

МЕТОДОЛОГИЯ РЕГУЛИРОВАНИЯ ЭНТРОПИИ В КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Е.В. Юркевич
(ИПУ им. В.А. Трапезникова РАН)
Н.В. Азаренкова
(МГТУ МИРЭА)
В.М. Долкарт
(ОАО «Корпорация «ВНИИЭМ»)

Поставлена задача минимизации ошибок в сообщениях, поступающих из базы данных корпоративной информационной системы. Предложены возможности регулирования энтропии как характеристики помех в информационном обеспечении системы управления космического аппарата. Показаны механизмы нарастания шума при запуске и реализации интернет-ресурсов, отвечающих за web-доступ к базам данных. Оценку энтропии базы данных предложено рассчитывать в соответствии с кластеризацией важности передаваемых сообщений по методологии IDEF1 и с выявлением исхода эксперимента наиболее вероятного по биномиальному распределению Бернулли. Для промежутка времени, определённого разработчиком, предложен механизм модернизации аппаратной части корпоративной информационной системы.

Ключевые слова: интернет-ресурсы, базы данных, энтропия, ошибки в сообщениях, корпоративная информационная система, кластеризация потоков информации, механизм IDEF1, наиболее вероятный исход эксперимента, биномиальное распределение Бернулли, модернизация hard-ware.

Постановка задачи обеспечения надёжности работы информационных комплексов

С целью минимизации затрат на обеспечение штатной работы комплекса управления космическим аппаратом (КА) ставится задача разработки методологии анализа динамики реакций информационных систем на внешние воздействия. При реализации и запуске крупных ресурсов, отвечающих за web-доступ к базам данных (БД) часто возникает необходимость в регулировании нарастающей неопределённости информации о текущем состоянии комплекса управления.

Будем полагать, что рост такой неопределённости приводит к погрешностям в работе системы управления. В этом случае для обеспечения надёжности управления КА воспользуемся моделью К. Шеннона¹ [1]: $H + I = 1$, где I – количество полезной информации; H – энтропия базы данных, т. е. количество содержащегося в

¹В работе принята модель К. Шеннона, в которой под информацией понимаются только сигналы, полезные для получателя. Неполезные сигналы, т. е. шум, помехи, оцениваются энтропией. Количество информации (в полезном сигнале) и количество энтропии (в сигнале помехи) рассчитывается алгоритмически одинаково. В данной работе примем, что, если сигнал, поступающий из базы данных, является точной копией сигнала, принятого с контролируемого объекта, то это означает отсутствие энтропии в базе данных. Такое отсутствие шума означает максимум информации.

ней шума. Примем, что в рассматриваемой задаче оптимизации работы базы данных, в связи с большим разнообразием характеристик используемых сообщений, выгоднее управлять количеством шума, чем количеством полезной информации.

Анализ механизмов проектирования и результатов информационного обеспечения работы комплексов управления приводит к выводу об актуальности проблемы совершенствования методов регулирования энтропии в web-ресурсах корпоративных информационных систем (КИС).

В данной работе, с целью разработки предложений по предотвращению критических сбоев в работе КА, рассмотрим задачу оптимизации воздействий на КИС с учётом её работы в каждом из периодов жизненного цикла.

Собственно КИС будем рассматривать как совокупность функциональных блоков, реализованных в виде программно-технических средств. Применительно к задаче формирования механизма создания КИС, обеспечивающего реализацию и доступ к web-ресурсам в сложных системах управления, используем модель [2]:

$$Q_{\Sigma} = \min_t \max_{(1-H_{\alpha})(1-H_{\beta})} f(P, L, S, (1-H_{\alpha}), (1-H_{\beta}), t), \quad (1)$$

где Q_{Σ} – финансовые затраты на формирование

информационного обеспечения КИС; P – множество задач функциональных блоков КИС; L – множество целей использования функциональных блоков КИС; S – множество характеристик условий работы функциональных блоков КИС; H_α – оценка энтропии в реакции КИС в виде вероятности информационного искажения при устранении исправного функционального блока; H_β – оценка энтропии в реакции КИС в виде вероятности информационного искажения при не устранении бракованного функционального блока; t – характеристика затрат ресурсов на формирование информационного обеспечения функционального блока при T – общих затратах ресурсов на создание КИС.

Если величины P, L, S рассматривать как ограничения, то множество допустимых стратегий создания КИС (C) определим с помощью отображения оператором O_1 :

$$O_1: R P L S \Rightarrow C, \quad (2)$$

где R – множество всех стратегий создания КИС.

С учётом затрат ресурсов на формирование информационного обеспечения для решения задач, поставленных перед КА, предлагается прогнозировать выбор стратегии с помощью отображения оператором O_2 , т. е.

$$O_2: L P S t C \Rightarrow Y, \quad (3)$$

где Y – множество результатов работы элементов КИС.

