ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

УДК 519.8+ 681.31

А.В. Викулов, А.А. Кочешков

АНАЛИЗ И МОДЕЛИРОВАНИЕ ДИНАМИЧЕСКИХ СВОЙСТВ ИНФОРМАЦИОННЫХ СИСТЕМ

Нижегородский государственный технический университет имени Р.Е. Алексеева,
Институт радиоэлектроники и информационных технологий

Представлена модель динамических свойств информационной системы в пространстве состояний. Ставится задача наблюдения за состоянием системы. Предлагается задача оценки параметров модели.

Ключевые слова: управление, наблюдение, информационная система, моделирование, пространство состояний

Введение

Современные информационные системы (ИС), в отличие от простых хранилищ документов, являются динамичными и развивающимися объектами. Наряду с задачами сбора, хранения и поиска данных, большое значение имеют текущие процессы управления и наблюдения за состоянием ИС.

В сетевой среде эффективное управление взаимодействием клиентов и серверов позволяет рационально распределять ресурсы, избежать множества лишних операций по передаче и обработке данных. На стороне серверов характер изменчивости данных влияет на такие процессы, как репликация, резервное копирование, индексация в поисковых системах. На стороне клиентов необходимо решить вопрос о периодичности формирования запросов к наблюдаемым серверам, чтобы соотнести выгоду от новой информации с затратами на ее получение и обработку. Решение задачи наблюдения за информационной системой должно учитывать общую цель обработки информации и минимизировать обобщенный критерий, в котором штраф от задержки получения информации о состоянии информационной системы соотносится с затратами на ее получение.

Формализовать такую оптимизационную задачу и получить необходимые априорные данные для объектов информационного характера сложно, поэтому чаще всего проблема решается простейшим эвристическим способом, что в итоге дает отклонение или в область вероятных существенных потерь информации, или в область неоправданных затрат на повторную обработку одной и той же информации. Построение адекватной математической модели, описывающей динамические процессы в ИС, дает основу для обоснованной и более выгодной стратегии наблюдения и обслуживания.

Для эффективного планирования использования ресурсов ИС клиентами необходимо решение задачи наблюдения за ее динамикой. Процесс наблюдения в динамических системах функционально может быть разбит на два этапа:

- проведение измерений;
- преобразование полученной информации по определенным алгоритмам.

Первый этап реализуется с помощью средств наблюдения и каналов связи, второй –

программным путем. Подавляющая часть исследований посвящена оптимизации второго этапа как наиболее легко перестраиваемого и носящего обобщенный характер. При этом модель измерительной системы обычно принимается фиксированной и заданной совместно с моделью наблюдаемого объекта. В то же время полнота и достоверность измерительной информации, получаемой на первом этапе, несомненно, определяют качество функционирования всей системы, поэтому выбор наилучшего варианта структуры измерительной системы и режимов ее работы во времени является важной задачей.

Предлагается решать задачу методом пространства состояний, что позволяет учитывать и многомерность, и нестационарность стохастических процессов.

Грамотно спланированное наблюдение за состоянием ИС лежит в основе рационального распределения ресурсов ИС, управления пользователями, позволяет своевременно решать разнообразные проблемы, возникающие при ее внедрении и эксплуатации.

Теоретический анализ

Практически все информационные системы являются динамическими объектами, где постоянно происходят изменения как в структуре самой системы, так и в содержании ее элементов. Под изменением системы будем считать выполнение хотя бы одной из следующих трех операций: добавление, удаление и внутреннее изменение элемента данных. Данное определение базируется на основе анализа типичных операций в реальных системах и удобно тем, что позволяет задавать количественные характеристики и формально описывать динамику наблюдаемых объектов.

Динамика информационной системы состоит в изменении ее состояния и параметров во времени. Состояние будем характеризовать множеством изменений элементов данных и структурных связей между ними, произошедших внутри системы за указанный временной интервал.

Отличительной особенностью динамических информационных систем является отсутствие явно выраженного физического закона их функционирования. Изменения в таких системах происходят случайным образом, поэтому единственным способом оценить параметры наблюдаемого объекта является наблюдение за его динамикой, накопление и исследование статистики случайных процессов, служащих моделью состояния системы. Свойства этих случайных процессов могут непосредственно использоваться в алгоритмах ковариационного анализа для построения моделей методами теории временных рядов, базирующихся на использовании моделей авторегрессии и скользящего среднего [1, 2].

