

plates can be used to generate the specified tests for database systems performance analyzing and for database system user behavior classification. It is developed the software for query templates classification that was integrated into the software system for information systems classification and configuration. The software applying allowed to increase the database productivity using testing and to reduce the tuning time by pre-classification.

Keywords: relational databases, SQL query templates, database testing, information system classification, transaction processing systems.

Рецензент д-р техн. наук, проф. Одес. нац. політехн. ун-ту Крісілов В.А.

Надійшла до редакції 3 лютого 2014 р.

УДК [004.891.3:62-33]:519.1

О.С. Савельева, д-р техн. наук, доц.,
Д.А. Пурич, інженер,
Одес. нац. політехн. ун-т,
Г.В. Налева, канд. техн. наук, доц., Одес. нац.
мор. акад.

МОДЕЛИРОВАНИЕ СОСТОЯНИЯ РЕЗЕРВИРОВАННЫХ СЕТЕВЫХ СИСТЕМ С ЛАТЕНТНЫМИ ЭЛЕМЕНТАМИ С ПОМОЩЬЮ СКРЫТОЙ МАРКОВСКОЙ МОДЕЛИ

Введение. Как известно, техническая диагностика решает обширный круг задач, в частности, весьма важной является оценка состояния системы, ее надежности. Процедура технической диагностики осуществляется без разборки изделия, а ее основной задачей является распознавание состояния системы в условиях ограниченной информации. Контроль технического состояния системы при эксплуатации, по результатам которого принимается решение о вероятности безотказной работы в заданных условиях с установленными параметрами, является важной составной частью системы обеспечения надежности объекта. При этом решения должны приниматься с учетом соответствующих правил, основываться на методах отказов, а для принятия обоснованных решений привлекаться методы теории статистических решений [1].

Анализ последних исследований и публикаций. Процедуры распознавания в технической диагностике основываются на соответствующих моделях, устанавливающих связь между состояниями технической системы и их отображениями в пространстве диагностических сигналов. В основе многих методов, которые при этом используются, — нахождение математической зависимости между повреждениями системы и ее работоспособностью [1, 2]. Основные существующие подходы (вероятностный и детерминистский) принципиально не отличаются. При детерминистском подходе задача сводится к разделению пространства состояний на области диагнозов; при вероятностном — к построению решающего правила, по которому известные параметры, характеризующие с определенной вероятностью состояние системы, относятся к одному из возможных состояний (признаков). Модели, базирующиеся на статистических характеристиках сигнала, описывают Гауссовы, пуассоновские, марковские, а также по-

DOI: 10.15276/opu.1.43.2014.34

© О.С. Савельева, Д.А. Пурич, Г.В. Налева, 2014

добные им процессы. Преимуществом их является возможность одновременного учета признаков различной физической природы, характеризующих вероятность проявления последних при различных состояниях системы.

Задача усложняется, когда исследователю необходимо решать ряд вопросов, связанных с определением показателей надежности системы и идентификацией ее структурных повреждений при наличии в системе скрытых для визуального наблюдения участков [3]. Для таких систем дополнительную диагностическую информацию можно получать, например, с использованием двух моделей: информационной и скрытой Марковской (СММ) [3, 4]. Для этого сетевые структуры в процессе моделирования представляются двумя множествами: скрытой переменной (структура, параметры недоступной части системы) и наблюдаемой переменной (структура, параметры доступной части системы).

Целью работы является повышение эффективности технической диагностики с частично недоступными для мониторинга и ремонта элементами путем разработки и создания алгоритма обучения скрытой марковской модели поведения таких систем.

Изложение основного материала. Задачу нахождения переменных и состояний с использованием СММ рассмотрим для систем, которые могут быть представлены в виде разветвленной структуры с нагруженным резервированием и имеющих частично недоступные для их непосредственного мониторинга части. Такие системы частично недоступны и для анализа, который обязательно предшествует восстановлению.

Для оценки структурной отказоустойчивости резервированных систем существуют различные методы. В основе многих из них лежит нахождение математической зависимости между повреждением структуры системы и ее остаточной работоспособностью. Однако подобный анализ частично скрытой системы затруднен как раз тем, что ее состояние, целостность и повреждения структуры не могут быть определены непосредственно: простым осмотром или каким-либо зондированием. В этом случае приходится пользоваться некоторыми косвенными сигналами, получаемыми от системы [3].

