

СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ВЕРОЯТНОСТЕЙ ГИПОТЕЗ О СОСТОЯНИИ УЗЛОВОГО ТРАФИКА В ЗАДАЧАХ МОНИТОРИНГА

К.С. Ткаченко, А.А. Скидан

ФГАОУ ВО «Севастопольский государственный университет»
РФ, г. Севастополь, ул. Университетская, 33
E-mail: itiks@sevsu.ru

Рассматриваются вопросы, связанные с классификацией состояний трафика в узле в условиях переменной нагрузки, характерной для задач мониторинга. Предлагается решение с использованием многоместной классификации, что позволяет повысить статистическую устойчивость оценок. Приведен численный пример, подтверждающий эффективность предлагаемых решений.

Ключевые слова: узловой трафик, нестационарная нагрузка, статистическое оценивание гипотез.

Введение. Мониторинг критически-ми объектами основан на использовании многопроцессорных вычислительных систем и распределенных сред (PC), что позволяет обеспечить требуемые показатели надежности и гарантированности. Функционирование таких систем основано на решениях задач синхронизации параллельной и распределенной обработки данных в реальном масштабе времени [1].

Как отмечено в работе [2], существующие системы мониторинга «основанные преимущественно на физико-химических методах, являются трудоемкими и дорогостоящими, дают фрагментарные сведения, охватывают традиционно узкий круг загрязняющих агентов, не обеспечивают непрерывного мониторинга и своевременного обнаружения внезапного выброса загрязнений». Отсутствуют непрерывность, оперативность, полный цикл. В связи с этим, актуальными являются задачи статистического анализа информации о трафике в узлах распределенных сетей мониторинга и оценивания на ее основе вероятностей гипотез о нагрузке сети для свое-временного принятия решений об адекватности изменения производительности обработки данных в узлах. Источниками изменения интенсивности входных потоков информации в системах мониторинга является изменение частоты опроса датчиков в зависимости от степени опасности концентрации вредоносных загрязнителей, так и унифицированных уровней наблюдаемых количественных характеристик, изменение чувствитель-

ности датчиков. Это приводит к ряду ухудшений функционирования при неблагоприятных погодных условиях. С другой стороны, для оперативного комплексного контроля следует с высокой частотой мониторировать факторы, критические для окружающей среды.

Функционирование системы поддержки принятия решений об изменениях производительности обработки должно быть основано на использовании оценок стоимостных функций, матриц потерь и так далее. В свою очередь, задача их идентификации требует построения комплексов моделей, описывающих трафик в системе обработки данных на основе трафиков в отдельных ее узлах.

Узел во многих важных для инженерной практики случаях может быть описан как система массового обслуживания (СМО). Известны аналитические приближения для СМО типа M/M/K/N в нотации Кендалла. Но, к сожалению, аналитически полученные оценки системных характеристик не могут быть рассчитаны на целевых встроенных платформах в режиме жесткого реального времени. Для практических целей, к которым, несомненно, относится и управление узлом, в первом приближении достаточно получить статистические вероятности пребывания узла в том или ином состоянии. Таких состояний может быть несколько в зависимости от используемого подхода.

Число таких состояний для больших распределенных систем чрезвычайно велико, что, с одной стороны, существенно усложняет вычислительную за-

дачу, с другой стороны, приводит к неоправданной детализации. Поэтому целесообразна кластеризация множества состояний на подмножество в интересах решаемой задачи. Поэтому зададимся направлением по уменьшению этого числа состояний. Предельно упрощая, их всего два – рабочее и нерабочее.

Воспользуемся принципом разумного компромисса, например, применяя гипотезы с числом альтернатив более чем два [3]. При этом появляется возможность использования гипотез для большего числа состояний. Для простоты будут рассматриваться три состояния. Положим, что в результате наблюдения оценена загрузка системы

$$\rho = \frac{\lambda}{\mu} < 1, \text{ где } \lambda - \text{интенсивность вход-$$

ного простейшего потока заявок, μ – производительность обработки заявок. В зависимости от диапазона принимаемых значений ρ может быть отнесено к одному из трех интервалов обобщенных характеристик входного потока при фиксированной производительности $\mu = \text{const}$ и переменной интенсивности входного потока λ . В области изменения $\rho \in [\rho_0; \rho_{\lim}]$ выделяются две рабочие точки для решения задачи ρ_1 и ρ_2 , которые разбивают этот участок на три интервала и будем полагать, что если принадлежит:

- $\rho = (\rho_0; \rho_1]$ – поток ненагружен, узел тоже ненагружен;
- $\rho = (\rho_1; \rho_2]$ – поток нагружен, узел тоже нагружен;
- $\rho = (\rho_2; \rho_{\lim})$ – поток перегружен, узел тоже перегружен.

