

вычислить интеграл
$$\varphi_k(\tau_i) = \int_{-\infty}^{\infty} \varphi_k(t_i) p(t/\tau_i) dt_i.$$

При некоторых конкретных видах $\varphi_k(t_i)$ этот интеграл достаточно просто вычисляется. Зная величины $\varphi_k(\tau_i)$, можно получить и явный вид оценок $\hat{\theta}_k$ параметров θ_k .

Как видно из изложенного выше, во всех трех ситуациях необходимо выполнять идентичные вычислительные действия, связанные с обработкой массивов векторов и матриц для получения оценок параметров тренда с последующим нахождением матрицы ковариаций этих оценок. Последняя позволяет провести анализ качества полученных оценок. Применение технологий многоагентных систем к программной реализации матричных алгоритмов позволит не только провести параллельные вычисления оценок трендов временных рядов для различных ситуаций, но и выбрать ту из них, которая адекватна реальным данным. Выбор сопровождается анализом качества оценок и сообщением этой информации пользователю.

5. ЗАКЛЮЧЕНИЕ

Создание программного обеспечения анализа временных рядов при случайных моментах измерений с использованием многоагентных технологий расширит возможности организации мониторинга экономических процессов. Последние отличает большая подвижность вследствие структурных сдвигов в народном хозяйстве, смены механизма управления, изменения склонности к политике потребления и сбережения, движения цен, научно-технического прогресса, кредитно-финансовой политики. Погружение

будущего специалиста в реальную среду современных информационных технологий должно происходить не только на занятиях по информатике и программированию, но и при изучении смежных дисциплин, связанных с обработкой информации. Выполнение курсовых работ в конкретных предметных областях, в частности по анализу временных рядов, с использованием современных информационных технологий способствует тому, что студент от теоретического знания технологий программирования переходит к их самостоятельному применению.

ЛИТЕРАТУРА

1. РААМ-97(1997) Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-agent Systems. London
2. Трахтенгерц Э. А. Взаимодействие агентов в многоагентных системах. // Автоматика и телемеханика, № 9, 2000 г.
3. Поттосина С.А., Терпугов А.Ф. Фильтрация дважды стохастических рекуррентных точечных процессов // Радиотехника, 1991.-№12.-с.20-25.
4. Поттосина С.А., Терпугов А.Ф. Линейная фильтрация случайных процессов при измерениях в случайные моменты времени // Изв.высш. учебн.завед., сер.Физика, 1994.-№2.-с.67-72.
5. Поттосина С. А., Трушкин С. Ю. Сравнительный анализ алгоритмов фильтрации флуктуирующих пуассоновских потоков // Автоматика и вычислительная техника, Мн.1993.
6. Идрисов Ф.Ф. Выделение трендов временных рядов при измерениях в случайные моменты времени // Изв.высш. учебн.завед., сер.Физика, 1995.-Т.38.-№3.-с.3-10.
7. Идрисов Ф.Ф. Выделение трендов временных рядов при наличии ошибок в измерениях моментов времени // Изв.высш. учебн.завед., сер.Физика, 1996.- Т.39.-№4.-с.11-16.

К ВОПРОСУ ОБНАРУЖЕНИЯ ПРЕДАВАРИЙНЫХ СОСТОЯНИЙ ТЕХНИЧЕСКИХ СИСТЕМ

В.В. Шахов

Институт Вычислительной Математики и Математической Геофизики Сибирского Отделения Российской Академии Наук, пр. Лаврентьева, 6, 630090, РОССИЯ, тел. (3832) 39-62-11, shakhov@rav.sccc.ru

АННОТАЦИЯ

Рассматривается статистическая задача обнаружения экстремального режима функционирования технической системы, обусловленного отказом некоторых компонент системы

при сохранении общей работоспособности. Предполагается, что объект может функционировать в предаварийном режиме некоторое фиксированное время без потерь.

Требуется обнаружить наступление указанного режима, при этом объявляя ложную тре-

вогу, реакция на которую влечет некоторые расходы, не чаще чем заданное число раз в течении определенного времени.

Таким образом, возникает задача о разладке с фиксированным допустимым запаздыванием и заданным уровнем ложной тревоги.