Выражение (3) будем рассматривать как модель формирования требований к КИС, создаваемой в соответствии со стратегией $c \in C$ при затратах ресурсов t . Причём t ограничены пределами, характерными для каждого этапа жизненного цикла КИС.

Оценку надёжности информационного обеспечения КИС представим отображением оператором O_3 :

$$O_3: L P S t \Rightarrow U, \quad (4)$$

где U – множество оценок вероятности достижения целей $l \in L$.

В общем виде оценка возможностей достижения цели $l \in L$ может быть сведена к условиям:

$$\begin{aligned} W(l, p, s, t, y) &= 0; \\ Q(l, p, s, t, y) &> 0; \end{aligned} \quad (5)$$

$$H(l, p, s, t, y) \Rightarrow \text{extr},$$

где W, Q, H – векторы требований типа равенства, неравенства и экстремума, соответственно.

С учётом выражений (2) – (5) задачу оптимизации информационной проработки КИС запишем:

$$\begin{aligned} H(l, p, s, t, y(l, p, s, t, c)) &\Rightarrow \text{extr} \quad (6) \\ C &\supseteq c \end{aligned}$$

при ограничениях:

$$\begin{aligned} W(l, p, s, t, y(\cdot)) &= 0; \\ Q(l, p, s, t, y(\cdot)) &> 0, \end{aligned}$$

где $y(\cdot)$ – конструктивное описание отображения оператором O_2 .

Таким образом, в соответствии с целевой функцией, записанной в данной работе выражением (1), ставится задача формирования механизма минимизации суммарных затрат на обеспечение функциональной надёжности работы КИС при минимизации вероятности информационного искажения в случае устранения исправного средства или не устранения средства бракованного.

Механизм оптимизации ресурсов для обеспечения функциональной надёжности работы КИС

ГОСТ Р МЭК 61508 «Функциональная безопасность электрических, электронных, программных электронных систем контроля и управления, связанных с безопасностью», вводя термин «функциональная безопасность», ориентирует аналитиков и разработчиков соответствующих систем на возможности снижения риска неполучения ожидаемого результата от управленческого воздействия при фиксированном влиянии внешних факторов. Стандарт определяет, что существенные сбои в работе системы управления могут быть порождены неправильными функциональными требованиями к аппаратным средствам, а также к программному обеспечению и к описанию допустимых влияний внешней среды. Положения этого стандарта предполагают необходимость выделения функциональных блоков, определяющих достижимость цели построения системы управления, и оценку вероятности достижения этой цели.

Обобщая положения стандарта ГОСТ Р МЭК 61508 в применении к нашей задаче, будем использовать понятие «функциональная надёжность». Под таким термином предлагается понимать часть об-

щей надёжности системы, которая определяется оценкой риска несоответствия её реакции на внешние воздействия алгоритму достижения её целевой функции.

В настоящей работе в качестве функционального блока, определяющего устойчивость управления КА, рассмотрим КИС. Вероятность достижения цели её построения будем оценивать через энтропию в базах данных и, соответственно, в информационных потоках при обращении к этим базам данных.

В нашем рассмотрении энтропия является только количественной характеристикой информационного шума. В связи с этим для обеспечения адекватности описания каждого из потоков предлагается их классифицировать по важности передаваемых сообщений. В результате с помощью механизма ранжирования данных IDEF1 (Information Definition) [3, 4] были выделены информационные потоки одинаковой важности.

Применительно к рассматриваемой задаче миссия IDEF1 рассматривалась как выявление потребности в информационных ресурсах, необходимых для обеспечения устойчивости управления КА. Проведённый нами вычислительный эксперимент показал, что применение расчётного метода позволяет своевременно вносить коррективы в конструкцию системы управления.

В соответствии с выражением (2) идентифицировались направления транзакций согласно выбранной стратегии формирования потоков информации. Системность предлагаемого подхода предполагает полноту учёта целей каждого из функциональных элементов КИС, и, соответственно, полноту стратегий создания КИС (R) в целом. К таким элементам в работе отнесены компьютеры, серверы, операторы, линии связи и т. п. [5].

С помощью выражения (3) предлагается формировать условия штатной работы каждого из выделенных элементов. Согласно выражению (4) по биномиальному распределению Бернулли предлагается оценивать надёжность информационных ресурсов элементов КИС. Эти оценки являются действенным инструментом для определения надёжности функциональных блоков, использующих информационное обеспечение, определенное техническими условиями на систему управления.

В применении к проектированию КИС расчёты энтропии информационного обеспечения на основе выявления наиболее вероятного исхода эксперимента показали достаточность ресурсов для удо-

влетворения требований к возможностям soft-ware, определённых разработчиками системы управления [6]. Рассмотрение оптимизационной задачи (6) позволило сделать вывод о том, что для снижения энтропии информационного обеспечения КИС ещё на этапе проектирования требуется предусматривать возможность модернизации hard-ware [5].