Например, как показывают результаты вычислительного анализа, проведенные далее, $\rho_1 = 0,4$, $\rho_2 = 0,6$.

Наша цель – разработать и исследовать подход для определения статистических оценок вероятностей гипотез о ненагруженности, нагруженности и перегруженности узла.

Краткое обобщение и аналитический анализ ряда публикаций. Для кластерных вычислительных систем подобная задача уже рассматривалась [4]. Существуют некоторые модели и анали-

тические оценки надежности обработки входного трафика. При моделировании резервируемых структур учитываются надежности резервов. Для использования в составе информационно-вычислительных центров расчет ведется по коэффициенту снижения производительности. В системах реального времени – на основе вероятности выполнения необходимого технологического цикла в заданный промежуток времени.

Известно рассмотрение возможных вариантов кластерных решений в зависимости от степени доступности и масштабируемости, например, при организации процессов безостановочного производства [5]. Сокращение запланированных и внеплановых простоев возможно благодаря переходу на различные узлы кластера в случае недоступности либо остановки аппаратной составляющей. Однако способ реализации решения задачи мониторинга трафика с использованием кластерных систем является весьма затратным.

В работе [6] предложены способы борьбы с перегрузками на основе алгоритма поиска варианта перераспределения нагрузки и его программная реализация. Решения важны для реальных систем компенсации перегрузок, которые в реальном времени создают дополнительную нагрузку. В условиях ограниченности пропускной способности каналов связи, их ненадежности и неоднородности, а также с низкой пропускной способностью звеньев сети, может быть увеличено количества одновременно подключенных приемных устройств.

В публикации [7] предлагается осуществить оценку информационной эффективности телекоммуникационной сети при использовании различных алгоритмов маршрутизации. Имитационным моделированием определяются основные характеристики процессов канального и сетевого уровня сегментов. К показателям эффективности относят производительность, среднюю временную задержку передаваемых пакетов, потери и коэффициент полезного действия. В полной мере мы не можем воспользоваться.

Рассмотрен подход к функционально-ориентированной узловой аппрокси-

мации задачи мониторинга [8]. Предложена классификационная нотация, которая может быть использована для отнесения анализируемой системы к одному из базовых классов. Описаны аналитические подходы к управлению узловой вычислительной мощностью.

Устойчивость определяется как минимальная сумма снижения пропускной способности тех линий, что приведут к необратимым отказам всех исходящих линий связи из узла-источника [9]. Когда внешний поток не превышает пропускную способность сети, выборочное снижение пропускной способности не определено.

Модель совместного распределения ресурсов и ассоциаций пользователей в качестве комбинационной игры с взаимо дополняемостью и взаимным влиянием, состоящий из корыстных игроков, заинтересованных исключительно в личном повышении производительности составляется в [10]. Предложенный механизм может значительно повысить эффективность функционирования сети с гетерогенными узлами и уменьшить отрицательное влияние различных эффектов, таких как аномалии физических адресов.

Выделение сетевых ресурсов рассматривается как проблема классического управления с обратной связью [11]. На основе результатов моделирования показано, что передаточная функция способна предоставить оптимальные значения распределения ресурсов для соседних узлов во время обычной, а также во время высокой степени перегрузки в сети.

Отказы, наступающие от перегрузки, в реально функционирующих сетях связи связаны с экстремальными значениями нагрузки, превышающей допустимый диапазон [12]. Использование графов Эрдеша-Реньи позволяет учесть влияние колебаний нагрузки.

Проблема определения состояния трафика ранее либо не рассматривалась, либо рассматривалась для двух разделенных состояний. Для случая трех и более состояний она является открытой.

Критериальная шкала: меньше нормы, норма и больше нормы. Согласно [3], критерий функциональной эффективности Е распознавания гипотез

о трафике в узле описывается масштабированной логарифмической зависимостью. Основываясь на результатах работы [3], критерий функциональной эффективности Е зависит от параметров: а – количество событий, состоящих в принадлежности распознаваемых реализаций к классу, характеризующему состояние процесса «норма», если они ему действительно принадлежат, б – «меньше нормы», если действительно, с – «больше нормы», если действительно, d – минимальное межклассовое кодовое расстояние.

