

УДК 519.217.2

## ОБЗОР МЕТОДОВ И РЕАЛИЗАЦИЯ АЛГОРИТМА СКРЫТОЙ МАРКОВСКОЙ МОДЕЛИ ДЛЯ АНАЛИЗА «ВЫЖИВАЕМОСТИ» ИННОВАЦИОННЫХ ПРОЕКТОВ

Рашидов А.Р., Шмидт И.А.

*Пермский национальный исследовательский политехнический университет, Пермь,  
e-mail: Rashidov.A.R@ieee.org, shmidt@msa.pstu.ac.ru*

В данной статье рассматривается вопрос определения «выживаемости» инновационных проектов. Под «выживаемостью» понимается вероятность, с которой инновационный проект будет реализован. Для оценки «выживаемости» инновационных проектов применяется математический аппарат скрытых Марковских моделей. В данном исследовании решается задача определения параметров модели, которая с наибольшей вероятностью воспроизводит заданные последовательности данных. В качестве исходных данных используются данные об инновационных проектах, содержащиеся в отчетах российских программ «УМНИК» и «СТАРТ», а также дополнительные данные, полученные в ходе исследования. В качестве алгоритма для определения параметров модели используется алгоритм Баума – Велша, который является одной из реализаций алгоритма Expectation-maximization. В заключительной части статьи описывается математический аппарат для практической реализации, в частности приводятся математическое описание алгоритма и методы реализации Марковских моделей.

**Ключевые слова:** инновационный проект, скрытая Марковская модель, Баума – Велша алгоритм, «выживаемость» проекта

## REVIEW OF METHODS AND IMPLEMENTATION OF THE HIDDEN MARKOV MODEL ALGORITHM FOR INNOVATIVE PROJECTS «VIABILITY» ANALYSIS

Rashidov A.R., Shmidt I.A.

*Perm National Research Polytechnic University, Perm, e-mail: Rashidov.A.R@ieee.org,  
shmidt@msa.pstu.ac.ru*

This article deals with determining of innovative projects «viability». «Viability» is the probability of innovative project being implemented. Hidden Markov Models are used for evaluation of this factor. The problem of determining parameters of model, which produce given data sequence with the highest probability, are solving in this research. Data about innovative projects contained in reports of Russian programs «UMNIK» and «START» and additional data obtained during study are used as input data for determining of model parameters. The Baum-Welch algorithm which is one implementation of expectation-maximization algorithm is used at this research for calculating model parameters. At the end part of the article mathematical basics for practical implementation are given (in particular mathematical description of the algorithm and implementation methods for Markov models).

**Keywords:** innovative project, Markov process, hidden Markov model, Baum-Welch algorithm, project «viability»

В настоящее время появляется много инновационных проектов в различных областях науки, которые не всегда находят применение и реализацию в своем развитии.

**Инновационный проект** представляет собой сложную систему взаимосвязанных и взаимоувязанных по ресурсам, срокам и исполнителям мероприятий, направленных на достижение конкретных целей и задач на приоритетных направлениях развития науки и техники [3].

Целью данного исследования является определение «выживаемости» инновационных проектов. Под «выживаемостью» в данной работе понимается вероятность, с которой инновационный проект будет реализован.

Актуальность исследования обусловлена следующими факторами:

– большое количество инновационных проектов, представляемых в различных программах;

– необходимость определения характеристик «идеального» инновационного проекта, т.е. такого проекта, который точно будет реализован в дальнейшем.

К основным фазам жизненного цикла инновационного проекта относятся:

1. Предынвестиционная фаза – выполнение различных исследований для планирования инновационного проекта.

2. Инвестиционная фаза – составление документации по реализации инновационного проекта.

3. Фаза торгов и контрактов – на данной фазе заключаются контракты на поставку различного оборудования и выполнения работ.

4. Фаза реализации проекта – разработка планов и графиков выполнения работ по реализации проекта.

5. Завершающая фаза проекта – на данной фазе происходит закрытие инновационного проекта, запуск объекта в экс-

плуатацию и начало реализации товара, анализируются результаты работы.