Как известно, скрытые марковские модели предполагают, что получить сами состояния системы невозможно, т.е. остается неизвестным, сколько этих состояний и какие между ними связи, это все — неизвестные параметры модели. Известными являются наблюдаемые величины $c(t)$, которые зависят от скрытых переменных $s(t)$. Задача состоит в определении скрытых параметров процесса, т.е. по имеющимся значениям $c(t)$ необходимо установить наиболее вероятные $s(t)$ и наиболее вероятную модель этого марковского процесса. Значение скрытой переменной $s(t)$ в момент времени t зависит только от значения скрытой переменной $s(t-1)$ в момент $t-1$. Формально это значит, что вероятность $s(t)$ при условии всех остальных установленных значений равна вероятности $s(t)$ при условии только предыдущего значения

$$p(s(t) = s_j | s(t-1) = s_{j_{t-1}}, \dots, s(1) = s_{j_1}) = p(s(t) = s_j | s(t-1) = s_{j_{t-1}}).$$

Значение наблюдаемой переменной $c(t)$ зависит лишь от значения скрытой переменной $s(t)$ (обе в момент времени t).

Если рассматривать только изменение структуры, то любое новое состояние получается из предыдущего удалением некоторого количества элементов и (или) связей между ними [2, 4]. Например, если система содержит 12 связей и если считать повреждениями удаление связей 1...12 в разных сочетаниях и количествах, то количество возможных вариантов состояний плюс исходное неповрежденное этой системы будет составляет

$$N = 2 \cdot 12! \sum_{l=1}^6 \frac{1}{l!(12-l)!} + 1 = 5019.$$

На практике связи могут повреждаться одновременно по две, три, вплоть до выхода из строя всех связей за одну итерацию, что резко увеличивает количество возможных состояний.

Распределение вероятностей начального состояния скрытой части структуры объекта $\pi = \{\pi_j\}$, где $\pi_j = p(q_1 = c_j)$. Здесь q_1 — это начальное состояние.

Наблюдаемая последовательность $c(t)$ представляет собой свойства доступной для мониторинга части объекта. Компоненты общей совокупности множества $C = \{c_1, \dots, c_j, \dots, c_M\}$ — числа, величина и размерность которых определяются пользователем в рамках предметной области. Это может быть, например, напряжение в одной из доступных связей механической системы (σ_i — механическое напряжение в балке, МПа), а графическое представление его изменения во времени может выглядеть, как показано на рис. 1.

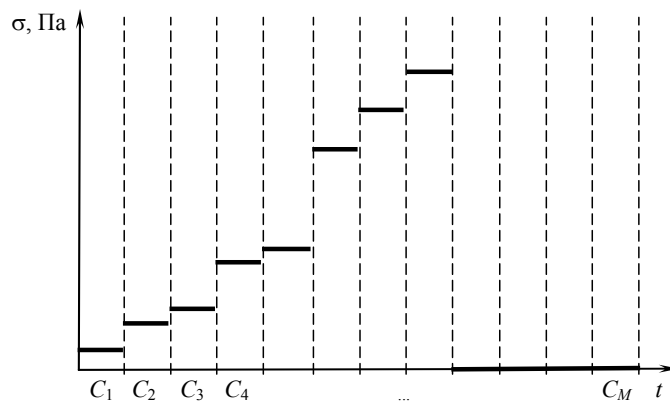


Рис. 1. Пример изменения значений наблюдаемой последовательности переменной

Пользователь определяет также величину M и правило разбиения интервала значений C во времени.

Матрица вероятностей переходов из i в j (или матрица переходов) $A = \{a_{ij}\}$, которая получается в результате предварительных исследований объекта диагностики, где

$$a_{ij} = p(q_{t+1} = s_j | q_t = s_i), \quad 1 \leq i, j \leq N$$

представляет собой вероятность того, что объект, находящийся в состоянии s_i , перейдет в состояние s_j .

Если для любых двух состояний в модели возможен переход из одного состояния в другое, то $a_{ij} > 0$ для любых i, j . В остальных СММ для некоторых i, j вероятность перехода $a_{ij} = 0$.

Третьей составляющей в СММ ($\lambda = (A, B, \pi)$) является матрица наблюдаемых величин $B = \{b_j(r)\}$:

$$b_{ij}(r) = p(s_r | c_j),$$

которую получают в результате обучения СММ.