Далее будем считать, что а – количество случаев нагруженного трафика, б – ненагруженного, с – перегруженного. Развивая работу [3] предлагаем выражение для d искать в виде

$$d = k \frac{0 + \rho_1 + \rho_1 + \rho_2 + \rho_2 + 1}{4}, \text{ то есть}$$

$$d = k \frac{2\rho_1 + 2\rho_2 + 1}{4}. \quad (1)$$

Такой вид (1) обусловлен его среднерифметичностью. Масштабный коэффициент k выбирается из условий существования логарифмических функций.

С целью дальнейшего анализа проведем вычислительные эксперименты по определению Е, фрагмент результатов сводится в табл. 1. Расчеты для этих целей велись при $\rho_1 = 0,4$, $\rho_2 = 0,6$, $k = 100$, и, соответственно, $d = 75$. Табл. 1 может быть визуализирована, причем для этого фрагмента считается $a=1$. Визуализацией является семейство параметрически заданных кривых показателя Е, параметром является b (рис. 1).

Анализ кривых на рис. 1 обнаруживает, что чем больше величина параметра семейства b, тем ниже график – ближе к оси абсцисс. Отсюда можно сделать вывод о том, что количество показателей ненагруженного трафика увеличивает степень прижатия к оси абсцисс. При некотором дальнейшем масштабировании, при изменении параметра с >> 5 с некоторой, наперед заданной точностью, кривые неразличимы.

После этого, на основе рассчитанного критерия Е можно уже получить гипотезы о том, что:

$$H_0 = \{\text{трафик узла ненагруженный}\} - \text{событие A};$$

$H_1 = \{\text{трафик узла нагруженный}\}$ – событие B;

$H_2 = \{\text{трафик узла перегруженный}\}$ – событие C.

Таблица 1. Результаты вычислительных экспериментов определения показателя эффективности E

№	a	b	c	E
1	1	1	1	2,2665
2	1	1	2	2,0933
3	1	1	3	1,9563
4	1	1	4	1,8424
5	1	1	5	1,7446
6	1	1	6	1,6586
7	1	1	7	1,5818
8	1	1	8	1,5123
9	1	1	9	1,4488
10	1	1	10	1,3902
11	1	2	1	2,0933
12	1	2	2	1,9563
13	1	2	3	1,8424
14	1	2	4	1,7446
15	1	2	5	1,6586
16	1	2	6	1,5818
17	1	2	7	1,5123
18	1	2	8	1,4488
19	1	2	9	1,3902
20	1	2	10	1,3358
21	1	3	1	1,9563
22	1	3	2	1,8424
23	1	3	3	1,7446
24	1	3	4	1,6586
25	1	3	5	1,5818

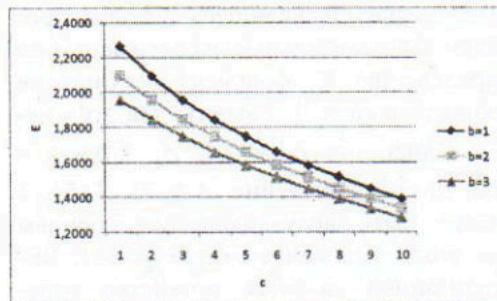


Рис. 1. Семейство параметрически заданных кривых критерия эффективности

Оценки вероятности этих гипотез крайне важны, например, для определения таких ключевых информационных структур как таблица поддержки принятия решений. Без них ничего нельзя сделать. Обозначается

$$f = a + b + c. \quad (2)$$

Тогда на основе (2) по известнейшей формуле получается:

$$P(H0) = \frac{a}{f}, \quad P(H1) = \frac{b}{f}, \quad P(H2) = \frac{c}{f}. \quad (3)$$

Выполним расчет по (3). Фрагмент результатов сводится в табл. 2.

