 больной УА и 60 – ХА. Контрольную группу составили 23 больных УА и 19 – ХА. Для ориентировочной оценки предполагаемого результата лечения и выявления, наиболее значимо влияющих на исход факторов произведен подсчет диагностических коэффициентов и диагностической ценности или информативности данных, использованных в данной работе. Две группы больных с хламидийным и уреаплазменным артритами были обследованы и сравнивались по определенным, преимущественно клиническим, критериям. При составлении диагностических таблиц руководствовались следующими предпосылками.

При прогнозировании плохих исходов возможны ошибки двух родов. Пациента, у которого имеет место хламидийный артрит можно ошибочно отнести в группу

пациентов с уреаплазменным артритом, и, наоборот. За ошибку первого рода принимаем величину $\alpha \le 0{,}01$ (т.е. 1%), а за ошибку второго рода величину $\beta \le 0{,}1$ (т.е. 10%). Формула принятия решения при последовательной процедуре распознавания в случае равенства априорных вероятностей этих двух состояний примет вид следующего неравенства (Гублер Е.В.,1978)

$$\frac{\alpha}{1-\beta} < \frac{P(x_1 \mid A_1)}{P(x_1 \mid A_2)} * \frac{P(x_2 \mid A_1)}{P(x_2 \mid A_2)} \cdots \frac{P(x_q \mid A_1)}{P(x_q \mid A_2)} < \frac{1-\alpha}{\beta},$$

где $x_1, x_2, \ldots, x_q$ – число пациентов, у которых найдены диагностические признаки 1, 2, …, q; A_1 и A_2 – общее число пациентов, входящих в группу соответственно с ХА и УА. Накопление диагностической информации продолжается, пока верно это неравенство. Оно становится неверным, когда достигнут один из порогов $\alpha/(1-\beta)$ или $(1-\alpha)/\beta$, процедура умножений отношений вероятностей диагностических признаков прерывается и выносится решение в пользу того состояния, порог которого достигнут, т.е. если достигнут порог $\alpha/(1-\beta)$ выносится решение: УА если порог $(1-\alpha)/\beta$ – ХА.

Формула принятия решения при последовательной диагностической процедуре приобретает вид следующего неравенства:

ДКпор (плохого исхода) < Σ_iДК(x_i) < ДКпор (благоприятного исхода),
где ДК=10 lg [$P(x_i|A_1)/P(x_i|A_2)$].

Так как в диапазоне обоих состояний встречаются одни и те же диагностические признаки, в диагностической таблице их располагают в порядке убывания информативности. Формула Кульбака [3] подсчета информативности в модификации Е.В. Гублера имеет следующий вид:

$$J(x_i) = \sum_i J(x_i) = \sum_i 10 \lg \frac{P(x_i \mid A_1)}{P(x_i \mid A_2)} * 0{,}5 * [P(x_i \mid A_1) - P(x_i \mid A_2)] .$$

Иными словами, информативность диагностического признака представляет собой сумму произведений диагностических коэффициентов и полуразности частот вероятности возникновения признака при состоянии A_1 и A_2, (при благоприятном и плохом исходе).

Проблема реактивных артритов требует дальнейшего изучения.

Автор благодарит д.м.н. Шафита С.Е. за помощь при обработке данных.

[1] Ревматические болезни. /под. ред. акад. Насоновой –М.: Медицина, 1997, 519с.
[2] Стерлинг Дж. Вест. Секреты ревматологии. –М.: Бином, С.-П.: Невский диалект, 1999, 832с.
[3] С.Кульбак. Теория информации и статистика. –М.: Изд-во «Наука» ГРФМЛ, 1967, 408с.