Классическим методом моделирования динамики систем является метод пространства состояний (МПС), который позволяет учитывать многомерность и нестационарность процессов. Согласно МПС линейная дискретная стохастическая модель объекта имеет вид

$$x_{k+1} = A_k x_k + C_k w_k \tag{1}$$

где x_k – n-мерный вектор состояния; $x_{k+1} = x[(k+1)T]$; A_k – матрица состояния объекта размерности $n \times n$, w_k – l-мерный вектор шума в объекте с математическим ожиданием $E\{w_k\} = 0$.

Состояние ИС характеризует количество элементов данных, классифицированных по типу и местам хранения. Изменение состояния вызывают действия обновления и модификации данных.

Функция состояния системы представляет собой дискретный случайный процесс

$$x_k = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^T .$$

Дискретизация этого процесса с постоянным периодом T во временной области производится при помощи кусочно-постоянной аппроксимации [3, 4]. В каждый $(k+1)$-й момент времени $x(k+1)$ есть число изменений в системе относительно предыдущего k-го шага.

В общем случае имеем n-мерный вектор состояния, определенный в k-й момент времени. Координаты данного вектора x_k – это множество изменений подсистем, входящих в состав наблюдаемого информационного объекта.

Состояние объекта в подавляющем большинстве случаев известно не полностью. Вся доступная информация о текущем состоянии x_k содержится в векторах измерений (наблюдений), проводимых в дискретные моменты времени $t_k \le kT$. При постоянном периоде измерений $t_k - t_{k-1} = T$ уравнение измерителя имеет вид

$$y_{k+1} = H_{k+1}x_{k+1} + \eta_{k+1} \qquad (2)$$

где $y_{k+1} = y[(k+1)T]$ – m-мерный вектор измерения; H_{k+1} – матрица измерителя размерности $m \times m$, η_{k+1} – m-мерный вектор шума измерений. Дискретный случайный процесс η_{k+1} физически интерпретируется как погрешность измерений и может быть как белым гауссовским шумом с математическим ожиданием $E\{\eta_{k+1}\} = 0$ и КФ $R_{k+1}(\eta_{k+1}) = E\{\eta_{k+1}\eta_{k+1}^T\}, k = \overline{0, N}$, так и кореллированным процессом.

В системе с управляемыми измерениями уравнение измерителя записывается в следующем виде, синхронном с дискретностью уравнения объекта (1):

$$y_{k+1} = H_{k+1}(S_{k+1})x_{k+1} + \eta_{k+1}(S_{k+1}), \qquad (3)$$

причем среди векторов y_{k+1} доступны только те, что соответствуют моментам проведения измерений. Остальные вектора можно характеризовать как фиктивные измерения.

Стратегия наблюдений на N-шаговом переходном процессе представляет собой последовательность

$$S^N = \{S_1, S_2, ..., S_N\}, S_i \in \{0,1\} . \qquad (4)$$

Для системы с постоянными измерениями S^N является единичной.

В системах с кратными периодами дискретности стратегия вида

$$S^N = \{...,1,0,0,...,0,1,0,0,...,0,...,1,0,0,...,0,...\}$$

характеризует систему с периодическим режимом съема информации.

Таким образом, дискретная модель системы наблюдения – это уравнение объекта, уравнение измерителя в синхронном с уравнением объекта виде и стратегия наблюдения S^N на N-шаговом переходном процессе.

Оценивание параметров модели

Наблюдаемые координаты состояния ИС представляют собой векторный случайный процесс с дискретным временем. В рамках корреляционной теории по собранной статистике могут быть построены модели временных рядов: авторегрессия и скользящее среднее.

Классическое уравнение авторегрессии с постоянными коэффициентами имеет следующий вид (5):

$$\xi_{k+1} = \sum_{i=0}^{\alpha} \Gamma_i \xi_{k-i} + w_k \qquad (5)$$

где $\xi_k \in R^p$ – вектор состояния процесса размерности p; $w_k \in R^p$ – вектор белого гауссовского шума с КМ Q_w, некоррелированный с; $\Gamma_j \in R^{p \times p}$ – матрицы коэффициентов размерности $p \times p$, $j = \overline{0, \alpha}$. Согласно терминологии теории динамических систем, уравнение (5) описывает линейную стохастическую систему с запаздываниями, вектор состояния которой является дискретным марковским процессом порядка $\alpha + 1$.