Таблица 2. Точечные выборочные оценки вероятностей $P(H0)$, $P(H1)$, $P(H2)$

№	a	b	c	f	$P(H0)$	$P(H1)$	$P(H2)$
1	1	1	1	3	0,3333	0,3333	0,3333
2	1	1	2	4	0,2500	0,2500	0,5000
3	1	1	3	5	0,2000	0,2000	0,6000
4	1	1	4	6	0,1667	0,1667	0,6667
5	1	1	5	7	0,1429	0,1429	0,7143
6	1	1	6	8	0,1250	0,1250	0,7500
7	1	1	7	9	0,1111	0,1111	0,7778
8	1	1	8	10	0,1000	0,1000	0,8000
9	1	1	9	11	0,0909	0,0909	0,8182
10	1	1	10	12	0,0833	0,0833	0,8333
11	1	2	1	4	0,2500	0,5000	0,2500
12	1	2	2	5	0,2000	0,4000	0,4000
13	1	2	3	6	0,1667	0,3333	0,5000
14	1	2	4	7	0,1429	0,2857	0,5714
15	1	2	5	8	0,1250	0,2500	0,6250

Для всех строк этой таблицы соблюдается условие нормировки вероятностей. Но этого недостаточно для инженерно-технических приложений. Поэтому для адекватного и непредвзятого описания информационной ситуации необходимо использовать формулу полной вероятности [13].

На основе рассчитанных значений можно получить возможность оценить вероятности гипотез:

$$\left\{ \begin{array}{l} P(A) = \sum_{i=0}^2 P(H_i) P(A|H_i), \\ P(B) = \sum_{i=0}^2 P(H_i) P(B|H_i), \\ P(C) = \sum_{i=0}^2 P(H_i) P(C|H_i) \end{array} \right. \quad (4)$$

В формуле (4) априорные величины $i=0,1,2$:

$P(A|H_i)$ – вероятность события A при условии, что гипотеза H_i имела место;

$P(B|H_i)$ – вероятность события B при условии, что гипотеза H_i имела место;

наконец, $P(C|H_i)$ – вероятность события C при условии, что гипотеза H_i имела место.

Эти самые априорные величины определяются экспертами – лицами, принимающими решения (ЛПР). О вкладе ЛПР в систему поддержки принятия решений (СППР) необходимо написать в публикации подробнее. Дело в том, что (5) можно использовать для управления

узлом, а именно, для построения СППР на их основе. С целью чего рассматривается схема алгоритма для итерационных человеко-машинных процедур принятия решений (ППР), и структура СППР.

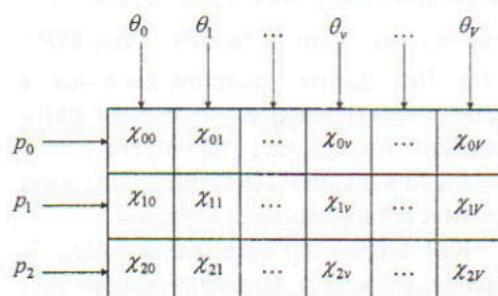


Рис. 2. Матрица принятия решений

На рис. 2 изображается матрица принятия решений. p_o , $o = 0, \dots, 2$ – вероятности выбираемых стратегий о состоянии узла. θ_v , $v = 0, 1, \dots$ – вероятности антагониста по отношению к рассматриваемому узлу. $X_{ov} = X(o, v)$ – частный оценочный функционал.

На рис. 3 некоторые циклические процессы для удобства восприятия и компактности упщены. А вот на рис. 4, в свою очередь, приводится структурная схема СППР.

На структурной схеме приводятся следующие подсистемы и их взаимосвязи:

- формирования оценок вероятностей
- формирует их по алгоритму на рис. 2;
- пользовательского интерфейса – взаимодействует с пользователем;

журнализирования – ведет протоколы;

формирования управляющих воздействий для узла – управляет целевым узлом;

настройки состояния – задает внутреннее состояние;

съема данных с узла – получает отклики с узла.

С целью иллюстрации на рис. 5 приводится ситуация характера ускорения диагностики в виде классической задачи теории вероятностей «два игрока». Эта задача связана с рассматриваемой, поскольку происходит разрешение антагонистического конфликта. Параметром семейства параметрически заданных кривых является вещественное p ($0,0 < p < 1,0$), аргументом является целое положительное n . Расчеты ведутся по формуле:

$$y(n) = 1 - (1 - p)^n. \quad (5)$$

В формуле (5) величина аргумента n изменяется в соответствии с интервальным выражением $n=1, 2, \dots, 100$, а параметр $p = p=0,1, 0,2, \dots, 0,9$.