В отличие от традиционных постановок задачи о разладке, когда требуется минимизировать среднее время между моментами возникновения и объявления разладки, здесь требуется минимизировать вероятность пропуска разладки. Причем разладка считается обнаруженной, если сигнал о ее возникновении подается не позднее заданного времени.

1. ВВЕДЕНИЕ

В виду сложности современных технических систем становится актуальной задача обнаружения неисправности компонент системы по наблюдению некоторых косвенных признаков. Является распространенной ситуация, когда часть устройств системы вышла из строя, но общая работоспособность сохраняется.

В указанном состоянии система может функционировать в течении некоторого времени, по истечении которого, потери становятся ощутимыми. В этом случае, аварию обнаружить труднее, чем в случае полного отказа системы. Однако, наблюдая в различные моменты времени значения некоторой характеристики системы, получаем последовательность одинаково распределенных случайных величин, анализируя которую, можно сделать вывод о наступлении экстремального режима (разладка) по смене параметров распределения наблюдаемого случайного процесса.

Момент, когда происходит смена параметров распределения наблюдаемого процесса, называется моментом возникновения разладки.

Событие, когда сигнал о разладке объявляется до наступления разладки, называется ложной тревогой. Требуется объявлять ее, по возможности, редко.

Здесь нет нужды объявлять разладку как можно ближе к моменту возникновения неисправности, что достигается за счет увеличения доли ложных сигналов, обработка которых требует определенных затрат. Надо не пропустить разладку, по возможности избегая расходов на обработку ложных тревог. С точки зрения качества критерия, не важно, когда подается сигнал о разладке в пределах зоны отсутствия потерь.

2. ПРАКТИЧЕСКИЕ ПРИМЕРЫ

Приведем несколько примеров, для иллюстрации вышесказанного.

Пусть имеется сеть с коммутацией пакетов, состоящая из двух узлов, соединенных двумя каналами (например, коммутатор и хост, получающий сообщения, Рис.1).

После выхода из строя одного канала система продолжает функционировать, данные передаются без искажений, но обслуживание идет с меньшей интенсивностью.

Если коэффициент использования системы больше единицы, то через некоторое время буфера в узлах переполняются, и пакеты начнут теряться, возникнут потери. Наблюдая времена задержки пакетов в сети, получаем последовательность случайных величин для анализа.

Допустимая задержка обнаружения разладки определяется размером буфера

Лопасты турбин гидроэлектростанций подвержены износу. Определить, критична ли степень износа, можно визуально, осушив предварительно турбину, что дорого, да и невозможно делать часто по техническим причинам.

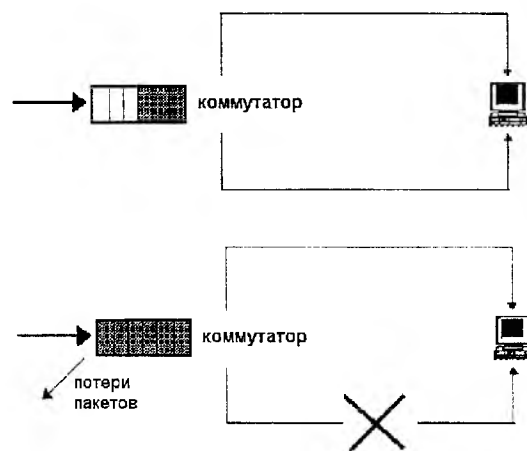


Рисунок 1. Сеть с коммутацией пакетов, состоящая из двух узлов

Более предпочтительно использовать косвенные характеристики, которые легко наблюдаются, в частности расход воды, вырабатываемая мощность и т.д. Считая значения характеристик независимыми случайными величинами, можно предположить гипотезу о наблюдении после разладки (возникновение микротрещин) того же самого закона распределения наблюдаемой величины, что и до, но с другим параметром.

Допустимая задержка обнаружения здесь может определяться временем эксплуатации

лопастей с момента появления трещин до откалывания кусков, приводящих к разрушению механизма.