В качестве исходных данных для оценки вероятности используются данные о проектах, участвующих в программах «УМНИК», «СТАРТ» и «УМНИК на СТАРТ», реализуемых Фондом содействия развитию малых форм предприятий в научно-технической сфере. Таким образом, «выживаемость» может быть определена как вероятность, с которой проекты пройдут все стадии в программах «УМНИК», «СТАРТ», «УМНИК на СТАРТ» и впоследствии станут успешно реализованы.

Программа «УМНИК» – это федеральная программа содействия развитию молодежных инновационных проектов [6]. В программе принимают участие только физические лица (программа ориентирована преимущественно на молодых специалистов, не имеющих опыта в предпринимательской деятельности).

Программа «СТАРТ» – это федеральная программа содействия развитию малых инновационных предприятий [4]. В программе принимают участие только юридические лица.

Программа «УМНИК на СТАРТ» – программа молодежных инновационных проектов, позволяющая победителям участвовать в программе «СТАРТ» [5]. В программе принимают участие победители программы «УМНИК», которые готовы создать малое инновационное предприятие для участия в программе «СТАРТ».

В ходе проведения данного магистерского исследования становится необходимым поиск метода определения «выживаемости» инновационных проектов. Были рассмотрены несколько методов, которые могут позволить адекватно оценить вероятность прохождения инновационными проектами всех стадий участия в федеральных программах («УМНИК», «СТАРТ», «УМНИК на СТАРТ») и дальнейшую реализацию.

Для определения «выживаемости» инновационных проектов рассмотрено несколько методов: искусственные нейронные сети, генетические алгоритмы и Марковская модель.

**Искусственная нейронная сеть** – это система, состоящая из многих простых вычислительных элементов, работающих параллельно, функция которых определяется структурой сети, силой взаимных связей, а вычисления производятся в самих элементах или узлах [7].

**Генетические алгоритмы** – алгоритмы, предназначенные для решения задач оптимизации. Примером подобной задачи может служить обучение нейронной сети,

то есть подбора таких значений весов, при которых достигается минимальная ошибка. При этом в основе генетического алгоритма лежит метод случайного поиска [1].

Случайный процесс в некоторой системе называется **Марковским**, если вероятность перехода системы в новое состояние зависит только от состояния системы в настоящий момент и не зависит от того, когда и каким образом система перешла в это состояние [2].

В результате исследования, для анализа «выживаемости» инновационных проектов, был выбран метод Марковских моделей, так как данный метод изначально учитывает Марковский характер жизненного цикла инновационного проекта.

Возникает проблема поиска неизвестных (скрытых) вероятностей перехода инновационных проектов по этапам участия в программах: «УМНИК», «СТАРТ» и «УМНИК на СТАРТ». Данная проблема решается с помощью подтипа Марковских процессов – скрытой Марковской моделью.

**Скрытая Марковская модель** – модель процесса, в которой процесс считается Марковским, причем неизвестно, в каком состоянии  $S_i$  находится система (состояния скрыты), но каждое состояние  $S_i$  может с некоторой вероятностью  $b_{oj}$  произвести событие  $o_j$ , которое можно наблюдать [2].

Обычно, при работе со скрытой Марковской моделью возникает несколько задач определения скрытых параметров. Для данного исследования воспользуемся задачей, которая определяет параметры модели  $\lambda = (P, O, \pi_1)$ , обеспечивающей максимальную вероятность  $P[O|\lambda]$ , где  $P = [p_{ij}]$  – вероятность состояния перехода,  $O = [o_j(m)]$  – вероятность наблюдений,  $\pi_1 = (\pi_{1,j})$  – вероятность начального состояния.

Структура Марковской модели должна быть известна. Необходимо определить начальные вероятности каждого состояния и вероятности перехода между состояниями в данной модели. Для решения поставленной задачи применяется алгоритм Баума – Велша [8].