Рис. 3. Схема алгоритма ППР



Рис. 4. Структурная схема СППР

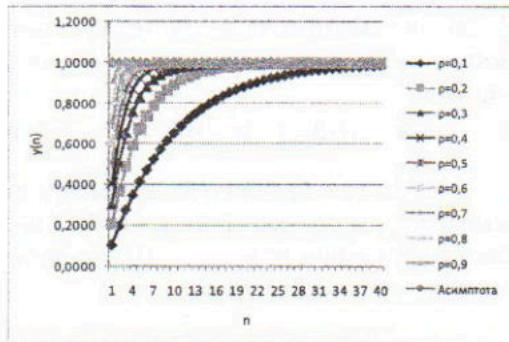


Рис. 5. Семейство параметрически заданных кривых $y(n)$, параметр $p=0,1,0,2,\dots,0,9$

Далее положим, что функция плотности вероятности для $P(A)$, $P(B)$ и $P(C)$ подчинена нормальному закону распределения. Это означает, что эти самые величины можно рассчитать по методологии многономенклатурных производственных систем, опираясь на плотность [14]. С целью упрощения громоздкости,

обозначается $Pl(x, \mu, \sigma) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$ — функция плотности, и

$Int(\mu, \sigma, x_{\min}, x_{\max}) = \int_{x_{\min}}^{x_{\max}} Pl(x, \mu, \sigma)dx$ — интеграл на отрезке $[x_{\min}, x_{\max}]$ от функции плотности. Функция $Int(\mu, \sigma, x_{\min}, x_{\max})$ не может быть выражена через элементарные функции. Наилучшим аналитическим результатом для нее является

$$\int_{x_{\min}}^{x_{\max}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \frac{1}{2} \left[\operatorname{erf}\left(\frac{x_{\max}-\mu}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{x_{\min}-\mu}{\sigma\sqrt{2}}\right) \right], \text{ где } \operatorname{erf} -$$

неэлементарная функция ошибок (интеграл вероятностей). Пусть полагаемые матожидания для $P(A)$, $P(B)$ и $P(C)$ составляет, соответственно, m_A , m_B , m_C . Также пусть три полагаемые сигмы не превышают границ отрезков и располагаются в отрезках для $P(A)$ $[\varpi_1; \varpi_2]$, для $P(B)$ $[\varpi_2; \varpi_3]$, и, в конце концов, для $P(C)$ $[\varpi_3; \varpi_4]$. Оттого сразу получается:

$$P(A) = Int(m_A, \sigma, \varpi_1, \varpi_2), \\ P(B) = Int(m_B, \sigma, \varpi_2, \varpi_3), \quad (6)$$

$$P(C) = Int(m_C, \sigma, \varpi_3, \varpi_4).$$

Например, для предельно распределенного и разнесенного по отрезкам случая, проводится серия вычислений, когда для известных m_A , m_B , m_C , ϖ_1 , ϖ_2 , ϖ_3 , ϖ_4 , σ , определяются $P(A)$, $P(B)$, $P(C)$. Результаты расчетов сводятся в табл. 3. Рекуррентное увеличение матожиданий ($m_B = m_A + m_A$; $m_C = m_B + m_A$) обусловлено влиянием смешения различных плотностей на границах отрезков.

Как видно из содержания табл. 3, оцененные в ходе вычислительного эксперимента вероятности в общем случае не образуют полную группу событий, то есть $P(A) + P(B) + P(C) \neq 1$, и не являются полностью независимыми.

О необходимости использования трех состояний вместо двух. Пусть рассматриваются две системы. В первой системе есть состояния S_0 , S_1 . Состояние S_0 — нормальное состояние, S_1 — неработоспособное состояние. Во второй системе есть состояния S_0 , S_1 , S_2 . Для нее S_0 — нормальное состояние, S_1 — пограничное состояние, когда беда близко, S_2 — неработоспособное состояние.

Для первой системы вероятности a_1 для перехода из S_0 в S_1 , a_2 — из S_1 в S_0 . Для второй системы вероятности b_1 для перехода из S_0 в S_1 , b_2 — из S_1 в S_0 , b_3 — из S_1 в S_2 , b_4 — из S_2 в S_1 . Графы переходов для процессов очевидны. Но могут быть в дальнейшем легко детализированы при усложнении. Составляются системы уравнений Колмогорова для финальных вероятностей состояний p_0 , p_1 , p_2 .