К ОБОСНОВАНИЮ ВЫБОРА КОМПОНЕНТ СОСТОЯНИЯ (ПРИЗНАКОВ) БОЛЬНОГО В СИСТЕМАХ ОПРЕДЕЛЕНИЯ «КАЧЕСТВА ЖИЗНИ»

Е.А.Грунина[1)], К.Г.Кирьянов[2)]

[1)]*Нижегородская медицинская академия* [2)]*Нижегородский госуниверситет*

Проблема выбора «информативных признаков» характеризующих состояние сложной системы любой природы зависит от целей последующего использования этих признаков и, в первую очередь, определяется степенью требуемой детализации, глубины разбиения сложной системы на части, с «точностью» до которых необходимо устанавливать различие систем. Проблема выбора «информативных признаков» в системах определения различий «качества жизни» больных людей [1,2], как и других, весьма сложных систем «живой» природы не являются исключением. В настоящей работе предлагается свести проблему обоснования выбора вида и числа нескольких ($i = 1, 2,\ldots, L$) одновременно наблюдаемых у больного компонент $\{y^i(t)\}$, векторных процессов $y(t)$, называемых в [1,2] «набором признаков», вводимых периодически в упомянутые выше системы определения «качества жизни». Выбор состава и количества L компонент на момент времени T считается лучшим, если так называемые оптимальные базовые параметры (БП) *математической модели* (ММ) «*источника*» из T [3] наборов приводят к уменьшению энтропии процесса $\{y^i(t)\}$, $i = 1, 2,\ldots, L$, $t = 1, 2,\ldots, T$ [4].

К БП относятся: Δt – интервал между съемом векторов данных $\{y^i(t)\}$ с источника-пациента, q – число уровней квантования признаков, n – «сложность» источника данных, характеризующая глубину зависимости значений соседних L-эк «признаков». В основе метода определения оптимальных БП используется принятая в строгих научных дисциплинах (теоретической и квантовой механике, физике, теории управления и т.д.) наиболее общая концептуальная ММ природных явлений «динамическая система» (ДС), обладающая наиболее мощной прогнозирующей способностью при вычислении будущего поведения фазовых траекторий по начальным или наблюдённым данным. Все другие ММ – частные случаи ДС.

Нами набор $\{y^i(t)\}$ преобразуется в $(\Delta t, q, n)$ – квантованный «текст» $\{y^i_k\}$, по которому затем (см., например, [3,4]) определяются оптимальные БП ММ ДС, соответствующие минимуму по q энтропии

$$\mathrm{E}(q) \equiv \mathrm{E}(q/\{y^i_k\}) = \log_2((q/\{y^i_k\})^{n(/\{yik\}q)}) = n(q/\{y^i_k\})\log_2 q(/\{y^i_k\}),$$

и далее по q_{onm} вычисляется $n_{onm} = n(q_{onm})$. Для наглядного сравнения наборов по энтропии нами были взяты две тестовые последовательности признаков: (в файлах h21.txt – с одного отведения с «пациента» и h1XX2211.txt – сумма трёх отведений). На рис.1 и рис.2 представлены графики зависимости условной энтропии E(q) и глубины зависимости значений соседних «признаков» $n(q)$ в интервале $2 < q < 1000$ тестовой последовательности $\{y^i(t)\}$, $i = 1, 2,\ldots, L$, $t = 1, 2,\ldots, T$, $L = 1$, $T = 230$ из 1-го файла из-за их большого объёма, по которым легко определяются базовые параметры: $q_{onm} = 220$: $n_{onm} = 2$, $\mathrm{E}(q_{onm}) = 15{,}5627$. Для второго файла получается

$E(q_{опт}) = 14{,}570844$, при $q_{опт} = 156$, $n_{опт} = 2$ и предпочтение следует отдать второму набору. Следует отметить, что для подбора информативных признаков в системах определения различий «качества жизни» может использоваться любая стратегия поиска экстремума, включая метод проб и ошибок, приводящая к уменьшению энтропии.