Алгоритм состоит из двух шагов. На первом шаге из известных параметров модели вероятность нахождения системы во время  $t$  в состоянии  $S_i(\gamma_t(i))$  и вероятность нахождения системы во время  $t$  в состоянии  $S_i$  и во времени  $t + 1$  в состоянии  $S_j$  данной последовательности наблюдений и модели  $(\xi_t(i, j))$  вычисляется по формулам

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O | \lambda),$$

$$\gamma_t(i) = \sum \xi_t(i, j),$$

где  $\xi_t(i, j)$  – вероятность нахождения во время  $t$  в состоянии  $S_i$  и во время  $t + 1$  в со-

стоянии  $S_j$  для данной последовательности наблюдений и модели,  $\gamma_t(i)$  – вероятность нахождения во время  $t$  в состоянии  $S_i$ .

На втором шаге модель  $\lambda$  вычисляется на основе значений  $\gamma_t(i)$ ,  $\xi_t(i, j)$ .

Повторяем первый шаг с вычисленной  $\lambda$ . При многократном повторении этой процедуры вероятность совпадения вычисленной функции с исходной возрастает. Алгоритм

сходится в тот момент, когда вероятность перестает изменяться.

В данном исследовании: набор состояний  $S = \{S_1, S_2, \dots, S_N\}$  – состояния по программам; набор наблюдений  $V = \{v_1, v_2, \dots, v_M\}$  – известные данные об инновационных проектах (количество инновационных проектов, количество заключенных контрактов и т.д.); модель  $\lambda = (P, O, \pi_1)$ .

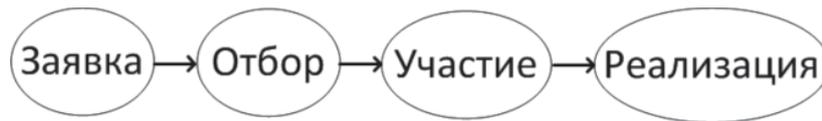


Рис. 1. Этапы прохождения инновационным проектом стадий программ

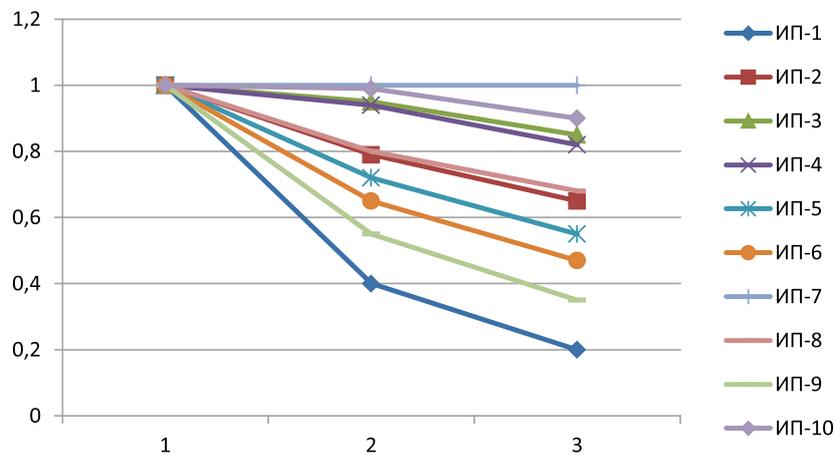


Рис. 2. Влияние критерия «Объем финансирования» на прохождение этапов инновационного проекта федеральных программ

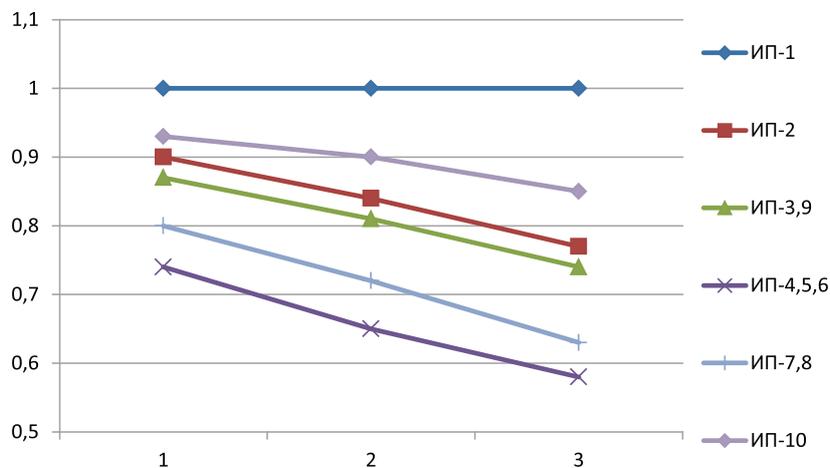


Рис. 3. Влияние критерия «Количество потенциальных заказчиков» на прохождение этапов инновационного проекта федеральных программ