Первая система описывается как:

$$S_0 : p_0 = a_2 p_1, \\ S_1 : p_1 = a_1 p_0, \quad (7) \\ p_0 + p_1 = 1.$$

Вторая система описывается как:

$$S_0 : p_0 = b_2 p_1, \\ S_1 : p_1 = b_1 p_0 + b_4 p_2, \quad (8) \\ S_2 : p_2 = b_3 p_1, \\ p_0 + p_1 + p_2 = 1.$$

Таблица 3. Результаты расчетов по (6)

№	m_A	m_B	m_C	σ	w_1	w_2	w_3	w_4	$P(A)$	$P(B)$	$P(C)$
1	5	10	15	3	1	9	17	25	0,8176	0,6207	0,2521
2	5	10	15	3	10	18	26	34	0,0478	0,0038	0,0001
3	5	10	15	3	19	27	35	43	0,0000	0,0000	0,0000
4	5	10	15	5	1	9	17	25	0,5763	0,4985	0,3218
5	5	10	15	5	10	18	26	34	0,1540	0,0541	0,0138
6	5	10	15	5	19	27	35	43	0,0025	0,0003	0,0000
7	5	10	15	7	1	9	17	25	0,4323	0,3981	0,3110
8	5	10	15	7	10	18	26	34	0,2059	0,1154	0,0547
9	5	10	15	7	19	27	35	43	0,0219	0,0074	0,0021
10	15	20	25	3	1	9	17	25	0,0227	0,1585	0,4962
11	15	20	25	3	10	18	26	34	0,7936	0,7248	0,3681
12	15	20	25	3	19	27	35	43	0,0912	0,0098	0,0004
13	15	20	25	5	1	9	17	25	0,1125	0,2603	0,4452
14	15	20	25	5	10	18	26	34	0,5671	0,5404	0,3848
15	15	20	25	5	19	27	35	43	0,2037	0,0794	0,0226
16	15	20	25	7	1	9	17	25	0,1729	0,2761	0,3735
17	15	20	25	7	10	18	26	34	0,4284	0,4168	0,3439
18	15	20	25	7	19	27	35	43	0,2406	0,1426	0,0715
19	25	30	35	3	1	9	17	25	0,0000	0,0000	0,0004
20	25	30	35	3	10	18	26	34	0,0098	0,0912	0,3681
21	25	30	35	3	19	27	35	43	0,7248	0,7936	0,4962
22	25	30	35	5	1	9	17	25	0,0007	0,0046	0,0226
23	25	30	35	5	10	18	26	34	0,0794	0,2037	0,3848
24	25	30	35	5	19	27	35	43	0,5404	0,5671	0,4452
25	25	30	35	7	1	9	17	25	0,0108	0,0303	0,0715
26	25	30	35	7	10	18	26	34	0,1426	0,2406	0,3439
27	25	30	35	7	19	27	35	43	0,4168	0,4284	0,3735

Исключая ненужные уравнения из систем (7) и (8) и перенося члены, можно их решить методом обратной матрицы любой математической САПР, в итоге:

$$\begin{pmatrix} 1-a_2 \\ 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{a_2}{a_2+1} \\ \frac{1}{a_2+1} \end{pmatrix}. \quad (9)$$

$$\begin{pmatrix} 1-b_2 & 0 \\ 0 & b_3 \\ 1 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{b_2}{b_2+b_3+1} \\ \frac{1}{b_2+b_3+1} \\ \frac{b_3}{b_2+b_3+1} \end{pmatrix} \quad (10)$$

В формулах (9) и (10) положим равновероятность переходов, что характерно для непрерывно атакуемых, но качественно и своевременно восстанавливаемых узлов. Тогда $a_2=0,5$ и $b_2=b_3=0,25$. Рассчитанный таким образом вектор вероятностей состояний для первой системы составляет $\left(\frac{1}{3} \frac{2}{3} \frac{1}{3}\right)^T$, а для второй

$\left(\frac{1}{6} \frac{2}{3} \frac{1}{6}\right)^T$. Таким образом, видно, что не-

работоспособной состояние для второй системы меньше, чем для первой

$\left(\frac{2}{3} > \frac{1}{6}\right)$. Это демонстрирует важность использования трех состояний.

Дальнейшее увеличение числа градаций оцениваемого признака, то есть увеличение числа состояний, не всегда оправдано. Можно повысить требования к точности распознавания до предела. Но это, в свою очередь, приведет к большим затратам на контроль и мониторинг, к структурной и аппаратной избыточности. В предельном случае, затраты на контроль во много раз превышают расходы на компенсацию потерь от потоков событий.

Распознавание состояний трафика в условиях аддитивных случайных величин. Обозначается: НСВ – непрерывная случайная величина; НРНСВ – нормально распределенная НСВ; РРНСВ – равномерно распределенная НСВ; ЭРНСВ – экспоненциально распределенная НСВ.