Одной из наиболее распространенных и опасных атак на распределенные вычислительные системы является атака, называемая "отказ в обслуживании" [1] или Distributed Denial of Service (DDoS) [2]. Атакующим создается шторм ложных запросов на подключение, приводящий к перегрузке системных ресурсов и зависанию системы.

Наблюдая количество запросов в единицу времени, получаем последовательность случайных величин для анализа.

С учетом сложившейся практики моделирования потока запросов распределением Пуассона ([3], [4]), имеем в случае атаки тоже Пуассоновский поток, только с большей интенсивностью.

Допустимая задержка определяется ресурсом очереди запросов на сервере и производительностью системы.

Для решения аналогичных задач о разладке используют подход, называемый Алгоритмом кумулятивных сумм (АКС), впервые применявшийся в работах [5-7], который, вообще говоря, является многократно применяемым последовательным анализом Вальда [8].

Основной идеей АКС является сравнение кумулятивной суммы $g(n)$, которая вычисляется по наблюдению некоторого числа n случайных величин, с некоторым пороговым значением h , оптимальный выбор которого представляет собой сложную самостоятельную задачу. В данной статье используется аналогичный подход, хотя схема алгоритма отличается от модификаций АКС [9], [10].

В общем случае, оптимальное пороговое значение h определяется экспериментально [11], но с учетом специфики проблемы можно получить аналитическое выражение.

3. МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ, МЕТОД РЕШЕНИЯ

Пусть имеется последовательность случайных величин $\{x_i\}$, такая что $x_1, \dots,$

x_{m-1} имеют функцию распределения F_1 , а x_m, x_{m+1}, \dots функцию F_2 . m - момент разладки, T - допустимое запаздывание.

Наблюдая $\{x_i\}$ требуется определить, в какой момент r подать сигнал о разладке.

Если $m \leq r \leq m + T$, то разладка обнаружена. Если $r > m + T$, то пропущена. Если $r < m$, то имеет место ложная тревога.

Требуется минимизировать вероятность пропуска разладки

$$P(r > m + T | F_2),$$

при этом избегая частого объявления ложной тревоги.

Для решения задачи предлагается следующий алгоритм:

1. Вычисляется сумма

$$S_T = \sum_{i=1}^T x_i$$

2. Если выполняется решающее правило, останов. Иначе вычисляется следующая сумма

$$S_n = S_{n-1} + x_n - x_{n-T},$$

где $n=T+1, T+2, \dots$

3. Переход на шаг 2.

Решающее правило выглядит следующим образом: $S_n > h, n=T, T=1, \dots$ т.е. момент подачи сигнала о разладке:

$$r = \inf \{n : S_n \geq h\}.$$

Таким образом, проблема сводится к выбору порогового значения, которое минимизирует вероятность пропуска разладки

$$P(S_n < h, n > m+T)$$

для заданного уровня ложной тревоги:

$$P(S_n > h, n \leq T) \leq \alpha,$$

где

α - некоторая константа. Данное пороговое значение будем называть *оптимальным* и обозначать h_{opt} , вероятность пропуска разладки обозначим β .

Введем следующие обозначения $\Gamma(x, n, \theta)$ - гамма распределение с параметрами n, θ

$Inv\Gamma(n, \theta, p)$ - функция, обратная гамма-распределению, возвращает значение x , если $\Gamma(x, n, \theta) = p$.

Учитывая свойства распределений Пуассона, Эрланга и экспоненциального, а также результаты [12] получаем следующую теорему.

Теорема. Пусть наблюдаемый процесс имеет распределение:

1. *экспоненциальное* с параметрами θ_1 до разладки и θ_2 после, тогда

$$h_{opt} = Inv\Gamma(T, \theta_1, 1-\alpha), \\ = \Gamma(h_{opt}, T, \theta_2);$$

2. *Эрланга* с параметрами N_1, θ_1 до разладки и N_2, θ_2 после, тогда

$$h_{opt} = Inv\Gamma(N_1 T, \theta_1, 1-\alpha), \\ = \Gamma(h_{opt}, N_2 T, \theta_2);$$

3. *Пуассона* с параметрами λ_1 до разладки и λ_2 после, тогда

$$\min \{ h : \sum_{k=0}^{h_{opt}} \frac{(T\lambda_1)^k}{k!} > e^{-T\lambda_1} (1 - \alpha) \},$$

$$\beta = \sum_{k=0}^{h_{opt}} \frac{(T\lambda_2)^k}{k!} e^{-T\lambda_2}$$

4. ЗАКЛЮЧЕНИЕ

Используя результаты теоремы можно вычислить оптимальный порог и оценить качество алгоритма для заданных параметров распределений, которые выбираются исходя из конкретной задачи.

