

В. В. Рыжаков, М. В. Рыжаков

**ТЕОРИЯ И АЛГОРИТМ ДИАГНОСТИРОВАНИЯ
БИОЛОГИЧЕСКИХ СИСТЕМ С БЕСКОНЕЧНЫМ ЧИСЛОМ
СОСТОЯНИЙ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ
СООТВЕТСТВУЮЩИХ Q -МАТРИЦ И МАТРИЦ
ВЕРОЯТНОСТЕЙ ПЕРЕХОДОВ**

Аннотация. *Актуальность и цели.* Биологические системы (популяции) живых организмов характеризуются массовым заселением отдельных территорий. При определенных экологических нарушениях в этом сообществе могут возникать соответствующие нарушения, которые можно использовать в качестве меры экологических нарушений при их количественном оценивании. *Материалы и методы.* При разработке процедур диагностирования биологических систем с бесконечным числом особей и их состояний учитывалась по известным публикациям особенность существования особей. Поэтому в работе предлагаются теория и алгоритмы, которые основаны на определенном вероятностном подходе. Этот подход использует треугольные матрицы с бесконечным числом вероятностей переходов, позволяющим учитывать бесконечное число особей и их состояний. Для получения элементов матрицы вероятностей переходов систем из одного состояния в другое использовались Q -матрицы и известные уравнения Колмогорова. *Результаты.* На основе использования полученной матрицы вероятностей переходов предложены алгоритмы и методика выборочного контроля состояний системы (популяции) организмов, в которой учитываются особенности использования биномиального и пуассоновского законов, влияющие на точность контроля. *Выводы.* Полученные результаты отличаются существенной утилитарностью в практических приложениях.

Ключевые слова: биологическая система, теория, алгоритм, диагностирование, Q -матрица, вероятности переходов.

V. V. Ryzhakov, M. V. Ryzhakov

**THE THEORY AND ALGORITHM FOR DIAGNOSIS
OF BIOLOGICAL SYSTEMS WITH AN INFINITE NUMBER
OF STATES ON THE BASIS OF APPLICABLE Q -MATRICES
AND TRANSITION PROBABILITY MATRICES**

Abstract. *Background.* Biological systems (populations) of living organisms are characterized by massive population of particular territories. Under certain environmental violations there may occur corresponding violations in this community that can be used as a measure of environmental violations for their quantitative assessment. *Materials and methods.* When developing the procedures for diagnosing biological systems with an infinite number of individuals and their states the peculiarities of specimen existence were taken into account according to some known publications. Therefore, the paper proposes a theory and algorithms which are based on a certain probabilistic approach. This approach uses triangular matrices with an infinite number of transition probabilities that allow taking into consideration an infinite number of specimens and their states. For the acquisition of probability matrix elements for system transitions from one state to another Q -matrices and known Kolmogorov equation were used. *Results.* Based on the use of the transition probability

matrix there were suggested algorithms and sampling control technique over the states of specimen system (population), which considers the peculiarities of using binomial and Poisson laws affecting the accuracy of control. *Conclusions.* The results obtained are characterized by substantial utility for practical purposes.

Key words: biological system, theory, algorithm, diagnosis, Q-matrix, transition probabilities.

Отдельные биологические системы (популяции) живых организмов характеризуются массовым заселением определенных территорий [1]. Число таких организмов может насчитывать миллионы. При определенных экологических нарушениях в этом сообществе могут возникать нарушения (изменения) состояний, которые можно оценить только с использованием вероятностных показателей (надежности). Для диагностирования состояний таких систем на основе вероятностных показателей предлагается использовать определенную теорию и алгоритм, позволяющие учитывать бесконечное (большое) число состояний [2].