Особенности использования методов расчёта энтропии при оптимизации КИС космического аппарата

Предлагаемый подход к оптимизации затрат ресурсов на обеспечение функциональной надёжности работы КИС рассмотрим на примере расчёта энтропии базы данных в управляющем комплексе КА. В качестве начального постулата будем полагать, что для взаимодействия КА и Центра управления полётом требуется доступ к web-ресурсам.

Ставится задача регулирования энтропии во время проведения эксперимента, запланированного на КА. Ключевые данные для контроля процессов, происходящих в КИС на уровне взаимодействия запросов и контейнеров, могут дать исследования энтропии в базе данных системы управления КА.

Согласно предлагаемой в данной работе модели механизма создания КИС, на энтропию влияют характеристики ситуаций, определяющих доступ к web-ресурсам. Согласно модели К. Шеннона [1] значение энтропии будем рассматривать как оценку приближения информации, полученной с КИС к реальному значению. Точность этого значения будет тем выше, чем больше ситуаций учтено при её проектировании.

Важной составляющей технического задания на разработку системы управления КА должны являться условия минимизации энтропии информации, содержащейся в базе данных КИС. Наиболее распространённым ограничением в реализации этого требования является количество элементов в системе управления.

Для учёта особенностей формирования баз данных КИС систематизируем характеристики поступающей информации с помощью алгоритмов кластеризации, описанных в [7]. Например, для анализа работы КА требуется учёт характеристик внешней среды и временных характеристик передачи информации между функциональными блоками КИС.

Традиционно эффективность решения таких задач оценивается с помощью последовательного анализа комбинаций затрат соответствующих ин-

формационных ресурсов на каждом из этапов работы, включая разработку тестовых сценариев, спецификаций и просмотров [8]. Причём, сверху вниз должны проверяться требования на корректность описания информационного взаимодействия между элементами рассматриваемой КИС, а также корректность описания связей между требованиями для каждого из них. Отдельное внимание должно уделяться особенностям в использовании уже имеющихся тестов.

Предлагаемый нами механизм поиска существенных характеристик, определяющих функциональную надёжность КИС, основан на нахождении критических точек в решении задач, поставленных перед КА. В таких точках возникает опасность появления несоответствия информации, получаемой с помощью расчётов, и данных, предоставляемых КА, полученных эмпирически. Выявление таких точек предлагается проводить с помощью графов [9].

Разработка компьютерных моделей работы функциональных блоков КИС, несмотря на специфическую направленность каждого из них, обычно характеризуется одними и теми же этапами. Причём на каждом из этих этапов требуется построение цепи технологических операций по тестированию, а каждая операция характеризуется фиксированным порядком действий [5].

Допустим, построены графы, описывающие работу КИС. В этих графах учитываются программные и технические особенности каждого блока, а также целевые характеристики и формы взаимодействия между ними. На каждом из них наглядно представляется механизм построения конкретной модели в виде узлов, ветвей и путей. Узел графа в данном случае обозначает точку, характеризующую точностью реакции блока на поступающее сообщение, а ветвь – это линия, соединяющая два узла (т. е. переход от одного информационного содержания базы данных к другому).

В решении нашей задачи ветви оценивались информационными ресурсами, необходимыми для перехода от узла к узлу. В качестве пути рассматривалась часть графа, образованная последовательными ветвями. Собственно информационные ресурсы оценивались по относительному размеру полученного сообщения. В качестве допустимого принято предположение, что для одинакового изменения работы различных блоков требуется восприятие одинаковых информационных ресурсов. Следовательно, при одинаковой важности получаемой информации для каждого блока, чем выше

энтропия воспринимаемого сообщения, тем меньше скорость восприятия информации, содержащейся в данном сообщении.

Отметим, что в данной работе все оценки проводились относительно времени реакции блока, имеющего минимальное время запаздывания. В тестировании КИС на функциональную надёжность особое значение имеет корректность введения контуров обратной связи, определяющих анализ результатов её использования.

В рассматриваемых задачах такие контуры могли быть двух видов. Контур обратной связи как замкнутый путь, состоящий из ряда ветвей, возвращающихся в исходный узел, и контур, состоящий из одной замкнутой ветви, образующей петлю.

Для нахождения существенных характеристик таких систем с помощью построенных графов нами использовались простые эквивалентные преобразования графов, позволяющие заменить последовательные и параллельные пути отдельными ветвями (например преобразование звезды в треугольник с помощью исключения одного узла) и решение систем уравнений, соответствующих построенным графам.