Также, для экспоненциального распределения и распределения Эрланга легко получить, если имеется в том необходимость, зависимость какого-либо параметра алгоритма от другого (например T от α) при заданном качестве обнаружения без дополнительного имитационного моделирования. Для чего можно использовать, например, Microsoft Excel, где реализованы функции вычисления гамма-функции обратной к гамма-функции.

ЛИТЕРАТУРА

[1].Медведевский И.Д., Семьянов П.В., Платонов В.В. Атака через Internet. Санкт-Петербург: НПО "Мир и семья-95", 1997.

[2].Xiaobing Zhang, S.Felix Wu, Zhi Fu, Tsung-Li Wu. Malicious Packet Dropping: How It Might Impact the TCP Performance and How We Can Detect It. IEEE Network Protocols, Proceeding International Conference, 2000, p.263-272.

[3].Башарин Г.П., Бочаров П.П., Коган Я.А. Анализ очередей в вычислительных сетях. М.: Наука, 1989.

[4].Бертсекас Д., Галагер Р. Сети передачи данных. М.: Мир, 1989.

[5].Page E.S. Continuous inspection schemes. Biometrika, 1954, v.41, N2, p.100-114.

[6].Page E.S. A test for a change in parameter occurring at an unknown point. Biometrika, 1955,v.42, N4, p.523-527.

[7].Page E.S. On problems in which a change in parameter occurs at an unknown point. Biometrika, 1957,v.44, N2, p.248-252.

[8].Вальд А. Последовательный анализ. М.: Физматгиз,1960, с. 328.

[9].Rodionov A.S., Shakhov V.V. About point-of-change problem in the information systems with exponential distribution. Proceeding of Digital Computer Communication Network Conference, Tel-Aviv (Israel), 1999, p.141 - 145 .

[10].Никифоров И.В. Последовательное обнаружение изменения свойств временных рядов. М.: Наука,1983.

[11]. Клигене Н., Телькснис Л. Методы обнаружения моментов изменения свойств случайных процессов. АИТ, 1983, н.10, с.5-56.

[12].Shakhov V.V. Simulation optimisation problem. Bull. NCC, Special Issue, 1999.

МНОГОМЕРНАЯ КЛАССИФИКАЦИЯ НА ОСНОВЕ МЕТРИКИ КОЛМОГОРОВА — СМИРНОВА

А.А.Кулешов¹, В.Н.Тамашевич²

¹ - Кафедра уравнений математической физики, механико-математический факультет, Белорусский государственный университет, просп. Ф. Скорины, 4, Минск, 220050, БЕЛАРУСЬ, тел. 209-50-47. kuleshov@cit.bsu.unibel.by.

² - Научно-исследовательский институт статистики при Министерстве статистики и анализа Республики Беларусь, Партизанский просп., 12-а, Минск, 220070, БЕЛАРУСЬ, тел. 249-34-13. niistat@postmaster.co.uk.

В настоящей публикации обсуждаются возможности применения и основные принципы работы программы кластерного анализа, созданной на базе компьютерной технической системы *Mathematica* и MS Excel. Математические алгоритмы, положенные в основу программы, в республике реализованы впервые. Описываемая программа не имеет аналогов ни в одном из известных коммерческих статистических пакетов.

Большинство объектов, изучаемых в социальных и экономических науках, может быть охарактеризовано кибернетическим понятием «сложная система». Одна из трудностей исследований в этих областях состоит в том, что почти не существует объектов, которые можно было бы рассматривать как отдельные (внесистемные) элементы. Так, экономика объединяет огромное число элементов, отличается многообразием внутренних связей и связей с другими системами (природной средой, экономической