В основе предлагаемой теории оценивания состояния систем предлагается использовать треугольную матрицу с бесконечным числом элементов вида [3]

$$Q = \begin{pmatrix} -v & v & 0 & 0 & \dots \\ 0 & -v & v & 0 & \dots \\ 0 & 0 & -v & v & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad (1)$$

которой соответствует также треугольная матрица вероятностей переходов $P(t)$:

$$P(t) = \begin{pmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) & \dots \\ 0 & P_{22}(t) & P_{23}(t) & \dots \\ 0 & 0 & P_{33}(t) & \dots \end{pmatrix}, \quad (2)$$

соответствующая в свою очередь цепям Маркова с непрерывным временем.

Используя матричную экспоненту [3]

$$P(t) \cdot e^{tQ} \quad (3)$$

и ее производную (уравнение Колмогорова)

$$P'(t) = P(t) \cdot Q \quad (4)$$

или

$$P'(t) = Q \cdot P(t), \quad (4a)$$

определим выражения элементов матрицы (2).

Из (3) следует

$$P(t=0) = E, \quad (5)$$

что соответствует условию

$$\begin{cases} P_{11}(t=0)=1 \\ P_{22}(t=0)=1 \\ \dots\dots\dots \\ P_{ii}(t=0)=1 \\ \dots\dots\dots \end{cases} \quad (6)$$

С учетом (3) и (4) или (4а) для $P_{ii}(t)$ получим

$$P'_{ii}(t) = -v \cdot P_{ii}(t), \quad (7)$$

или

$$-v = \frac{P'_{ii}(t)}{P_{ii}(t)}. \quad (8)$$

После интегрирования (8) и потенцирования будем иметь

$$P_{ii}(t) = e^{-vt}. \quad (9)$$

Далее на основе (4а) для $P'_{i,i+1}(t)$ можем записать

$$P'_{i,i+1}(t) = -vP_{i(i+1)}(t) + vP_{ii}(t). \quad (10)$$

Этому уравнению соответствует

$$P_{i,i+1}(t) = vt \cdot e^{-vt}. \quad (11)$$

Проверки в виде определения $P'_{i,i+1}(t)$ подтверждают справедливость этого выражения.

Аналогичные процедуры позволяют оценить

$$P_{i,i+2}(t) = \frac{(vt)^2}{2!} e^{-vt}. \quad (12)$$

Используя далее принцип индукции, можем записать

$$P_{i,i+k}(t) = \frac{(vt)^k}{k!} e^{-vt}. \quad (13)$$

С учетом (9), (11)–(13) матрица $P(t)$ (2) примет вид

$$\begin{pmatrix} e^{-vt} & \frac{vt}{1} e^{-vt} & \frac{(vt)^2}{2!} e^{-vt} & \dots \\ 0 & e^{-vt} & \frac{vt}{1} e^{-vt} & \dots \\ 0 & 0 & e^{-vt} & \dots \end{pmatrix}. \quad (14)$$

Элементы матрицы (14) есть выражение закона Пуассона, который характеризуется параметром ν . Его общий вид совпадает с (13).

Рассмотрим более подробно особенности этого закона.

Кратко покажем основные свойства потоков событий (например, смен состояний, отказов объектов), которые являются цепями Маркова и которые могут выражаться законом Пуассона.

1. Поток стационарен. Это означает, что вероятность появления k событий на любом промежутке времени зависит только от числа k и от продолжительности интервала наблюдения t и не зависит от начала отсчета. Подтверждение этого следует из формулы (13): вероятность появления k событий, характеризующих состояния системы, зависит от $\nu = n \cdot \lambda$, т.е. от числа n – объема выборки, и заданной интенсивности (λ) появления событий: $\lambda = const$. Исходя из этого, появление k событий не зависит от начала отсчета времени.

2. В выражении (13) не используется информация о появлении событий до начала рассматриваемого промежутка, что характеризует свойства отсутствия последствия.