Достаточной статистикой гауссовского марковского процесса (5) порядка $\alpha + 1$ являются математическое ожидание $\overline{\xi_k} = E\{\xi_k\}$, $(\alpha + 1)$ отсчетов матричной КФ

$$Q(i) = E\{\xi_k \xi_{k-i}^T\}, \ i = \overline{0, \alpha + 1}.$$

Для определения параметров уравнения (5) предлагается два метода. Рассмотрим метод точечной аппроксимации заданной КФ, предполагающий получение модели (5), при которой КФ формируемого процесса $Q_M(i)$ совпадает в $(\alpha+1)$-й точках по аргументу i с заданной КФ $Q(i), i=\overline{0,v}$.

Согласно работе [5], матричная КФ $Q(i), i=\overline{0,\alpha+1}, k=\overline{0,N}$, порождающая положительно определенную блочную матрицу $M(\alpha,\alpha)$

$$M(\alpha,\alpha)=\begin{bmatrix} Q(0) & Q(1) & \dots & Q(\alpha) \\ Q^T(1) & Q(0) & \dots & Q(\alpha-1) \\ \vdots & \vdots & \ddots & \vdots \\ Q^T(\alpha) & Q^T(\alpha) & \dots & Q(0) \end{bmatrix}, \tag{6}$$

является достаточной статистикой некого гауссовского центрированного марковского процесса порядка $\alpha+1$, описываемого уравнением (5).

Введем в рассмотрение взаимную КМ:

$$M(1,\alpha)=[Q(1),Q(2),...,Q(\alpha+1)]\ , \tag{7}$$

причем блоки матриц $R(\alpha,\alpha)$ и $R(1,p)$ являются элементами заданной КФ.

Блочная матрица выбирается в виде

$$\Gamma(\alpha)=M(1,\alpha)M(\alpha,\alpha)^{-1}\ . \tag{8}$$

КМ Q_w возбуждающего белого шума w_k находится из уравнения

$$Q_w=Q(0)-M(1,\alpha)M(\alpha,\alpha)^{-1}M(1,\alpha)^T\ . \tag{9}$$

Таким образом, заданная КФ $Q(i), i=\overline{0,\alpha+1}$ определяет параметры гауссовский плотности вероятности марковского процесса и, соответственно, параметры его модели в виде уравнения авторегрессии с постоянными коэффициентами.

Другой подход, позволяющий учесть при синтезе формирующего фильтра (5) $\gamma>\alpha+1$ значений $Q(i), i=\overline{0,\gamma}$, получим, применяя метод наименьших квадратов (МНК).

Сформируем блочную матрицу размерностью $\gamma\times\alpha+1$ блоков, следующего вида:

$$M(\gamma,\alpha)=\begin{bmatrix} & M(\alpha,\alpha) & & \\ Q(\alpha+1) & Q(\alpha) & \dots & Q(1) \\ \vdots & \vdots & \vdots & \vdots \\ Q(\gamma) & Q(\gamma-1) & \dots & Q(\gamma-\alpha) \end{bmatrix} \tag{10}$$

Взаимная КМ запишется:

$$M(1,\gamma)=[Q(1),Q(2),...,Q(\gamma+1)] \tag{11}$$

С помощью псевдообратной матрицы

$$M^+(\gamma,a)=\left[M(\gamma,\alpha)^T M(\gamma,\alpha)\right]^{-1}M(\gamma,\alpha) \tag{12}$$

находится оптимальное в смысле минимумов наименьших квадратов решение

$$\Gamma(\alpha)^T=M^+(\gamma,\alpha)M(1,\gamma)^T\ . \tag{13}$$

В качестве приближения КМ возбуждающего шума может быть принято значение

$$Q_w=Q(0)-\Gamma(\alpha)M(\alpha,\alpha)\Gamma^T(\alpha)\,. \qquad 14)$$

Рассчитав параметры модели $\Gamma(\alpha)$ и Q_w, вычислим КФ модели. Ковариационная функция модели $Q_M(i)$ на интервале $\alpha<i<\gamma$ вычисляется по формулам

$$\begin{cases} Q_M(0) = \sum\limits_{i,j=0}^{\alpha} \Gamma_{k,i} Q_M(|i-j|)\Gamma_{k,j}^T + Q_{Mw}, \\ Q_M(i) = \sum\limits_{j=0}^{\alpha} \Gamma_j Q_M(|i-j-1|), i = \overline{1,\gamma} \end{cases} \tag{15}$$