В связи с тем, что увеличение количества и протяжённости ветвей уменьшало надёжность получения ожидаемых результатов, а углубление обратной связи её повышало, то для оптимизации использования модели КИС проводились операции над графами, направленные на оптимизацию числа контуров обратной связи. Такие преобразования проводились на основании анализа экспертных заключений, в результате которых достигались необходимые упрощения.

В конечном итоге получался граф, выражающий в простой форме зависимость относительной оценки используемых информационных ресурсов от переменных, определённых экспертами.

Для повышения надёжности анализа работы функциональных блоков КИС требовалась оценка надёжности результатов анализа каждого из путей на построенном графе. Независимо от действий наблюдателя из Центра управления полётом (кроме ошибочных действий последнего) предлагается проводить сопоставление мнений экспертов с результатом эксперимента.

Эксперименты показывают, что наиболее часто условиями появления артефактов (в рамках рассматриваемой задачи) является некорректность её постановки, отражённая в инфологической составляющей модели [7].

В этом случае результат использования базы данных КИС может быть в одном из двух видов – безошибочном или ошибочном. Безошибочный результат соответствует всем требованиям технического задания на разработку КА. Ошибочным считается результат, если он не соответствует хотя бы одному из требований задания, т. е. когда в нём имеется хотя бы одна ошибка. Определим, что ошибкой полученного результата является каждое его несоответствие требованиям задания.

Особенностью технологии формирования базы данных, используемой в работе, являются риски появления явных и скрытых ошибок. Явные ошибки обнаруживаются проверкой по правилам, предусмотренным тестами на датологическую часть информационного обеспечения. Остальные дефекты (скрытые) можно выявить с помощью средств контроля инфологической составляющей. Такое положение накладывает дополнительные требования к методам анализа энтропии информационного обеспечения при проверке КИС на функциональную надёжность.

Другой особенностью явились различия в последствиях, к которым может привести наличие ошибки в базе данных. Если её наличие исключает использование полученных результатов по назначению (в терминологии данной работы – из-за недостаточности функциональной надёжности базы данных), то такая ошибка (или ошибки) рассматриваются как критические. КИС, разработанная с малозначительными ошибками, может использоваться по назначению, и наличие таких ошибок не влияет на описание процессов выполнения КА своих функций. В связи с этим в дополнение к рассмотренным двум видам выводов (безошибочный и ошибочный) целесообразно ввести ещё две характеристики состояния КИС – работоспособное и неработоспособное.

Работоспособным предлагается называть состояние КИС, когда получаемой с её помощью информации достаточно для формирования выводов о соответствии работы КА заданной программе полёта.

В нашем случае узлы графа, характеризующиеся возможностью перехода КИС из работоспособного состояния в неработоспособное, предложено называть критическими точками. Следовательно, задача анализа ветвей построенного графа на определение моментов появления функциональных ошибок сводится к поиску таких точек.

В данной работе в качестве бракующего условия было принято наличие ошибки, определяемой в со-

ответствующей критической точке. Анализ условий возникновения таких ошибок показал, что они появляются в двух случаях:

– вследствие недостатков разработки инфологической части информационного наполнения базы данных, что не связано с внешними воздействиями или процессами формирования программного обеспечения;

– при событии, называемом помеха в программировании. Однако несущественная помеха может не повлиять на результаты использования КИС, а существенная приводит к нарушению её работоспособности. Такое нарушение может явиться следствием появления ошибок, но не всегда наличие ошибок приводит к нарушению. Например недостатки программного обеспечения при формировании информационных ресурсов обуславливают возникновение соответствующих ошибок [8].

Итак, для минимизации вычислительных ресурсов при тестировании базы данных предлагается схема выделения операций, требующих особого внимания. Для этого необходимо выделить узлы, непосредственно влияющие на надёжность результата анализа и сформулировать причины появления бракующих условий. Важным этапом в определении таких причин является анализ энтропии информационного обеспечения КИС.

Пусть в нашем эксперименте выделено 8 этапов управления, а также рассчитана вероятность выполнения каждого этапа p_i . В том числе: запуск системы ($p_1 = 0,9$), вход в систему ($p_2 = 0,9$), обработка пары пароль и логин на соответствие ($p_3 = 0,7$), открытие системы с расширенным доступом ($p_4 = 0,9$), ответ на запрос пользователей ($p_5 = 0,8$), выход из основного меню системы ($p_6 = 0,9$), обработка команды выход ($p_7 = 0,8$), отказ системы ($p_8 = 0,1$).