Задача обучения СММ — важнейшая для большинства проектируемых СММ, поскольку она обеспечивает оптимизацию параметров СММ на основе обучающей наблюдаемой последовательности, то есть позволяет получать модель, наилучшим образом описывающую реальные процессы [5].

Для этого использован алгоритм “предположений и максимизаций” для поиска максимальной вероятностной оценки параметров СММ при заданном наборе наблюдений. Этот алгоритм известен как алгоритм Баума-Велша [5]

$$\lambda^* = \max_{\lambda} p(s_j | \lambda).$$

Вектор λ^* максимизирует вероятность наблюдений C .

Результатом работы алгоритма является подбор такой последовательности состояний скрытой части объекта $S = \{S_1, S_2, \dots, S_m, \dots, S_M\}$, которая лучше всего соответствует наблюдаемой последовательности доступной части $C = \{C_1, C_2, \dots, C_m, \dots, C_M\}$, то есть “объясняет” наблюдаемую последовательность.

Ясно, что абсолютно точно нельзя определить эту последовательность. Здесь можно говорить лишь о предположениях с соответственной степенью достоверности. Процесс решения как раз и состоит в том, чтобы эту вероятность выразить конкретными числами, “предъявляя” на каждой итерации последовательности C набор “обученных” моделей λ .

Наиболее простое решение состоит в том, чтобы определить вероятность появления последовательности наблюдений для каждой возможной последовательности состояний объекта, а затем сложить эти вероятности.

Для i -й последовательности состояний вероятность появления соответствующей последовательности наблюдений C

$$P(C|C_i, \lambda) = \prod_{j=1}^t P(c_j | s_{ij}, \lambda) = \prod_{j=1}^t b_{ij}(c_j).$$

Вероятность же появления самой i -й последовательности состояний

$$P(S_i | \lambda) = \pi_{qi1} \prod_{j=2}^t a[q_{i,j-1} - q_{i,j}].$$

По определению СММ вероятность наблюдения в каждом из состояний зависит только от самого состояния и не зависит от предыдущих состояний

$$P(C|\lambda) = \sum_{i=1}^{N^t-1} P(C|C_i, \lambda) P(S_i | \lambda) = \sum_{i=1}^{N^t-1} (\pi_{qi1} b_{qi1}(c_1) \prod_{j=2}^t a[q_{i,j-1} - q_{i,j}] b_{qij}(c_j)). \quad (1)$$