Найти решение системы (15) предлагается рекуррентной процедурой

$$\begin{cases} Q_{M\,k+1}(0) = \sum\limits_{i,j=0}^{\alpha} \Gamma_i Q_{M\,k-\min(i,j)}(|i-j|)\Gamma_{k,j}^T + Q_{M\,wk}, \\ Q_{M\,k+1}(i) = \sum\limits_{j=0}^{\alpha} \Gamma_j Q_{M\,k-\min(i-j-1,j)}(|i-j-1|), i = \overline{1,\gamma} \end{cases} \tag{16}$$

по шагам $k = \overline{0,N}$ до достижения установившегося решения $Q_M(i)$. Начальные условия $Q_{M\,0}(i) = Q(i)$.

Отождествив выходную последовательность ξ_k с процессом x_k, а шум фильтра w_k с шумом в объекте, можно перейти от уравнения авторегрессии к уравнению объекта в пространстве состояний.

Оценка процесса авторегрессии

Рассмотрим случай управляемого наблюдения, когда измеряются все координаты вектора состояния без погрешности, но в определенные моменты дискретного времени k.

Модель наблюдателя может быть сформирована зависимостями

$$H(S_k) = \begin{cases} I, \text{при } S_k = 1 \\ 0, \text{при } S_k = 0 \end{cases}, \tag{17}$$

где I – единичная nxn матрица, η_{k+1} – вектор белого шума с ковариационной матрицей

$$R_\eta(S_k) = \begin{cases} 0, \text{при } S_k = 1 \\ \infty, \text{при } S_k = 0 \end{cases}, \tag{18}$$

где условная запись $R_\eta = \infty$ воспринимается как неограниченное возрастание диагональных элементов матрицы R для процедуры предельных переходов.

В случае управляемого наблюдения с полным и точным измерением вектора состояния (с учетом вида $H(S_{k+1})$ и $R_\eta(S_{k+1})$) получим оптимальную оценку состояния в виде

$$\hat{\xi}_{k+1} = \begin{cases} y_{k+1}, S_{k+1} = 1, \\ \hat{\xi}_{k+1|k}, S_{k+1} = 0. \end{cases} \tag{19}$$

где $\hat{\xi}_{k+1|k} = \sum\limits_{i=0}^{\alpha} \Gamma_i \hat{\xi}_{k-i}$.

Тогда ошибка оценки $\tilde{\xi}_{k+1} = \xi_{k+1} - \hat{\xi}_{k+1}$

$$\tilde{\xi}_{k+1} = \begin{cases} 0, S_{k+1} = 1, \\ \xi_{k+1} - \hat{\xi}_{k+1}, S_{k+1} = 0. \end{cases} \tag{20}$$

Для планирования наблюдений и анализа точности необходимо вычислить ковариационную функцию ошибки оценки $\tilde{Q}_k(i) = E\{\tilde{\xi}_k \tilde{\xi}_{k-i}^T\}$, $i = \overline{0,v}$.

Критерием точности оценок служит дисперсия ошибки оценки. Рекуррентные уравнения для вычисления $\tilde{Q}_k(i)$ должны учитывать наличие измерения на соответствующей итерации.

$$\begin{cases} \tilde{Q}_{k+1}(0) = \sum_{i,j=0}^{\alpha} \Gamma_i \tilde{Q}_{k-\min(i,j)}(|i-j|)\Gamma_{k,j}^T + Q_{wk} \\ \tilde{Q}_{k+1}(i) = \sum_{j=0}^{\alpha} \Gamma_j \tilde{Q}_{k-\min(i-j-1,j)}(|i-j-1|), i=\overline{1,\gamma} \end{cases} \qquad (21)$$

Экспериментальная часть

На основе предложенных алгоритмов был реализован программный комплекс, предназначенный для исследования динамики локальных и удаленных информационных систем. Программный комплекс состоит из трех программных средств: монитора изменений информационной системы, обработчика данных и анализатора, предназначенного для построения моделей и стратегий наблюдения. Этап моделирования представляет собой запуск монитора для слежения и сбора статистики выбранного объекта. После требуемого периода наблюдений данные должны быть обработаны программой обработчиком измерений. Эта программа позволяет извлечь необходимые данные из хранилища, просмотреть их, определить координаты состояния и настроить параметры экспорта: отфильтровать данные по времени, выставить фильтр, выбрать формат файла. После того, как экспорт данных был произведен, полученный файл может быть использован для анализа в программе анализаторе или в любой другой совместимой программе. Программа анализатор строит модель динамики объекта, оценивает входящий процесс и выдает результаты в наглядном графическом и табличном видах.