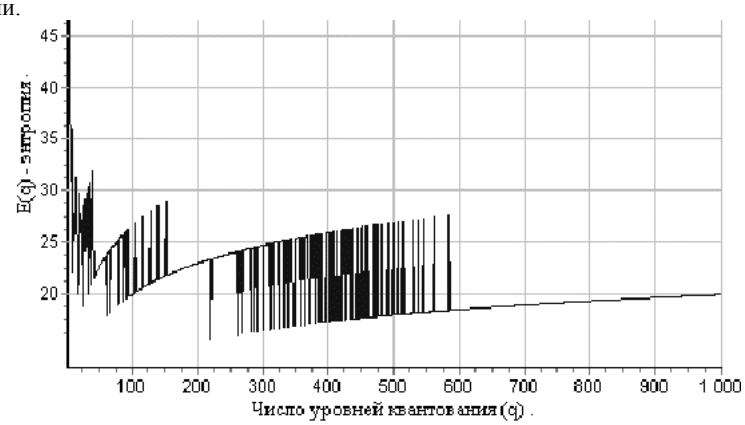


Рис.1

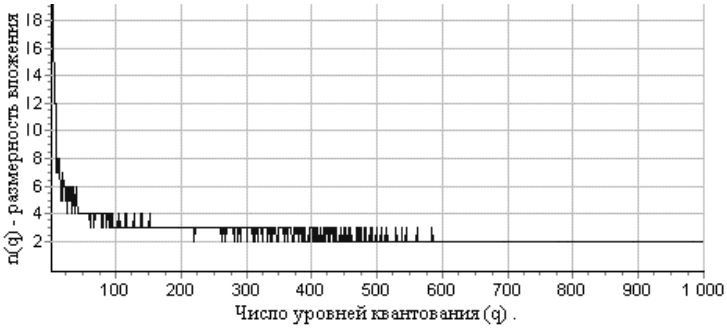


Рис.2

Работа выполнена при частичной поддержке РФФИ (грант 02-02-17573)

[1] Ware JE, Snow KK, Kosinski M, et al. *SF-36 Health Survey Manual and Interpretation Guide.* Boston, The Health Institute, New England Medical Center, 1993.

[2] Грунина Е.А., Борисов А.А. Система определения «качества жизни». В настоящем сборнике. С.344.

[3] Кирьянов К.Г. Идентификация базовых параметров динамики «сцен». /В кн. Труды четвертой научной конференбнции по радиофизике. /Ред. А.В.Якимов. – Н.Новгород: ТАЛАМ, 2000. С.176.

[4] Кирьянов К.Г. Генетический код и тексты: динамические и информационные модели. Учебное пособие, Изд-во ННГУ, 2001. 55с.

СИСТЕМА ОПРЕДЕЛЕНИЯ «КАЧЕСТВА ЖИЗНИ»

Е.А.Грунина[1)], А.А.Борисов[2)]

[1)]*Нижегородская медицинская академия*
[2)]*Нижегородский госуниверситет*

Понятие «качество жизни, связанное со здоровьем» впервые было введено в 1977 году, а в 1993 году Всемирная Организация Здравоохранения дала его определение: «Качество жизни – целостная характеристика физического, психического, эмоционального и социального состояния больного, основанная на его субъективном восприятии». Для его оценки используют физические, психические и социальные критерии. Принципиальными признаками понятия «качество жизни» (КЖ) являются многомерность, изменяемость во времени в зависимости от состояния пациента и участие пациента в оценке собственного состояния. Для оценки КЖ в 1992 году было создан специальный инструмент – опросник SF-36 [1]. Автор John E. Ware создал его с учетом минимальных психометрических стандартов для групповых сравнений, учитывая общие концепции здоровья или благополучия, то есть те параметры, которые *не являются специфичными* для различных возрастных или нозологических групп, а также групп, получающих определенное лечение. *Измерительная модель*, лежащая в основе конструкции SF-36, имеет 3 уровня: пункты (вопросы), 8 шкал, каждая из которых объединяет вместе от 2 до 10 пунктов, 2 суммарных измерения, которые объединяют вместе шкалы. При обработке результатов полученные ответы по всем пунктам формируют 8 шкал. Каждый пункт используется в обработке баллов только одной из шкал. Предполагается, что 8 шкал должны формировать 2 различные высокоупорядоченные группы (суммарные оценки психологического и физического здоровья).Выбрав компоненты оценки КЖ [2] и используя алгоритм автора [3], мы создали компьютеризированную систему оценки КЖ на основе опросника SF-36.