Пусть a и b – нижняя и верхняя граница для РРНСВ, а λ – параметр ЭРНСВ. Тогда известно, что плотности функций распределения сверток двух НСВ

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{\infty} f_{\xi_1}(t) f_{\xi_2}(x-t) dt, \quad \text{которые}$$

можно рассчитать в системе компьютерной алгебры.

Плотность распределения НСВ = НРНСВ + НРНСВ есть

$$\int_{-\infty}^{\infty} \frac{\exp\left(-\frac{t^2}{2}\right) \exp\left(-\frac{1}{2}(x-t)^2\right)}{\sqrt{2\pi} \sqrt{2\pi}} dt = \frac{\exp\left(-\frac{x^2}{4}\right)}{2\sqrt{\pi}}.$$

Плотность распределения НСВ = НРНСВ + РРНСВ есть

$$\int_a^b \frac{\exp\left(-\frac{t^2}{2}\right)}{\sqrt{2\pi}(b-a)} dt = \frac{\operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right)}{2a-2b}.$$

Плотность распределения НСВ = НРНСВ + ЭРНСВ есть

$$\begin{aligned} \int_0^x \frac{\exp\left(-\frac{t^2}{2}\right)}{\sqrt{2\pi}} (\lambda \exp(-\lambda(x-t))) dt = \\ = \frac{1}{2} \lambda \left[\operatorname{erf}\left(\frac{\lambda}{\sqrt{2}}\right) + 1 \right] \exp\left(\frac{1}{2}\lambda(\lambda-2x)\right). \end{aligned}$$

Плотность распределения НСВ = РРНСВ + РРНСВ есть

$$\int_a^b \frac{1}{(b-a)(b-a)} dt = \frac{1}{b-a}.$$

Плотность распределения НСВ = РРНСВ + ЭРНСВ есть

$$\int_a^b \frac{\lambda \exp(-\lambda(x-t))}{b-a} dt = \frac{\exp(-\lambda x)(\exp(a\lambda) - \exp(b\lambda))}{a-b}$$

($a > 0$).

Плотность распределения НСВ = ЭРНСВ + ЭРНСВ есть (поменять x на t)

$$\int_0^x (\lambda \exp(-\lambda t))(\lambda \exp(-\lambda(x-t))) dt = \lambda^2 \exp(-\lambda x).$$

Использование этих плотностей вероятностей может улучшить результат распознавания для смесей НСВ. Так, например, можно построить таблицы о видах и типах узлового трафика. Строки соответствуют видам, столбцы соответствуют типам. ЛПР, пользуясь табл. 4, может, позиционируясь на пересечение нужной строки и столбца (вида и типа), получить формулу. При более детальном рассмотрении трафика в узле можно выделить несколько фаз: буферирование, обработка, передача. Для каждой фазы различаются оценки функций распределения. Тогда общее время пребывания в узле находится как интеграл свертки из таблицы в зависимости от ситуации, причем для разных гипотез задействованы разные пределы интегрирования.

Таблица 4. Виды и типы узлового трафика

		Тип узлового трафика		
		НРНСВ	РРНСВ	ЭРНСВ
Вид узлового трафика	НРНСВ	$\frac{\exp\left(-\frac{x^2}{4}\right)}{2\sqrt{\pi}}$	$\frac{\operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right)}{2a-2b}$	$\frac{1}{2} \lambda \left[\operatorname{erf}\left(\frac{\lambda}{\sqrt{2}}\right) + 1 \right] \exp\left(\frac{1}{2}\lambda(\lambda-2x)\right)$
	РРНСВ	$\frac{\operatorname{erf}\left(\frac{a}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{b}{\sqrt{2}}\right)}{2a-2b}$	$\frac{1}{b-a}$	$\frac{\exp(-\lambda x)(\exp(a\lambda) - \exp(b\lambda))}{a-b}$
	ЭРНСВ	$\frac{1}{2} \lambda \left[\operatorname{erf}\left(\frac{\lambda}{\sqrt{2}}\right) + 1 \right] \exp\left(\frac{1}{2}\lambda(\lambda-2x)\right)$	$\frac{\exp(-\lambda x)(\exp(a\lambda) - \exp(b\lambda))}{a-b}$	$\lambda^2 \exp(-\lambda x)$

Заключение. Задача оценивания параметров трафика в узле является по-прежнему актуальной, так как от этого зависит результативное решение задач мониторинга критических объектов. В основе предлагаемого решения лежит использование трехточечного метода, что позволяет получить большее правдоподобие и достоверность.