Результаты

Ряд информационных ресурсов сети Интернет имеет ярко выраженный динамический характер. В качестве примера был взят информационный портал РосБизнесКонсалтинг, расположенный на веб-сайте http://www.rbc.ru.

На начальном этапе изучения свойств объекта с постоянным периодом в 1 час собирались данные об обновлениях текстовых блоков и графических элементов. Обозначим процесс изменения текстовых блоков как $(x_k)_1$, а графических $(x_k)_2$.

Процессы изменения данных на сайте имеют коррелированный характер, что связанно с периодичностью суточной активности. В дневной период интенсивность изменений в несколько раз выше, чем в ночные часы. Это позволяет планировать наблюдения за информацией на сайте и экономить ресурсы.

Имея адекватную модель системы, можно рассмотреть задачу управляемого наблюдения. Чтобы решить эту задачу, необходимо выбрать стратегию наблюдения, рассчитать оценку состояния системы и проанализировать ее точность. В качестве примера, зададим стратегию таким образом, что наблюдение будем производить в моменты времени k, кратные T:

$$S_k = \begin{cases} 1, k\, кратно T \\ 0, k\, не\, кратно T \end{cases}.$$

Вычисленная по формуле (20) оценка ($(\tilde{\xi}_k)_1$ и $(\tilde{\xi}_k)_2$ соответственно) позволяет вычислить дисперсию ошибки оценки по формуле (21), которая служит критерием точности. График реализации исходного процесса и его оценки представлен на рис. 1.

Поскольку целью наблюдения является максимально точное оценивание, примем за критерий качества квадратичную форму ошибки оценки ($(\tilde{Q}_k)_1$ и $(\tilde{Q}_k)_2$ соответственно, $(\tilde{Q}_k)_{вз}$ – взаимная ошибка оценки). Дисперсия ошибки оценки представлена на рис. 2.

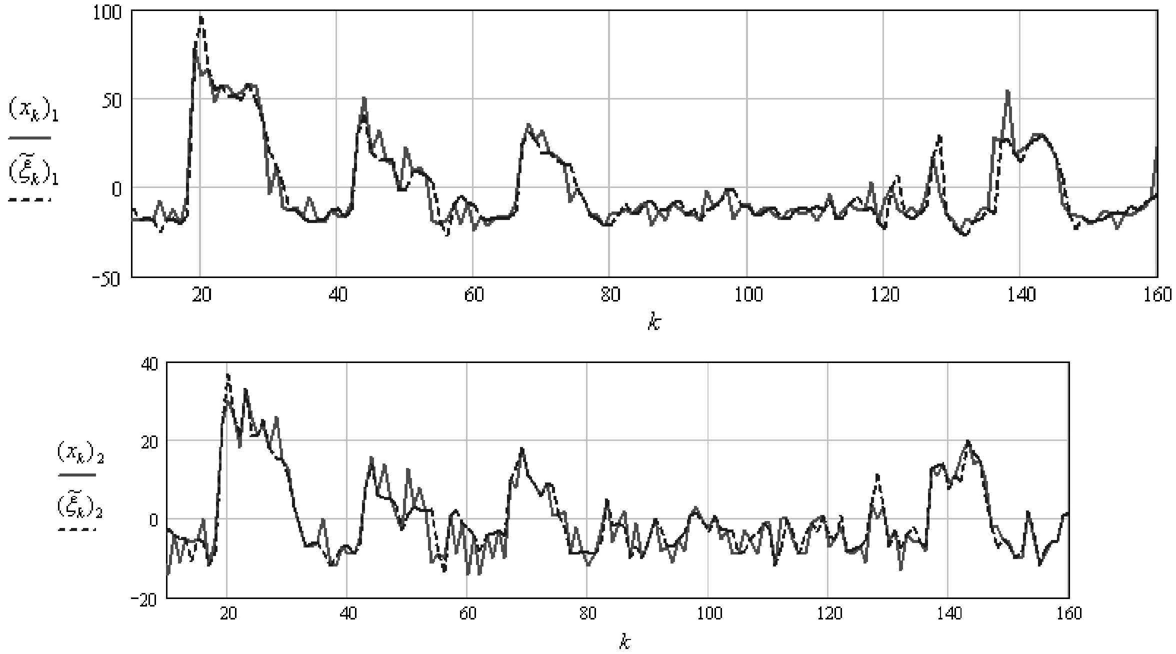

Рис. 1. Графики реализации исходного процесса x_k и его оценки $\tilde{\xi}_k$ для *T*=2