3. Убедимся, что выражение (13) отражает свойство ординарности. Положим $k=0$ и $k=1$, найдем вероятности непоявления событий и появления одного события (одного наблюдения событий):

$$\begin{cases} P(0) = \frac{1}{1!} \cdot e^{-\nu t} = e^{-\nu t}; \\ P(1) = \frac{(\nu t)^1}{1!} \cdot e^{-\nu t} = \nu t e^{-\nu t}. \end{cases} \quad (15)$$

Следовательно, вероятность появления более одного события

$$P(k > 1) = 1 - [e^{-\nu t} + \nu t e^{-\nu t}]. \quad (15a)$$

Пользуясь разложением

$$e^{-\nu t} = 1 - \nu t + \frac{(\nu t)^2}{2!} - \dots, \quad (15b)$$

получим

$$\begin{aligned} P(k > 1) &= 1 - \left[1 - \nu t + \frac{(\nu t)^2}{2!} + \nu t \left(1 - \nu t + \frac{(\nu t)^2}{2!} \right) \right] = \\ &= -\frac{(\nu t)^2}{2!} + \frac{(\nu t)^2}{1!} + \frac{(\nu t)^3}{2!} \doteq \frac{(\nu t)^2}{2!}. \end{aligned} \quad (16)$$

При малых значениях νt вероятность появления более одного события пренебрежимо мала, что следует из сравнения $P(1)$ и $P(k > 1)$.

Из изложенного следует, что закон Пуассона – модель потока событий, который является цепью Маркова.

Чтобы смысл параметров закона был более понятен, укажем на его взаимосвязь с другими законами.

Он является предельным выражением биномиального распределения вероятностей

$$P(k(t)) = \frac{n!}{k!(n-k)!} \cdot (\lambda t)^k (1-\lambda t)^{n-k}, \quad (17)$$

где $\lambda t = g$ – вероятность обнаружения одного события за время наблюдения t ; λ – интенсивность появления событий $\left(\frac{1}{\tau}\right)$; $k(t)$ – число событий, наблюдаемых за время t .

При числе экспериментов $n \rightarrow \infty$ и при условии, что $v = n\lambda$ остается конечной величиной, можем записать

$$P(k(t)) = \frac{n(n-1)\dots(n-k+1)}{k!} \cdot (\lambda t)^k (1-\lambda t)^{n-k} \doteq \frac{(vt)^k}{k!} \cdot e^{-vt}. \quad (18)$$

Представляет интерес проверка выполнения условия (марковости матрицы) нормирования суммы вероятностей в каждой строке матрицы (14). Для этого определим сумму $S(t)$:

$$S(t) = \sum_{k=0}^{\infty} \frac{(vt)^k e^{-vt}}{k!} = e^{-vt} \sum_{k=0}^{\infty} \frac{(vt)^k}{k!}. \quad (19)$$

Из математического анализа известно, что при $vt \ll 1$ сумма ряда

$$\sum_{k=0}^{\infty} \frac{(vt)^k}{k!} = e^{vt}. \quad (20)$$

Тогда

$$S(t) = e^{-vt} \cdot e^{vt} = 1. \quad (21)$$

Полученный результат (21) – свидетельство того, что условие нормировки вероятности в строках матрицы (14) (марковости) выполняется, при этом любой ее элемент удовлетворяет условию ≥ 0 .

С учетом характера приложений матрицу (14) можно интерпретировать так. Ее первая строка соответствует вероятности перехода системы (партии организмов) из исходного состояния $C1$ в состояние $C2$ с одним отклонением, в состояние $C3$ с двумя отклонениями и т.д. – теоретически до n_0 отклонений (всей генеральной совокупности). Сумма всех вероятностей в этой строке должна приближаться к 1.

Вторая строка соответствует исходному состоянию $C2$, когда генеральная совокупность составляет $(n_0 - 1)$ элементов и т.д.

Из изложенного также можно сделать заключение, что закон Пуассона можно использовать для оценивания эффективности статистического (выборочного) контроля (диагностирования) состояния объекта.