Обследовали 94 студентов IV курса лечебного факультета НГМА, средний возраст 20,9 лет (SD 1,2), из них 46 мужчин и 48 женщин; 46 больных сахарным диабетом 2 типа, средний возраст 66,2 лет (SD 8,4), из них мужчин 14, а женщин 32; и 15 больных ревматоидным артритом, средний возраст 54,4 лет (SD 9,6), из них мужчин 5 и женщин 10.

Студенты по 6 показателям КЖ были сопоставимы со сверстниками из США, на показатели ролевого физического функционирования (РФФ) и ролевого эмоционального функционирования (РЭФ), отражающие адаптацию к повседневной работе, учебе, семейной жизни, были достоверно ниже.

Больные сахарным диабетом и ревматоидным артритом показали разные типы отклонения показателей КЖ от возрастной нормы. Показатели КЖ у больных сахарным диабетом были ниже, чем у сверстников в США, за исключением социального функционирования (СФ). Показатели КЖ у больных ревматоидным артритом были ниже, чем у сверстников в США, за исключением РФФ и РЭФ. РЭФ у больных ревматоидным артритом оказалось выше, чем у студентов. У больных РА вы-

ше показатель боли, ниже физическое функционирование (ФФ) и СФ, но выше РФФ и РЭФ, чем при сахарном диабете.

Выводы

1. Система оценки КЖ – это модель оценки благополучия человека.
2. Опросник SF-36 – современное средство оценки КЖ у разных групп больных и здоровых, валидированное для применения в разных странах.
3. Создана оригинальная компьютеризированная система подсчета результатов SF-36.
4. Показатели КЖ по SF-36 у российских студентов отличаются от сверстников из США.
5. Больные разными заболеваниями демонстрируют разные типы изменения КЖ
6. Ролевое физическое и эмоциональное функционирование у больных ревматоидным артритом выше, чем у студентов, что может отражать как частичную ролевую дезадаптацию студентов, так и хорошую ролевую адаптацию больных ревматоидным артритом.

[1] Ware JE, Snow KK, Kosinski M, et al. *SF-36 Health Survey Manual and Interpretation Guide.* Boston, The Health Institute, New England Medical Center, 1993.
[2] Кирьянов К.Г., Грунина Е.А. К обоснованию выбора компонент состояния (признаков) больного в системах определения «качества жизни». В настоящем сборнике. С.342.
[3] Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, The Health Institute, New England Medical Center, 1994.

АНАЛИЗ ДИАЛЕКТОВ ГЕНЕТИЧЕСКИХ КОДОВ НА ОСНОВЕ ИХ КОДОНОГРАММ

А.А.Баранова[1)], М.Я.Эйнгорин[2)]

[1)]*Нижегородский госуниверситет,* [2)]*НИиОКП «Скит»*

Отображение в 3-мерном пространстве генетического кода белкового биосинтеза, основанного на 64-х 3-символьных позиционных элементах – кодонах: $X_1X_2X_3$, где $X_i \subset \{\boldsymbol{u}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{g}\}$, i = 1,2,3, – в виде кодонограмм позволяет наглядно иллюстрировать большинство сложившихся особенностей данного кода.

Близкие по составу нуклеотидов кодоны на кодонограмме расположены по линиям решетки. На рис.1 изображен код УБК. Выделены стабильные для 17 известных диалектов [1-4] элементарные КГ (ЭКГ) [2, 3, 4], объединяемые по первой паре нуклеотид.