При этих условиях рассчитаем b -арную энтропию для КИС, где p_i – вероятность выполнения i -го этапа управления при $n = 8$. Тогда

$$H_b(S) = -\sum_{i=1}^n p_i \log_b p_i;$$

$$H_b(S) = -0,9\log_0,9 - 0,9\log_0,9 - 0,7\log_0,7 - 0,9\log_0,9 - \\ - 0,8\log_0,8 - 0,9\log_0,9 - 0,8\log_0,8 - \\ - 0,1\log_0,1 = 0,52819.$$

Значение 0,52819 будет достигнуто при прохождении всех этапов управления. В нашем случае (с условным запасом) будем считать, что критерием жизнеспособности КИС является

Таблица 1

Энтропия на этапах работы КИС

Этапы работы КИС	<i>P</i>	<i>H_b(S)</i>
Запуск системы	0,9	0,01283
Вход в систему	0,9	0,01702
Обработка пары пароль и логин на соответствие	0,7	0,12596
Открытие системы с расширенным доступом	0,9	0,01702
Ответ на запрос пользователей	0,8	0,08406
Выход из основного меню системы	0,9	0,01702
Обработка команды «выход»	0,8	0,08406
Отказ системы	0,1	0,45686

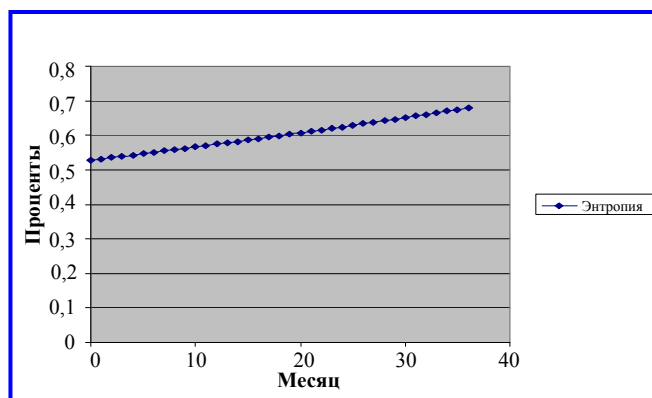


Рис. 1. Динамика энтропии в КИС в течение 36 месяцев

Таблица 2

Расчёт энтропии по фьючерсам (36 месяцев)

Фьючерсы	Энтропия	Фьючерсы	Энтропия
0	0,52819	19	0,603043
1	0,53188	20	0,607264
2	0,53561	21	0,611517
3	0,53935	22	0,615798
4	0,54313	23	0,620108
5	0,546937	24	0,624449
6	0,550765	25	0,628820
7	0,554621	26	0,633222
8	0,558503	27	0,637655
9	0,562412	28	0,642118
10	0,566349	29	0,646613
11	0,570313	30	0,651139
12	0,574305	31	0,655696
13	0,578325	32	0,660286
14	0,582374	33	0,664908
15	0,586450	34	0,669563
16	0,590555	35	0,674250
17	0,594689	36	0,678969
18	0,598851		

значение вероятности прохождения этапа работы равно 0,59.

Примем следующие допущения: в начальном состоянии КИС является работоспособной. Под начальным состоянием будем считать разработанную КИС, прошедшую все виды тестирования, но ещё не воспринимавшую обращения пользователей.

Для каждого из этапов значения энтропии приведены в табл. 1.

В табл. 1 значения для *P* условны. Будем считать (как было определено), что в работоспособной системе приемлемой для ответа является *P* = 0,59.

В натуральных расчётах эта величина определяется экспертно на основании статистических данных.

В ходе вычислительного эксперимента были приняты следующие условия: количество пользователей не менялось; количество запросов пользователей за один временной интервал увеличивалось с небольшой прогрессией. По итогам статистических данных за 12 месяцев был получен эмпирический коэффициент изменения состояний КИС *K_s* = 1,007.

Для расчёта изменения энтропии информационного обеспечения КИС во времени была предложена формула:

$$H > H_1 = H_0K; H_2 = H_1K; H_3 = H_2K = (H_1K)K = (H_0K)KK = H_0K^3.$$

Отсюда следует $H(\tau) = H_0K^\tau$, где τ равно временному интервалу (в нашем случае месяцу). При этом не рассматривается $\tau_i = 0$. На рис. 1 показан пример расчёта энтропии за три года.

Допустим полётное задание КА предполагает необходимость его работы в течение трёх лет. В табл. 2 приведён пример расчёта энтропии на три года.

Вычислительный эксперимент показал, что при расчёте изменения энтропии в течение трёх лет порог принятой нами максимально допустимой вероятности успешного исхода испытаний равный 0,59 достигается за 30 месяцев, т. е. по истечении 2,5 лет, необходима модернизация, устраняющая фактический хаос в базе данных КИС.

Метод ранжирования данных IDEF1 осуществляется посредством методологии семейства Manufacturing ICAM (Integrated computer-aided manufacturing) для решения задач моделирования сложных систем. В нашем случае метод IDEF1 обеспечивает моделирование ин-

формационных потоков внутри КИС, позволяющие отображать и анализировать их структуру и взаимосвязи. Такой метод позволяет отображать и анализировать модели деятельности широкого диапазона систем [7]. При этом широта и глубина обследования процессов в КИС определяется самим разработчиком, что позволяет не перегружать создаваемую модель излишними данными.