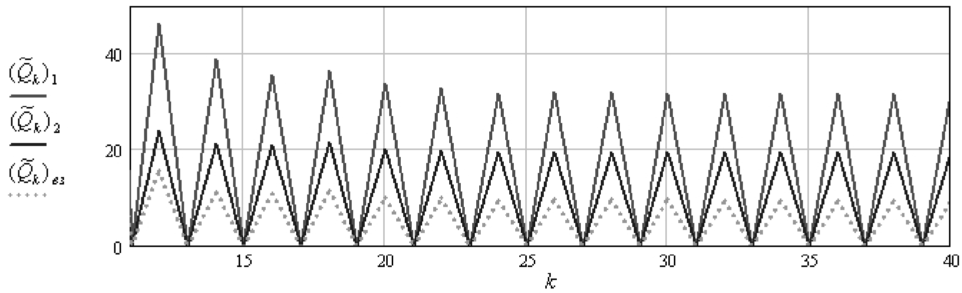

Рис. 2. График дисперсии ошибки оценки $\tilde{Q}_k$ для *T*=2

В результате проведения серии экспериментов был получен график зависимости точности оценки от величины интервала между наблюдениями (рис. 3).

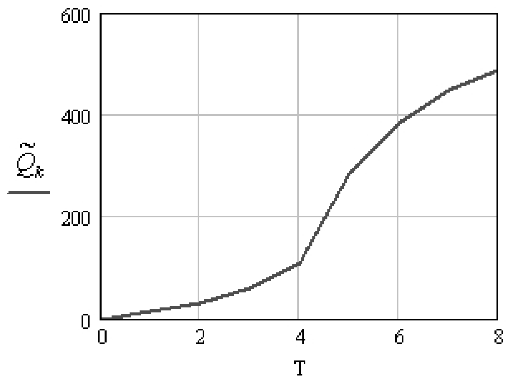

Рис. 3. Точность оценки

Задав необходимую точность оценки, можно выбрать интервал между наблюдениями. Это позволит оператору выбрать наиболее подходящую стратегию наблюдения за состоянием ИС.

Библиографический список

1. **Бокс, Дж.** Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс. – М.: Мир, 1974. – 315 с.
2. **Хеннан, Э.** Многомерные временные ряды / Э. Хеннан. – М.: Мир, 1974. – 576 с.
3. **Медич, Д.** Статистически оптимальные линейные оценки и управление / Д. Медич. – М.: Энергия, 1973. – 440 с.
4. **Богуславский, И.А.** Прикладные задачи фильтрации и управления / И.А. Богуславский. – М.: Наука, 1983. – 400 с.
5. **Кондратьев, В.В.** Фильтрация и анализ линейных дискретных систем управления при непараметрическом задании коррелированных шумов / В.В. Кондратьев, А.А. Кочешков // Автоматика и телемеханика. 1985. № 6. С. 67–76.

Дата поступления
в редакцию 05.10.2012

A.V. Vikulov, A.A. Kocheshkov

ANALYSIS AND MODELLING OF INFORMATIONAL SYSTEMS DYNAMICAL PROPERTIES

Nizhny Novgorod State Technical University named after R.E. Alekseev,
Institute of radio-electronics and informational technologies

Purpose: The problems of informational system state observation and control are analyzed with particular attention to the rational system resources usage. The definition of the state is given. The software based on the proposed methods was implemented.
Design/methodology/approach: A theoretical framework is proposed based on discrete-time control theory methods and stochastic process corelation theory. The informational system’s dynamical properties discrete model based on state-space approach has been considered. The problem of model parameters evaluation is proposed.
Findings: It is possible, for example, to apply either theoretical framework or the software to improve the quality of the system maintenance
Research limitations/implications: The present study provides a starting-point for further research in the sector of information technologies

Key words: control, observation, information system, modeling, state-space.