Использование предложенного подхода для контроля состояний окружающей среды способствует разрешению ряда возникающих задач, например, по получению апостериорных аналитических оценок вероятностей принятия ги-

потез о текущем состоянии изолированной системы экологического мониторинга на прикладном уровне, а именно, о нахождении степени отклика в одной из трех допустимых градаций по возможному наносимому ущербу.

СПИСОК ЛИТЕРАТУРЫ

1. Охтилев М.Ю., Соколов Б.В., Юсупов Р.М. Теоретические и технологические основы концепции проактивного мониторинга и управления сложными объектами // Известия ЮФУ. Технические науки. С. 162–174.

2. Гайский П.В., Трусевич В.В., Забурдаев В.И. Автоматический биоэлектронный комплекс, предназначенный для раннего обнаружения отравляющих загрязнений пресных и морских вод // Морской гидрофизический журнал. 2014. № 2. С. 44–53.
3. Довбыш А.С., Берест О.Б. Трёхальтернативная обучающаяся система поддержки принятия решений для автоматизации технологического процесса // Вестник Томского государственного университета. Томск, 2014. № 4 (29). С. 31–40.
4. Волков В.А., Чудинов С.М. Метод оценки надежности кластерных вычислительных структур и отказоустойчивости приложений с недетерминированным поведением // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. 2011. № 13-1 (108), Вып. 19/1, С. 121–126.
5. Ренев Р.О. Кластеризация как способ оптимизации информационных ресурсов // Спецтехника и связь. 2012. № 1 С. 62–63.
6. Манакова И.П. Борьба с перегрузками в системах онлайн-вещания // Технические науки – от теории к практике. 2015. № 4 (41). С. 16–27.
7. Литвинов К.А., Пасечников И.И. Информационная эффективность модели телекоммуникационной сети при различных алгоритмах маршрутизации с использованием параметра кибернетической мощности телекоммуникационной сети // Вестник Тамбовского университета. Серия: Естественные и технические науки. 2015. № 1. С. 232–237.
8. Функционально-ориентированная узловая аппроксимация задачи мониторинга распределенных сред / А.В. Скатков, К.С. Ткаченко [и др.] // Системы контроля окружающей среды. Севастополь: ИПТС. 2016. Вып. 4 (24). С. 42–48.
9. Yazicioglu A.Y., Rozbehani M., Dahleh M.A. Resilience of Locally Routed Network Flows: More Capacity is Not Always Better // arXiv:1608.04155v1 [cs.SY] 14 Aug 2016 (дата обращения: 01.09.2016).
10. Touati M. Controlled Matching Game for Resource Allocation and User Association in WLANs / Mikael Touati, Rachid El-Azouzi, Marceau Coupechoux, Eitan Altman and Jean-Marc Kelif // arXiv:1510.00931v2 [cs.GT] 2 May 2016 (дата обращения: 01.09.2016).
11. Singha N., Gupta R., Singh Y.N. Resource Allocation in Peer-to-Peer Networks: A Control Theoretical Perspective // arXiv:1509.07989v2 [cs.NI] 24 Dec 2015 (дата обращения: 01.09.2016).
12. Mizutaka S., Yakubo K. Robustness of scale-free networks to cascading failures induced by fluctuating loads // arXiv:1507.00121v1 [physics.soc-ph] 1 Jul 2015 (дата обращения: 01.09.2016).
13. Венцель Е.С. Теория вероятностей. М.: Высш. шк., 1998. 576 с.
14. Скатков А.В., Балакирева И.А. Обеспечение гарантированности системы управления многофункциональным производством КМОП // Радиоэлектронные и компьютерные системы, 2009. № 6 (40). С. 26–31.

STATISTICAL ESTIMATION OF HYPOTHESES PROBABILITIES FOR MONITORING TASKS OF NODAL TRAFFIC STATE

K.S. Tkachenko, A.A. Skidan

Federal State Educational Institution of Higher Education «Sevastopol State University»
Russian Federation, Sevastopol, Universitetskaya St., 33

We discuss problems associated with classifications of nodal traffic under the conditions of variable loads are typical for monitoring tasks. The solution using multi-classification which improves the statistical stability of the estimates proposed. There is the numerical example that confirms efficiency of the proposed solutions.

Keywords: nodal traffic, variable load, hypotheses statistical evaluation.