Например, существует КИС, в которую запросы пользователей могут приниматься только в режиме реального времени. В этом случае разработчикам необходимо определить пропускную способность каналов связи для обработки запросов, полученных on-line. Однако в общем случае для распределения нагрузки на оборудование разработчики должны учитывать введение и режима off-line, т. е. предусмотреть режим обработки запроса с неограниченными временными рамками.

Особенностью IDEF1 является возможность использования не только баз данных на компьютерных носителях, но и информационных возможностей сотрудников, характеристик рабочих мест, сообщений, переданных по телефону, и т. д. Неоднородность каналов информационной связи является важной причиной повышения энтропии в информационном обеспечении системы управления КА. В данной работе появление таких неопределенностей предлагается компенсировать с помощью распараллеливания потоков, создающих затор при обращении к соответствующему блоку информации. Для этого предлагается введение технологии ранжирования важности данных, и в соответствии с информацией о заторе, поступающей от специальных программ мониторинга работы КИС, предлагается формировать самостоятельные потоки с данными одного ранга.

В такой структуре КИС должны быть определены возможные пути прохождения потоков информации. Каждый пользователь должен иметь свой уникальный идентификатор ID (identifier). По его значению система распознает уровень доступа к ресурсам, разрешённый данному пользователю. Каждый пользователь для получения, внесения, изменения каких-либо данных должен обращаться в КИС через пользовательский модуль (web-интерфейс). Далее в зависимости от запроса пользователя поток поступившей от него информации должен следовать по заранее определенному маршруту.

Биномиальное распределение. В данной работе все ответы на запросы рассматриваются как

последовательность независимых случайных величин $\{X, X, \dots\}$, вероятность успешных испытаний равна P и имеет одно и то же распределение $Bi(l, p)$. В этом случае можно полагать, что имеется бернуллиевская последовательность, определяющая распределение количества успешных серий испытаний, содержащих n независимых случайных экспериментов [10].

Представим, что под серией испытаний понимается единовременное обращение пользователей к КИС. В данной работе каждая из проведённых серий испытаний содержала по 10 экспериментов, т. е. по десять обращений одновременно. В каждой серии испытаний менялось количество успешных и неуспешных исходов экспериментов. Под неуспешным исходом понимался отказ КИС (ошибка или сбой), возникший при некотором числе обращений.

Таблица 3

Расчёт вероятности успеха проводимых испытаний

№ испытания	$P(\Phi_k)$	Количественные соотношения успешных и неуд. экспериментов
1	0,3486784401	Все 10 эксперим. – успех
2	0,3874204890	9 эксперим. – успех, 1 – неуд. экспериментов
3	0,1937102445	8 эксперим. – успех, 2 – неуд.
4	0,0573956280	7 эксперим. – успех, 3 – неуд.
5	0,0111602610	6 эксперим. – успех, 4 – неуд.
6	0,0014880348	5 эксперим. – успех, 5 – неуд.
7	0,0001377810	4 эксперим. – успех, 6 – неуд.
8	0,0000087480	3 эксперим. – успех, 7 – неуд.
9	0,0000003645	2 эксперим. – успех, 8 – неуд.
10	0,0000000090	1 эксперим. – успех, 9 – неуд.
11	0,0000000001	0 успеха, 10 – неуд. экспериментов
Бином	1,0000000000	