Линии auX_3 и X_1ug – местоположение стартовых кодонов, линии uaR_3 (uX_2a) – стоп кодонов большинства диалектов, содержат специфический РНК нуклеотид u. Нуклеотидный состав КГ выражен характерным графическим рисунком на кодонограмме, подобие которого отмечено для групп auX_3, uaX_3, uX_2a, содержащих оди-

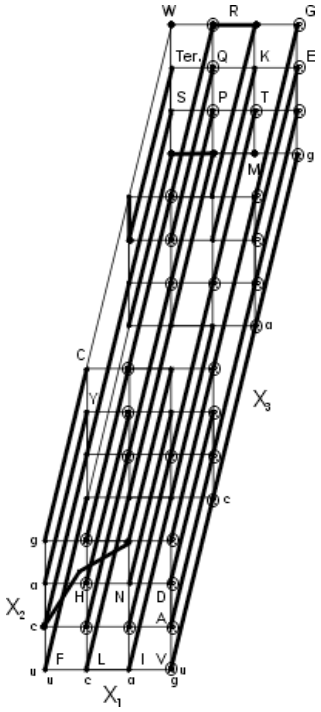


Рис. 1

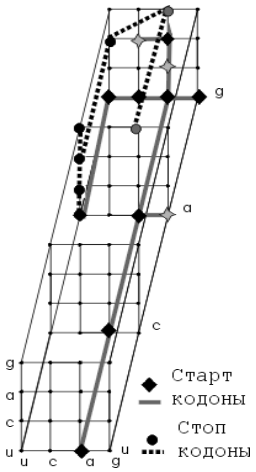


Рис. 2

наковые нуклеотиды. Старт и стоп КГ, расширенные на все диалекты, даже пересекаются на кодоне uua (рис. 2).

На трех ЭКГ: cu1(cuX_3), ua1, ag1, – собраны основные изменения между диалектами. Вводятся нехарактерные для данных ЭКГ аминокислоты, и редкое сочетание полярных и неполярных аминокислот в пределах одной ЭКГ.

[1] Эйнгорин М.Я. Основы кодирования и управления в молекулярной биологии. – Н.Новгород: НГМА, 2001, 117с.

[2] Эйнгорин М.Я. «Кодирование в основах молекулярной биологии», январь 2002 г. Online. http://www.uic.nnov.ru/~emy/biology.html.

[3] Баранова А.А., Эйнгорин М.Я. «Кодонограммы диалектов», февраль 2002 г. Online. http://www.uic.nnov.ru/~emy/codo_ru.html.

[4] Эйнгорин М.Я. «К закономерностям кодирования диалектов», февраль 2002 г. Online. http://www.uic.nnov.ru/~emy/zacon.html.

К АНАЛИЗУ ГЕНЕТИЧЕСКИХ ТЕКСТОВ

А.А.Баранова[1)], М.Я.Эйнгорин[2)]

[1)]*Нижегородский госуниверситет,* [2)]*НИиОКП «Скит»*

Определение функциональных участков в генетических последовательностях сложных организмов в настоящее время имеет большое значение для дальнейшего практического применения в генной инженерии.

Для описания взаимодействия молекул с нуклеотидной цепочкой часто достаточно выделения характерных участков в ее первичной структуре: в нуклеотидном тексте ДНК или РНК [1,2]. В последовательном характерном участке допускается неоднозначное сочетание символов, в том числе, ограниченное бинарными признаками нуклеотидов. Бинарные последовательности, отражающие свойства элементов нуклеотидной цепи, вводятся в работе [3].

При непосредственном сравнении 2-ичные представления генетических текстов учитывают большее количество факторов подобия, по сравнению с 4-ичной последовательностью нуклеотид.

Рассмотрим показатели подобия участков в нуклеотидной (N0) и в бинарной (Na) последовательностях гена Thermoascus aurantiacus endo-1,4-beta-xylanase A precursor (xynA) gene [4]. Ген содержит кодирующие и некодирующие белок участки, которые чередуются между собой. В процессе биосинтеза некодирующие участки удаляются. Возьмем 11 кодирующих участков, из которых далее формируется единая матрица для синтеза белка. Выберем 50 нуклеотидов до начала кодирующего участка и 40 в нем, всего по N=90 нуклеотид для каждого участка. Двоичную последовательность (Na) построим по бинарному признаку нуклеотидов: пурин или пиримидин.