Суть выборочного контроля такова: из контролируемой партии (генеральной совокупности) биологических организмов отбирается выборка элементов и подвергается полному контролю. Если доля отклонений в выборке не превышает установленного значения, то всю популяцию (сообщество) можно считать не пострадавшей. Если доля отклонений в выборке превышает установленное значение, то популяция подвергается дальнейшей проверке.

Теперь кратко суть алгоритма выборочной проверки (диагностирования) можно представить в виде отдельных пунктов.

1. Отбирается выборка объемом n из n_0 – всей генеральной совокупности биологических организмов в исследуемом пространстве.

2. Назначается предельное значение числа отклонений $[k]$ в партии n .

3. При $k < [k]$ партия считается нормальной и вся совокупность n_0 принимается; если $k > [k]$, то элементы популяции $(n_0 - n)$ подвергаются дальнейшей дополнительной проверке.

4. При выполнении выборочного контроля (диагностирования) необходимо гарантировать установленную долю отклонений в принятой партии объема n . Для этого необходимо выбрать значения n и $[k]$, которые в свою очередь будут определять оперативную характеристику как вероятность события ($k < [k]$) при доле отклонений (g) и объеме выборки n :

$$P_0(k < [k]) = \sum_{k=0}^{[k]} P(k; n; g), \quad (22)$$

где 0 – индекс при $P_0(k < [k])$ – означает оперативная характеристика; k – текущее значение отклонений в партии объема n , $k \in \{1, 2, \dots, [k]\}$.

При достаточно малом отношении n/n_0 рекомендуется [4] использовать биномиальный закон распределения вероятностей (17) для вычисления $P_0(k < [k]; n; g)$ (22). Для этого (17) следует переписать так:

$$P_0(k; n; g) = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!} \cdot g^k \cdot (1-g)^{n-k}, \quad (23)$$

и тогда (22) можем представить в виде

$$P_0(k < [k]; n; g) = \sum_{k=0}^{[k]} P(k; n; g) = \sum_{k=0}^{[k]} \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{k!} g^k \cdot (1-g)^{n-k}. \quad (24)$$

Теперь (23) как приближение (22) в виде (13) запишем так:

$$P_0(k < [k]; n; g) \doteq \sum_k \frac{(g \cdot n)^k}{k!} \cdot e^{-(g \cdot n)}. \quad (25)$$

Из сравнения (24) и (25) следует, что при вариации g и n , но при $n \cdot g = \text{const}$, по формуле (24) будем получать различные результаты, а по формуле (25) – один и тот же результат на все случаи указанных вариаций. Это подтверждают данные из табл. 3.1.8 [4], в которых вариация объема выборки n определяется дискретным множеством $\{15; 30; 60; 150; 300\}$, а вероятность (g) – множеством $\{0,5; 0,2; 0,1; 0,05; 0,02\}$ при $g \cdot n = 3 = \text{const}$.

При анализе данных, соответствующих $[k] = 3$, и при $30 \leq n \leq 300$ оказывается, что пуассоновский закон дает результаты, совпадающие с биномиальным распределением, а при $[k] = 4$ дает занижение P_0 : при $n = 300$ – примерно на 0,001; при $n = 150$ – примерно на 0,002; при $n = 60$ – примерно на 0,045; при $n = 30$ – примерно на 0,1. При $[k] = 5$ эта тенденция сохраняется, но с меньшими значениями, при $[k] = 6$ тенденция меняет знак. При определенных исследованиях этот факт может вызвать сомнения в правильности методики расчетов, построенной на основе закона Пуассона [4]. Поэтому целесообразно вводить поправки в результат оценивания, основываясь на указанных закономерностях.