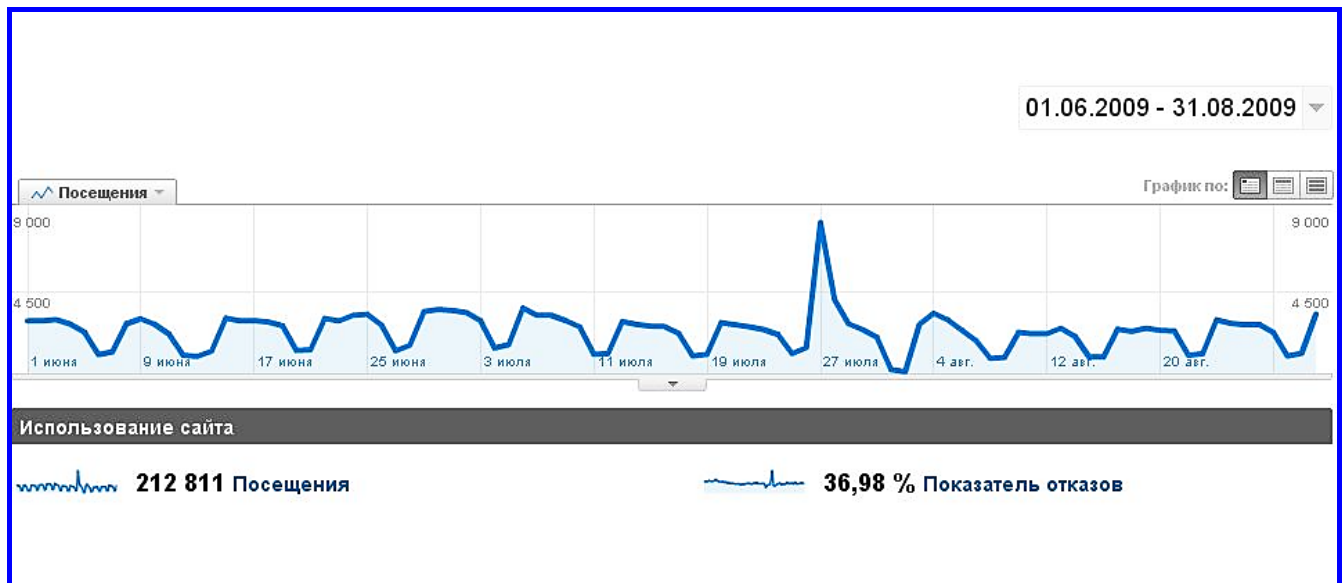


Рис. 2. Статистика посещений сайта (обращений к КИС) с 01.06.2009 по 31.08.2009

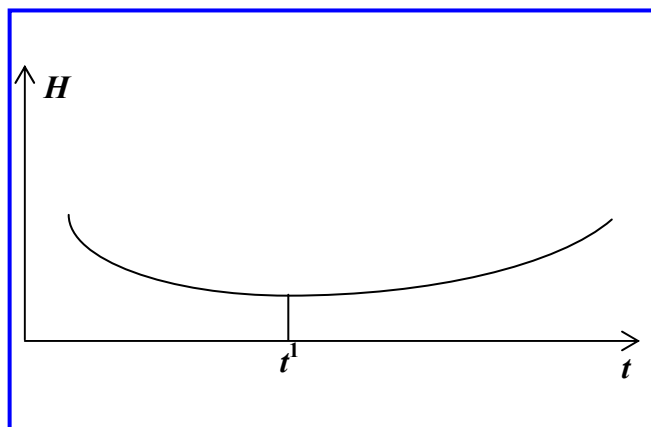


Рис. 3. Качественная зависимость энтропии информационного обеспечения от количества ресурсов, используемых для обработки потока обращений к КИС

В терминах данной работы ошибка рассматривалась как следствие повышенной энтропии в базе данных.

Числовой расчёт показывает, что в рассматриваемом вычислительном примере число отказов КИС минимально.

В табл. 3 сведены результаты биномиального распределения. Проведено 11 серий испытаний, в каждой серии проводилось по 10 экспериментов. Пусть $P(\Phi k)$ – вероятность успеха проводимой серии испытаний.

Экспериментально показано, что при проведении очередной серии показатель $P(\Phi k)$ изменяется. В нашем случае при проведении серии 1 испытаний на выходе все 10 экспериментов экспертно признаны успешными ($P(\Phi k) = 0,3486784401$). При

проведении серия 11 испытаний все 10 экспериментов экспертно признаны как неуспешные ($P(\Phi k) = 0,0000000001$).

Для оценки корректности вычислений было проведено суммирование всех показателей $P(\Phi k)$, полученных в ходе операции. Наличие суммы равной единице является критерием соответствия распределения биномиальному расчёту.

Таким образом, проведённые расчёты показывают, что в соответствии с моделью формирования требований к КИС (3), для каждого этапа её жизненного цикла, можно выбрать оператор O_2 . С помощью этого оператора устанавливается соответствие ресурсов t , имеющихся в существующей реализации КИС, и количества запросов, требуемых для успешного выполнения заданий КА. В результате выявляются условия для выбора стратегии $s \in C$, позволяющей получать успешные результаты испытаний внутри серии.

Модернизация аппаратной части. Рассматриваемая технология обеспечения доступа к web-ресурсам в сложных системах управления ориентирована на обеспечение надёжности КИС. В данной работе эта технология представлена выражением (4).

Ставится задача оптимизации сочетания требований к информационному обеспечению КИС и возможностей hard-ware, на котором она реализована.

Вычислительный эксперимент решения такой оптимизационной задачи рассчитывался на примере сайта Российской академии народного хо-

зайства и государственного управления при Президенте РФ (<http://www.rane.ru/>). Анализ динамики посещений этого сайта позволил зафиксировать диаграмму обращений к КИС, показанную на рис. 2.

На диаграмме статистики обращений видно, что основной пик пришёлся на 27 июля. Всего за указанный период было 212 811 посещений.

К показателям отказа относится просмотр главной страницы сайта с отсутствием перехода в какой-либо раздел. 27 июля 2009 г. сервер не выдержал нагрузку, поэтому произошёл отказ системы. Сайт стал недоступен.

