УДК 681.324

А.В. СКАТКОВ, Д.Ю. ВОРОНИН, Д.Н. ДАНИЛЬЧУК

Севастопольский национальный технический университет, Украина

АНАЛИЗ ГАРАНТОСПОСОБНОСТИ РАСПРЕДЕЛЕННЫХ СИСТЕМ С АДАПТИВНЫМ БАЙЕСОВСКИМ УПРАВЛЕНИЕМ

Рассмотрена задача адаптивного управления распределенной подсистемой обработки и компрессии данных с целью выполнения QoS-требований для среды коммутации и обеспечения таких необходимых составляющих гарантоспособности системы мониторинга, как достоверность и готовность. Приведен алгоритм решения поставленной задачи. Полученные результаты подтверждают актуальность применения данного подхода.

гарантоспособность, достоверность, готовность, адаптивное управление, метод Байеса, распределенная система, статистический метод распознавания

Введение

Тенденция вхождения компьютерных технологий в жизнь человека заставляет признать тот факт, что обеспечение гарантоспособности компьютерных систем является крайне важной задачей. Известно, что гарантоспособные системы обладают рядом свойств, соблюдение которых исключает существенные материальные убытки и катастрофы различного масштаба [1]. Особенно важной становится задача обеспечения гарантоспособности систем мониторинга, так как неполадки при работе непосредственно приводят к дезорганизующим воздействиям на объект наблюдения.

Распределенные системы в общем случае представляют собой сложные технические комплексы, объединенные между собой средой коммутации, например, компьютерной сетью. Для выполнения возложенных функций распределенная система должна не только предоставлять пользователю необходимые услуги, но и обеспечивать их должное качество – "качество обслуживания" (Quality of Service, QoS) [2]. Упрощенно QoS требования содержат в себе следующие составляющие:

1) передача информации за минимальное время и строго по назначению;

2) достоверность и надежность передаваемой информации.

Эти требования соответствуют таким свойствам гарантоспособности, как достоверность, безотказность, конфиденциальность, готовность [1]. В рамках данной статьи предлагается подход, который поможет адаптивно минимизировать объем передаваемой информации и тем самым не только обеспечить выполнение QoS-требований для среды коммутации, но и обеспечит такие необходимые составляющие гарантоспособности системы мониторинга, как достоверность и готовность. Задачу контроля за состоянием компьютерной сети принято разделять на две подзадачи: мониторинг и анализ. Задача анализа состоит в осмыслении полученных при мониторинге данных и в разработке предложений по улучшению работы сети. В современных исследованиях по разработке средств контроля основной упор делается на задачу анализа. Однако не следует недооценивать задачу организации мониторинга, так как решение этой задачи нетривиально и состоит в нахождении компромисса между полнотой, частотой мониторинга и его стоимостью. Таким образом, разработка структур и логики работы различных систем мониторинга является крайне **актуальной** задачей.

Постановка задачи

Описательно постановка задачи состоит в следующем. Пусть имеется система мониторинга (рис. 1). Задачей рассматриваемой системы мониторинга

(СМ) является наблюдение за состоянием узлов компьютерной сети (УС). В качестве инициатора запуска системы мониторинга может выступать сам УС, либо администратор сети. В первом случае УС самостоятельно отправляет информацию о своем состоянии СМ. Во втором случае УС получает команду на отправку указанных данных. Данные, отправляемые УС, будем называть информационным пакетом (ИП). ИП от УС передается на группу узловых коммутаторов (УК), которые распределяют поступившие пакеты между несколькими настраиваемыми информационными фильтрами (ИФ). ИФ предназначены для выделения из ИП определенной информации, в соответствии с предустановленной программой. Управление ИФ осуществляется блоком управления информационными фильтрами, являющимся отдельным блоком или частью центрального блока управления.