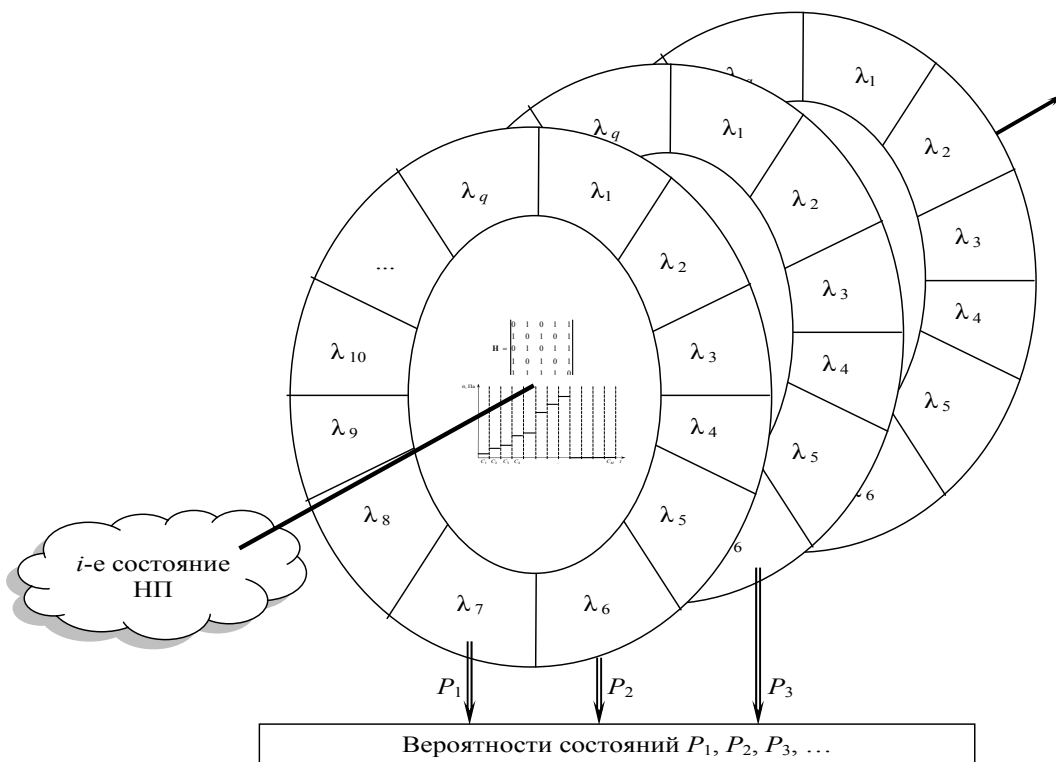


Рис. 2. Схема задачи диагностики скрытой части объекта

Решение выражения (1) найдено с помощью алгоритма прямого-обратного хода. Для этого вводится переменная $\tau_t(i)$ — вероятность того, что к моменту времени t система будет находиться в i -м состоянии, а последовательность порожденных ею до этого момента наблюдений равна $c_1, c_2, \dots, c_t$.

Последовательно находятся вероятности, в начале

$$a_0(i) = \pi_i b_i(c_1),$$

и для текущего состояния для всех t от 1 до τ и для всех j от 1 до N

$$\alpha_t(j) = \left[\sum_{i=1}^N \alpha_{t-1}(i) a_{ij} \right] b_j(c_t).$$

Определяется искомая вероятность

$$P(C|\lambda) = \sum_{i=1}^N \alpha_t(i).$$

Это прямой ход. Обратный проход вычисляет условные вероятности, начиная из состояния i на шаге t , проинициализировав $\beta_t(i)=1$, а затем по индукции

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(c_{t+1}) \beta_{t+1}(j).$$

При диагностике последовательность наблюдений разбивается на отрезки, и применяется алгоритм прямого или обратного хода для вычисления вероятности соответствия данного фрагмента определенному “слову” словаря, составленного при обучении СММ.

В некоторых источниках предлагается считать “слово” распознанным, если вероятность превышает некоторое пороговое значение [4]. Однако такой подход резко снижает точность и адекватность моделирования, поскольку назначение порога всегда является субъективной процедурой, кроме того, у разных элементов одного и того же объекта могут быть различные пороги.

Результаты. В результате для скрытой части сетевого объекта формируется новая матрица состояний, в которой соответствующие буквы p с индексами, символизирующие вероятности существования той или иной связи на той или иной итерации, заменяются конкретными значениями вероятностей (пример на рис. 3).

$$\begin{matrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & p_{12} & 0 & 0 & p_{15} & 0 \\ p_{21} & 0 & p_{23} & 0 & p_{25} & 0 \\ 0 & p_{32} & 0 & p_{34} & 0 & p_{36} \\ 0 & 0 & p_{43} & 0 & 0 & 0 \\ p_{51} & p_{52} & 0 & 0 & 0 & p_{56} \\ 0 & 0 & p_{63} & 0 & p_{65} & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0,34 & 0 & 0 & 0,66 & 0 \\ 0,34 & 0 & 0,78 & 0 & 0,12 & 0 \\ 0 & 0,78 & 0 & 0,47 & 0 & 0,89 \\ 0 & 0 & 0,47 & 0 & 0 & 0 \\ 0,66 & 0,12 & 0 & 0 & 0 & 0,50 \\ 0 & 0 & 0,89 & 0 & 0,50 & 0 \end{pmatrix} \\ a & б & в \end{matrix}$$

Рис. 3. Результат “работы” скрытой марковской модели: а — начальное состояние информации о структуре скрытой части сетевого объекта; б — состояние информации о структуре скрытой части сетевого объекта перед началом работы СММ; в — состояние информации о структуре скрытой части сетевого объекта после работы СММ.

Здесь предложено задачу “округления” полученной на СММ вероятности до диагностических значений (“1” или “0”) решать с помощью информационной морфологической модели [3].

Выводы. Рассматриваемые сетевые структуры с резервированием и частично недоступными для мониторинга элементами могут быть интерпретированы для построения и обучения СММ следующим образом: скрытая переменная представляет собой свойства (структуру, параметры) недоступной для мониторинга части объекта, а наблюдаемая переменная представляет собой свойства (структуру, параметры) доступной для мониторинга части объекта. Такой подход позволяет повысить эффективность разработанного метода, прежде всего, за счет расширения области его применения на новый класс частично недоступных объектов.