В соответствии с выражением (1) в качестве целевой функции данной работы рассматривается минимум суммарных затрат ресурсов на *soft-ware* и *hard-ware* при минимизации вероятности информационного искажения в случае устранения исправного средства или не устранения средства бракованного. Согласно выражению (4) была найдена регрессия вида

$$H_{\Sigma} = f(P, L, S, t), \quad (7)$$

где H_{Σ} – суммарная энтропия информационного обеспечения КИС; P – множество задач функциональных блоков КИС; L – множество целей использования функциональных блоков КИС; S – множество характеристик условий работы функциональных блоков КИС; t – характеристика затрат ресурсов на формирование информационного обеспечения КИС.

Желание разработчика КИС минимизировать сложность её аппаратной реализации приводило к повышению энтропии информационного обеспечения из-за повышения ограничений в условиях её работы. Усложнение конструктивных решений повышало количество отказов из-за увеличения количества критических точек.

Анализ аналитической записи выражения (7) показал, что функция $H(t)$ имеет ярко выраженный экстремум, определяющий оптимальную величину затрат ресурсов на создание *soft-ware* и *hard-ware* при фиксированных оценках P , L , S (рис. 3).

На основании расчётов с помощью предлагаемых моделей было решено модернизировать аппаратную часть системы, путём добавления второго сервера. Точка t^1 была найдена с выбором сервера Nginx, который обладает рядом преимуществ по сравнению с уже имеющимся сервером Apache.

При последующих испытаниях система выдержала нагрузку, что оправдывает наличие дополнительного сервера Nginx.

Выводы

Оценка функциональной надёжности КИС с помощью расчёта энтропии в базах данных показала, что анализ биномиального распределения по вероятности Бернулли явился действенным методом выявления наиболее вероятного результата испытаний.

Введение метода ранжирования данных повышало точность определения критических точек в системе управления КА, а выделение потоков данных одного ранга позволило уменьшить частоту сбоев в транзакциях. При этом введение параллельных потоков информации минимизировало задержки в прохождении сообщений от КИС.

На основании анализа динамики энтропии информационного обеспечения КИС было получено, что в рассмотренном случае максимально допустимый риск сбоя транзакций возникает на тридцатом месяце её жизненного цикла.

Экспериментально определено, что через 2,5 года эксплуатации требуется повышение мощности процессора и увеличение объёма оперативной памяти на сервере БД, и после введения конструктивных изменений требуется дополнительный расчёт энтропии информационного обеспечения на ближайшие 2 – 3 года.

Литература

1. Шеннон К. Работы по теории информации и кибернетике / К. Шеннон. – М. : Иностранная литература, 1963. – С. 259 – 268.
2. Юркевич Е. В. Введение в теорию информационных систем / Е. В. Юркевич. – М. : ИД «Технологии», 2007. – 272 с.
3. Методология IDEF1: [Электронный ресурс]. – Режим доступа <http://www.lcard.ru/>.
4. Метод моделирования информации : [Электронный ресурс]. – Режим доступа <http://www.idef.com/IDEF1.htm>.
5. Азаренкова Н. В., Проворова И. П. Модернизация энтропии в корпоративных информационных системах на основе детализации входящих потоков / Н. В. Азаренкова, И. П. Проворова // Аспирант и соискатель. – 2013.
6. Биномиальное распределение. – Режим доступа <http://algotlist.manual.ru/math/matstat/binomial/index.php>.
7. Современные компьютерные технологии управления информационно-аналитической деятельностью / Э. А. Трахтенгерц, Е. Л. Иванюков, Е. В. Юркевич. – М. : СИНТЕГ, 2007. – 371 с.
8. Липаев В. В. Функциональная безопасность программных средств / В. В. Липаев // М. : СИНТЕГ, 2004. – 348 с.

9. Берж К. Теория графов и ее применение: [перевод с франц.] / К. Берж; под ред. И.А. Вайнштейна. – М. : Иностранная литература, 1962. – 320 с.
10. Законы распределения дискретной случайной величины : [Электронный ресурс]. – Режим доступа <http://apollyon1986.narod.ru/docs /TViMS/NP/lekziiv/lekziya7.html>.

Поступила в редакцию 04.06.2013

Евгений Владимирович Юркевич, д-р техн. наук, профессор,
зав. лабораторией ИПУ им. В.А. Трапезникова РАН, т. (495) 334-88-70; e-mail: yurk@ipu.ru.
Наталья Викторовна Азаренкова, аспирант, т. 8 (916) 253-19-84, e-mail: Azarka1@yandex.ru.
Владимир Михайлович Долкарт, д-р техн. наук, главный научн. сотрудник,
т. (495) 624-86-65, e-mail: vniiem@orc.ru.