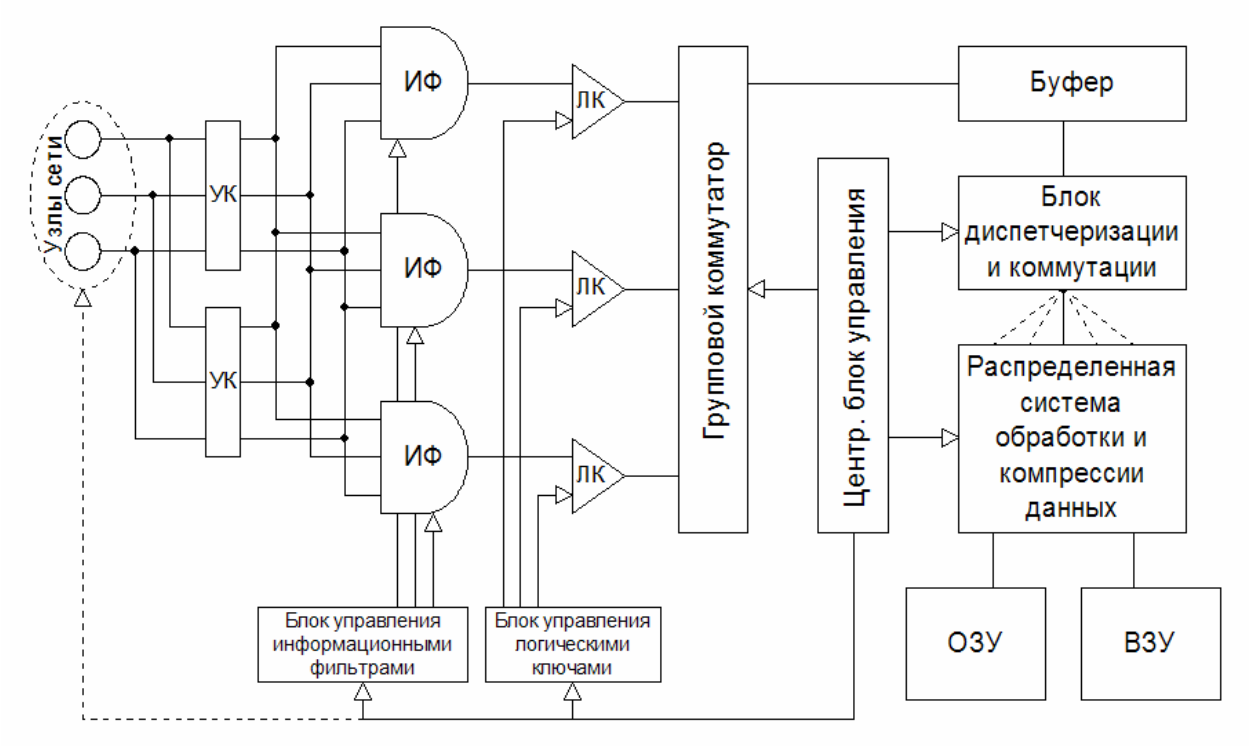

Рис. 1. Структурная схема системы мониторинга

Информацию, получаемую на выходе ИФ независимо от содержания, будем называть блоком данных. Управление перемещением блоков данных осуществляется при помощи блока управления логическими ключами, который определяет моменты времени, в течение которых блок данных передается в групповой коммутатор через логические ключи. Групповой коммутатор перемещает блок данных через буфер в блок диспетчеризации и коммутации (блок ДиК).

Функционирование группового коммутатора определяется центральным блоком управления. Из буфера блок данных через блок ДиК передается в распределенную систему обработки и компрессии данных (РСОиКД). В РСОиКД происходит:

1) обработка блока данных, после чего получается информация, которую будем называть информацией мониторинга;

2) уплотнение информации мониторинга для дальнейшей записи либо в оперативную память (ОЗУ) либо во внешнюю память (ВЗУ).

Управление РСОиКД также осуществляется центральной системой управления. Информация мониторинга, помещенная в ОЗУ, может использоваться в обработке поступивших блоков данных, а помещенная в ВЗУ, используется для накопления статистики о состоянии УС.

Распределенную систему обработки и компрессии данных можно представить совокупностью уз-

лов, осуществляющих обработку и компрессию информации. Для каждого такого узла необходимо решить задачу планирования и управления. В качестве УУ выступает центр диспетчеризации и управления. Входными данными для центра является первичный поток данных, состоящий из блоков данных системы мониторинга. На выходе имеем обработанный и сжатый поток данных. Степень компрессии зависит от условий дальнейшего хранения и использования информации. Если планируется, что сжатый блок будет записан в ОЗУ и будет использоваться в обработке поступивших блоков данных, то степень его сжатия должна быть минимальна, так как это повлияет на такой показатель гарантоспособности системы мониторинга, как готовность. Сжатый блок, помещаемый в ВЗУ, используемый для накопления статистики о состоянии УС, должен быть сжат с максимальной степенью компрессии. Эту информацию следует хранить, так как она обеспечивает такое свойство гарантоспособности системы мониторинга, как достоверность. В процессе обработки также будет проявляться вредная информация, имеющая дезорганизующее воздействие. Таким образом, на выходе центра диспетчеризации и управления должна присутствовать совокупность команд, необходимая для утилизации вредной информации из системы мониторинга. При функционировании центра диспетчеризации и управления используются условные вероятности, полученные при помощи подхода Байеса.