Литература

1. Рябинин, И.А. Надежность и безопасность структурно-сложных систем / И.А. Рябинин; Воен. инженер.-техн. ун-т Рос. акад. наук, Ин-т проблем машиноведения. — СПб.: Изд-во С.-Петерб. ун-та, 2007. — 275 с.
2. Оборский, Г.А. Синергетический подход в моделировании марковских процессов / Г.А. Оборский, О.С. Савельева, Н.А. Котенко // Пр. Одес. політехн. ун-ту. — 2011. — Вип. 1(35). — С. 164 — 168.
3. Савельева, О.С. Метод дистанционной структурной диагностики низкочастотной аналоговой сети, частично недоступной мониторингу / О.С. Савельева, В.Г. Максимов, Д.А. Пурич // Пр. Одес. політехн. ун-ту. — 2012. — Вип. 2(39). — С. 208 — 213.

4. Савельєва, О.С. Підвищення надійності систем дистанційного діагностування / О.С. Савельєва, О.Л. Становський, Д.О. Пуріч // Наукові вісті [Інституту менеджменту та економіки "Галицька академія"]. — 2009. — № 15(1). — С. 58 — 63.
5. The HTK Hidden Markov Model Toolkit Book (for HTK Version 3.4) / S. Young, G. Evermann, M. Gales et al. — Cambridge University Engineering Department, 2006. — 368 p.

References

1. Ryabinin, I.A. Nadezhnost' i bezopasnost' strukturno-slozhnykh sistem [Reliability and Safety of the Structurally Complex Systems] / I.A. Ryabinin; Voen. inzhener.-tekhn. un-t, Ros. akad. nauk, In-t problem mashinovedeniya [Military Engineering University, Russian Acad. Sciences, Institute of Problems of Mechanical Engineering]. — Saint Petersburg, 2007. — 275 p.
2. Oborskiy, G.A. Sinergeticheskiy podkhod v modelirovanii markovskikh protsessov [Synergetic approach in modeling Markov processes] / G.A. Oborskiy, O.S. Savel'yeva, N.A. Kotenko // Pratsi Odeskoho politekhnichnoho universytetu // [Proc. Odesa Polytech. Univ.]. — 2011. — Iss. 1(35). — pp. 164 — 168.
3. Savel'yeva, O.S. Metod distantsionnoy strukturnoy diagnostiki nizkochastotnoy analogovoy seti, chastichno nedostupnoy monitoringu [Method of remote structural diagnostics of a low-frequency analogue network partially inaccessible to monitoring] / O.S. Savel'yeva, V.G. Maksimov, D.A. Purich // Pratsi Odeskoho politekhnichnoho universytetu [Proc. Odesa Polytech. Univ.]. — 2012. — Iss. 2(39). — pp. 208 — 213.
4. Savielieva, O.S. Pidvyshchennia nadiinosti system dystantsiinoho diahnostuvannia [Improving the reliability of remote diagnostics] / O.S. Savielieva, O.L. Stanovskyi, D.O. Purich // Naukovi visti [Instytutu menedzhmentu ta ekonomiky "Halytska akademiia"] Science news [Institute for Management and Economics "Galician Academy"]. — 2009. — #15(1). — pp. 58 — 63.
5. Young, S. The HTK Hidden Markov Model Toolkit Book (for HTK Version 3.4) / S. Young, G. Evermann, M. Gales et al. — Cambridge University Engineering Department, 2006. — 368 p.

АНОТАЦІЯ / АННОТАЦИЯ / ABSTRACT

О.С. Савельєва, Д.О. Пуріч, Г.В. Налєва. Моделирование stanu резервованных сетевых систем с латентными элементами за допомогою прихованої марківської моделі. Розглядаються питання, пов'язані з розв'язанням одного із завдань технічної діагностики — оцінкою стану технічної системи. На підставі контролю технічного стану резервованої системи приймається рішення про ймовірність її безвідмовної роботи в заданих умовах з встановленими параметрами. В основі методів, які при цьому використовуються, — знаходження математичної залежності між ушкодженнями системи і ймовірністю її безвідмовної роботи. Завдання ускладнюється, коли необхідно вирішувати питання, пов'язані з визначенням показників надійності системи й ідентифікацією її структурних ушкоджень при наявності в системі прихованих для візуального спостереження ділянок. Додаткову діагностичну інформацію отримано з використанням двох моделей: інформаційної і прихованої марківської моделі. У результаті для прихованої частини резервованої системи отримано нову матрицю станів з конкретними значеннями ймовірностей.