Табл. 1

N0	1	2	3	4	5	6	7	8	9	10	11
1	1,00	0,26	0,24	0,20	0,19	0,23	0,20	0,27	0,27	0,21	0,30
2	0,26	1,00	0,28	0,27	0,27	0,30	0,30	0,29	0,27	0,30	0,32
3	0,24	0,28	1,00	0,26	0,33	0,32	0,26	0,28	0,27	0,34	0,29
4	0,20	0,27	0,26	1,00	0,26	0,29	0,32	0,31	0,20	0,28	0,19
5	0,19	0,27	0,33	0,26	1,00	0,29	0,39	0,27	0,29	0,28	0,36
6	0,23	0,30	0,32	0,29	0,29	1,00	0,30	0,28	0,24	0,27	0,20
7	0,20	0,30	0,26	0,32	0,39	0,30	1,00	0,33	0,27	0,30	0,26
8	0,27	0,29	0,28	0,31	0,27	0,28	0,33	1,00	0,33	0,30	0,33
9	0,27	0,27	0,27	0,20	0,29	0,24	0,27	0,33	1,00	0,21	0,31
10	0,21	0,30	0,34	0,28	0,28	0,27	0,30	0,30	0,21	1,00	0,28
11	0,30	0,32	0,29	0,19	0,36	0,20	0,26	0,33	0,31	0,28	1,00

В табл. 1, 2 приведены попарные коэффициенты подобия *Kp* участков гена, выраженных нуклеотидной и бинарной последовательностями:

$Kp = \sum_{N} k_i \ / N,$

$k_i = 1$, если соответствующие символы совпали, 0 - иначе, $i = (1 \ldots N)$.

Табл. 2

Na	1	2	3	4	5	6	7	8	9	10	11
1	1,00	0,44	0,44	0,42	0,44	0,46	0,49	0,52	0,61	0,46	0,51
2	0,44	1,00	0,49	0,51	0,58	0,57	0,60	0,59	0,50	0,54	0,56
3	0,44	0,49	1,00	0,58	0,53	0,57	0,53	0,48	0,52	0,59	0,49
4	0,42	0,51	0,58	1,00	0,58	0,54	0,53	0,57	0,50	0,52	0,47
5	0,44	0,58	0,53	0,58	1,00	0,59	0,58	0,59	0,54	0,48	0,53
6	0,46	0,57	0,57	0,54	0,59	1,00	0,57	0,58	0,47	0,49	0,57
7	0,49	0,60	0,53	0,53	0,58	0,57	1,00	0,61	0,50	0,57	0,51
8	0,52	0,59	0,48	0,57	0,59	0,58	0,61	1,00	0,58	0,49	0,52
9	0,61	0,50	0,52	0,50	0,54	0,47	0,50	0,58	1,00	0,49	0,54
10	0,46	0,54	0,59	0,52	0,48	0,49	0,57	0,49	0,49	1,00	0,50
11	0,51	0,56	0,49	0,47	0,53	0,57	0,51	0,52	0,54	0,50	1,00

Как видно из табл. 1,2 бинарные последовательности имеют более высокий коэффициент подобия *Kp*. Кроме того, число совпадающих значений соответствующих символов по всем 11 участкам (n) выше для бинарной последовательности. В табл. 3 приведено распределение соответствующих символов последовательностей по n.

Табл. 3

N	2	3	4	5	6	7	8	9	10	11
Код N0	-	5	36	25	16	5	1	1	1	0
Код Na	-	-	-	-	28	28	21	10	3	0

Множество подобных участков бинарных последовательностей включает множество подобных нуклеотидных участков, что может быть использовано при определении неоднозначных по нуклеотидному составу функциональных участков.

[1] Ярыгин В.Н., Васильева В.И., Волков И.Н., Синельщикова В.В. Биология. В 2кн. Кн.1./ Под ред. В.Н.Ярыгина. 2-е изд. –М.: Высш. шк. 1999, с.105.
[2] Паспорта - Эндонуклеаза рестрикции. Online. http://www.vekton.spb.ru/_private/project/bio.htm
[3] Эйнгорин М.Я. Основы кодирования и управления в молекулярной биологии. – Н.Новгород: НГМА, 2001, 117с.
[4] The National Center for Biotechnology Information. Online. http://www.ncbi.nlm.nih.gov/