Список литературы

1. Мазей, Ю. А. Видовой состав и распределение раковинных амёб в некоторых водоемах и водотоках Среднего Поволжья / Ю. А. Мазей, А. В. Киреев, Е. А. Малышева // Известия ПГПУ им. В. Г. Белинского. – 2011. – № 25. – С. 519–523.
2. Рыжаков, В. В. Теория и алгоритмы диагностирования и прогнозирования состояния технических систем длительного применения на основе цепей Маркова: моногр. / В. В. Рыжаков, М. В. Рыжаков. Депонирована ВИНТИ РАН 27.03.2013, № 88-В2013. – 55 с.
3. Кельберт, М. Я. Вероятность и статистика в примерах и задачах. Марковские цепи как отправная точка теории случайных процессов и их приложения / М. Я. Кельберт, Ю. М. Сухов. – М.: Изд-во МЦНМО, 2010. – 559 с.
4. Смирнов, Н. В. Курс теории вероятностей и математической статистики для технических приложений / Н. В. Смирнов, И. В. Дунин-Барковский. – М.: Наука, 1969. – 512 с.

References

1. Mazei Yu. A., Kireev A. V., Malysheva E. A. *Izvestiya PGPU im. V. G. Belinskogo* [Proceedings of Penza State Pedagogical University named after V. G. Belinsky]. 2011, no. 25, pp. 519–523.
2. Ryzhakov V. V., Ryzhakov M. V. *Teoriya i algoritmy diagnostirovaniya i prognozirovaniya sostoyaniya tekhnicheskikh sistem dlitel'nogo primeneniya na osnove tsepey Markova: monogr.* [Theory and algorithms of diagnosing and forecasting the condition of long-term operation technical systems on the basis of Markov circuits: monograph]. Deponirovana VINITI RAN 27.03.2013, № 88-V2013. 55 p.
3. Kel'bert M. Ya., Sukhov Yu. M. *Veroyatnost' i statistika v primerakh i zadachakh. Markovskie tsepi kak otpravnyaya tochka teorii sluchaynykh protsessov i ikh prilozheniya* [Probability and statistics in examples and problems. Markov circuits as a starting point of the theory of random processes and application thereof]. Moscow: Izd-vo MTsNMO, 2010, 559 p.

4. Smirnov N. V., Dunin-Barkovskiy I. V. *Kurs teorii veroyatnostey i matematicheskoy statistiki dlya tekhnicheskikh prilozheniy* [Course of probability theory and mathematical statistics for technical application]. Moscow: Nauka, 1969, 512 p.
-

Рыжаков Виктор Васильевич

доктор технических наук, профессор,
заведующий кафедрой технического
управления качеством, заслуженный
деятель науки РФ, Пензенская
государственная технологическая
академия
(г. Пенза, проезд Байдукова, 1а)

E-mail: rvv@pgta.ru

Ryzhakov Viktor Vasil'evich

Doctor of technical sciences, professor,
head of sub-department of technical quality
control, Honoured Scientist of the Russian
Federation, Penza State Technological
Academy
(Penza, 1a Baydukova passage)

Рыжаков Михаил Викторович

старший преподаватель, кафедра
прикладной механики, заведующий
лабораторией прикладных
нанотехнологий, Московский
физико-технический институт
(государственный университет)
(Московская область, г. Долгопрудный,
Институтский переулок, 9)

E-mail: mryzhakov@applmech.mipt.ru

Ryzhakov Mikhail Viktorovich

Senior lecturer, sub-department of applied
mechanics, head of the laboratory
of applied nanotechnologies, Moscow
Institute of Physics and Technology
(State University)
(Moscow Region, Dolgoprudny,
9 Institutsky Lane)

УДК 519.21.248+57+574

Рыжаков, В. В.

Теория и алгоритм диагностирования биологических систем с бесконечным числом состояний на основе использования соответствующих *Q*-матриц и матриц вероятностей переходов / В. В. Рыжаков, М. В. Рыжаков // Известия высших учебных заведений. Поволжский регион. Естественные науки. – 2013. – № 2 (2). – С. 41–48.