Математическая модель

Известно, что основным преимуществом статистических методов распознавания является возможность одновременного учета признаков различной физической природы, так как они характеризуются безразмерными величинами – вероятностями их появления при различных состояниях системы [3]. Метод, основанный на обобщенной формуле Байеса, относится к методам технической диагностики и популярен благодаря простоте и эффективности. К недостаткам можно отнести большой объем предварительной информации и «угнетение» редко встречающихся диагнозов.

Как видно из постановки задачи центр управления для каждого узла системы принимает решение:

1) обработанный блок информации сжать для записи в ОЗУ;

2) обработанный блок информации сжать для записи в ВЗУ;

3) утилизировать обработанный блок информации.

Согласно подходу Байеса, эти три ситуации представляют собой множество диагнозов:

$$D = \begin{cases} D_1 - \text{сжать для ОЗУ}; \\ D_2 - \text{сжать для ВЗУ}; \\ D_3 - \text{утилизировать}. \end{cases} \quad (1)$$

Аналогично можно выделить множество признаков (симптомов):

$$k_q = \begin{cases} k_{q_1} - \text{необходим для обработки}; \\ k_{q_2} - \text{необходим для статистики}; \\ k_{q_3} - \text{вреден}. \end{cases} \quad (2)$$

В качестве входных данных для подхода Байеса имеется вектор априорной вероятности выбора диагноза $P(D_i)$ и матрица условных вероятностей появления признаков вида:

	$k_{q1}/\ D_i$	$k_{q2}/\ D_i$	$k_{q2}/\ D_i$
D_1	…	…	…
D_2	…	…	…
D_3	…	…	…

Элементом матрицы является $P(k_j / D_i)$ – вероятность появления признака k_j при получении диагноза D_i.

В нашем случае имеется только одно подмножество признаков ($q = 1$). Однако при наличии нескольких подмножеств, количество матриц условных вероятностей для признаков также будет увеличиваться.

По формуле Байеса можно найти вероятность выбора D_i при наличии k_j.

$$P(D_i / k_j) = \frac{P(D_i) \cdot P(k_j / D_i)}{\sum_s P(D_s) \cdot P(k_j / D_s)}. \quad (3)$$

Получив $P(D_i / k_j)$, можно выбрать решение как $\max_i P(D_i / k_j)$.

Можно еще ввести пороговое значение $P^{П}(D_i)$ для того, чтобы обрабатывать ситуацию равновероятных решений. Однако с введением пороговой функции целесообразно повременить.

В нашем случае имеется только одно подмножество признаков ($q = 1$). Но следует упомянуть, что при наличии нескольких подмножеств справедлива формула

$$P(D_i / k_{qj}) = \frac{P(D_i) \cdot \prod_j P(k_{qj} / D_i)}{\sum_s P(D_s) \cdot \prod_j P(k_{qj} / D_s)}, \quad (4)$$

где q – количество подмножеств признаков.

Применение подхода Байеса подразумевает модификацию вектора априорной вероятности выбора диагноза $P(D_i)$ и матрицы условных вероятностей признаков $P(k_j / D_i)$. Согласно [3], необходимо ввести следующие величины:

N – количество принятых решений;

N_{ij} – количество принятых решений с использованием диагноза D_i на основе признака k_j.

Если принимается решение с использованием диагноза D_μ, то производятся поправки по формулам:

$$P(D_i) = \begin{cases} P(D_i)\dfrac{N}{N+1}; i = 1,2,...,n; i \neq \mu; \\ P(D_\mu)\dfrac{N}{N+1} + \dfrac{1}{N+1}; i = \mu, \end{cases} \quad (5)$$

$$P(k_{js} / D_\mu) = \begin{cases} P(k_{js} / D_\mu)\dfrac{N_{\mu j}}{N_{\mu j}+1}; s \neq r; \\ P(k_{js} / D_\mu)\dfrac{N_{\mu j}}{N_{\mu j}+1} + \dfrac{1}{N_{\mu j}+1}; s = r. \end{cases} \quad (6)$$

Адекватную работу программной реализации подхода Байеса проиллюстрируем с помощью графиков, изображенных на рис. 2. Пример отображает зависимость изменения вектора вероятностей диагнозов от истории полученных признаков. В качестве вектора признаков имеем: 22222111333333.