Ключові слова: діагностика, мережеві системи, прихована марківська модель, оцінка стану.

О.С. Савельєва, Д.А. Пуріч, Г.В. Налєва. Моделирование состояния резервированных сетевых систем с латентными элементами с помощью скрытой марковской модели. Рассматриваются вопросы, связанные с решением одной из задач технической диагностики — оценкой состояния технической системы. На основании контроля технического состояния резервированной системы принимается решение о вероятности ее безотказной работы в заданных условиях с установленными параметрами. В основе методов, которые при этом используются, — нахождение математической зависимости между повреждениями системы и вероятностью ее безотказной работы. Задача усложняется, когда необходимо решать вопросы, связанные с определением показателей надежности системы и идентификацией ее структурных повреждений при наличии в системе скрытых для визуального наблюдения участков. Дополнительная диагностическая информация получена с использованием двух моделей: информационной и скрытой марковской модели. В результате для скрытой части резервированной системы получена новая матрица состояний с конкретными значениями вероятностей.

Ключевые слова: диагностика, сетевые системы, скрытая марковская модель, оценка состояния.

O.S. Saveleva, D.A. Purich, G.V. Naleva. Modeling of redundant network systems with latent elements using hidden Markov model. Problems associated with the decision of one of the tasks of technical diagnostics — state estimation of the technical system. Based on the technical state control redundant system the decision on its probability of failure-free operation under specified conditions, with the specified parameters. The basis of the methods that are used at the same time —

to find the mathematical relationship between the injuries system and its functions. The task is complicated when you need to deal with measurement reliability of the system and identification of its structural damage if the system hidden for visual observation plots. Additional diagnostic information was obtained using the two models — the information and the hidden Markov model. As a result, for the hidden part of a redundant system a new state matrix with specific values of probabilities is received.

Keywords: diagnostics, network systems, hidden Markov model, state estimation.

Рецензент д-р. техн. наук, проф. Одес. нац. политехн. ун-та Крылов В.Н.

Поступила в редакцию 29 мая 2014 г.

УДК 504.3.054:614.712.001.57

В.Я. Гамоліч, математик,
А.Г. Бутенко, канд. техн. наук, доц.,
С.Ю. Смык, специалист,
Одес. нац. политехн. ун-т

ВЕРОЯТНОСТНАЯ МОДЕЛЬ КОМБИНИРОВАННОЙ СИСТЕМЫ ОЧИСТКИ ВОЗДУХА

Введение. Современные нормы экологической безопасности предъявляют высокие требования к качеству очистки газовых выбросов в атмосферу от полидисперсной пылевой массы. На промышленных предприятиях, особенно небольших, и в настоящее время широко используются инерционные уловители, эффективность которых не соответствует современным требованиям. Возможности повышения их эффективности очистки за счет конструктивного усовершенствования фактически исчерпаны, а замена на более эффективные сдерживается как технологическими особенностями производства, так и экономическими факторами.

Анализ последних исследований и публикаций. Известно, что эффективность работы уловителя зависит не только от их конструкции, но и от свойств пыли, в частности, фракционного состава. Предложенная комбинированная система очистки газа позволяет значительно повысить эффективность уловителей за счет совершенствования схемы движения полидисперсного потока и связанного с этим перераспределения пылевой массы по фракционному признаку. Поток с тонкими фракциями направляется в уловитель циркуляционного контура, а поток с крупными — в основной уловитель. Поскольку коэффициент улавливания крупных фракций достаточно высокий, то проскок пыли в окружающую среду является незначительным. Этим обеспечивается экологический эффект. Опыт внедрения комбинированной системы очистки (КСО) показал многократное снижение массы выбрасываемой в окружающую среду пыли, по сравнению с обычной прямоточной схемой [1].

Цель работы является — построение вероятностной модели КСО, позволяющей прогнозировать изменение показателей системы во времени.

Изложение основного материала. Характерной особенностью КСО является изменение во времени ее основных показателей. Все время работы системы можно разделить на два периода:

— очистки, состоящий из двух этапов: разгонного, когда параметры изменяются, асимптотически приближаясь к стабильным, и этапа стабильных показателей;

DOI: 10.15276/opr.1.43.2014.35

© В.Я. Гамоліч, А.Г. Бутенко, С.Ю. Смык, 2014