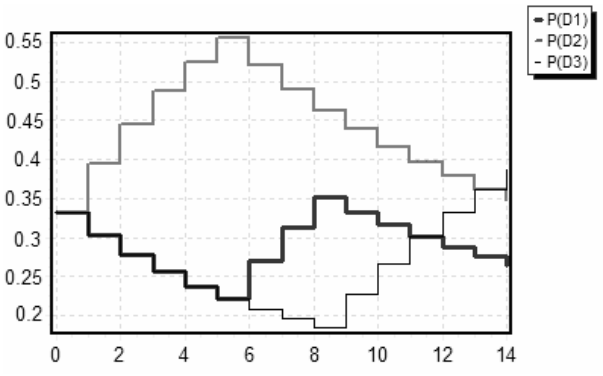

Рис. 2. График изменения вероятностей появления диагнозов от истории поступления признаков

Как видно из графика, при получении второго признака происходит пропорциональное увеличение вероятности второго диагноза и пропорциональное уменьшение вероятности двух оставшихся диагнозов. Это продолжается до тех пор, пока не встречается первый признак (6 шаг). Начиная с шестого шага и до девятого, происходит рост вероятности первого диагноза и уменьшение двух других. Начиная с девятого шага происходит рост вероятности третьего диагноза и уменьшение двух других.

Для оценки эффективности предложенного подхода введем функцию:

$$W = \frac{t_{ПР}}{t_{ПР\max}} + \frac{V_{ПД}}{V_{ПД\max}} + \frac{\sum k \neq d}{n} \to \min,$$

где $\dfrac{t_{ПР}}{t_{ПР\max}}$ – степень готовности процессора;

$\dfrac{V_{ПД}}{V_{ПД\max}}$ – степень готовности среды передачи данных;

$\dfrac{\sum (k \neq d)}{n}$ – степень достоверности диагнозов.

В табл. 1 приведено влияние диагнозов на увеличение различных составляющих оценочной функции *W*.

Таблица 1

Влияние диагнозов на увеличение различных составляющих оценочной функции

	Коэффициент готовности процессора	Коэффициент готовности среды передачи данных	Степень достоверности диагнозов
D_1	15	0	7
D_2	0	17	7
D_3	4	0	7

Таблица 2

Результаты исследования

№ шага	Признак	Метод Байеса		Random	
		Диагноз	Эффективность	Диагноз	Эффективность
1	2	2	17	2	17
2	2	2	34	1	39
3	2	2	51	1	61
4	2	2	68	2	78
5	2	2	85	2	95
6	1	1	100	1	11
7	1	1	115	2	134
8	1	1	130	1	149
9	3	3	134	2	173
10	3	3	138	1	195
11	3	3	142	1	217
12	3	3	146	2	241
13	3	3	150	1	263
14	3	3	154	1	285

В табл. 2 приведен расчет функции эффективности для случая с использованием метода Байеса, и для сравнения была рассчитана функция эффективности для случайного выбора диагноза.

Выводы

Как видно из полученных результатов, эффективность метода Байеса на 30% выше, чем у случайной дисциплины выбора диагноза. Это говорит о том, что такие необходимые составляющие гарантоспособности, как достоверность и готовность были существенно улучшены. Следовательно, предлагаемый подход представляет ценность и является актуальным.

«В процессах управления сложными объектами адаптация занимает почетное место» [4]. В качестве дальнейшей работы планируется исследования других методов адаптации, в том числе основанных на нейронных сетях.

Литература

1. Харченко В.С. Гарантоспособность и гарантоспособные системы: элементы методологии // Радиоэлектронные и компьютерные системы. – № 5 (17). – С. 7-19.

2. Иванов А.Б., Соколов И.В. От сквозного контроля сети к контролю качества услуг. – [Электрон. ресурс]. – Режим доступа: http://www.syrus.ru.

3. Биргер И.А. Техническая диагностика. – М.: Машиностроение, 1978. – 240 с.

4. Растригин Л.А. Адаптация сложных систем. Методы и приложения. – Рига: Зинатне, 1981. – 375 с.

Поступила в редакцию 6.02.2007

Рецензент: д-р техн. наук, проф. Ю.К. Апраксин, Севастопольский национальный технический университет, Севастополь.