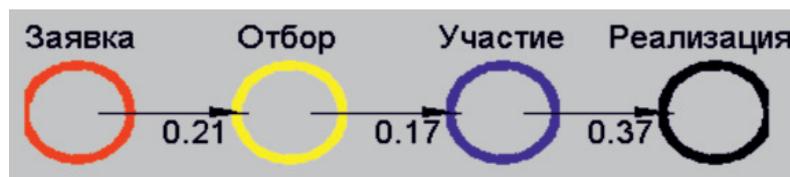


Рис. 4. Предсказанные, скрытые вероятности перехода инновационного проекта

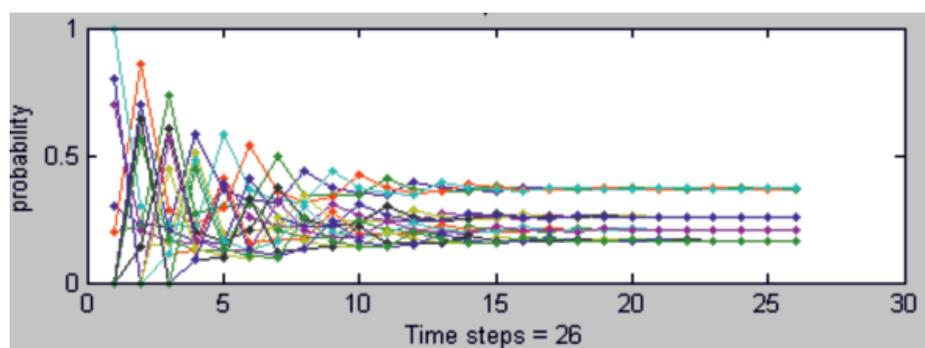


Рис. 5. Графическое представление итераций алгоритма Баума – Велша

Для исследования «выживаемости» инновационных проектов были выбраны необходимые данные по нескольким инновационным проектам из отчетов Фонда содействия развитию малых форм предприятий в научно-технической сфере за 2010–2013 гг.

Критериями, необходимыми для адекватного анализа «выживаемости» инновационных проектов, являются:

- Объем финансирования (т.р.).
- Окупаемость.
- Срок реализации.
- Количество статей по тематике.
- Количество потенциальных заказчиков.

Следующим шагом необходимо оценить влияние каждого критерия на этапы прохождения инновационным проектом стадий программ («УМНИК», «СТАРТ», УМНИК на СТАРТ»). Для более адекватной оценки влияния критериев объединим схожие этапы программ. В результате для оценки имеется три наиболее важных этапа (рис. 1).

На этапе оценки влияния каждого критерия инновационного проекта на стадии прохождения федеральных программ, наблюдается зависимость значения вероятности успешного прохождения какой-либо стадии программ от значения критерия. Для наглядности выведем графики критериев «Объем финансирования (т.р.)» (рис. 2) и «Количество потенциальных заказчиков» (рис. 3). На оси  $y$  отмечаются вероятности

успешного прохождения стадии программ, а на оси  $x$  отмечаются стадии программ (1 – заявка, 2 – отбор, 3 – участие).

На завершающем этапе данного исследования необходимо практически реализовать алгоритм Баума – Велша и обучить скрытую Марковскую модель для предсказания вероятностей перехода какого-либо между стадиями федеральных программ («УМНИК», «СТАРТ», «УМНИК на СТАРТ»).

Обучение скрытой Марковской модели происходит с помощью алгоритма Баума – Велша и статистическими данными имеющейся выборки инновационных проектов в программном комплексе MATLAB.

Далее происходит получение данных о «новом», не входящем в статистическую выборку инновационном проекте. Данные об этом инновационном проекте загружаются в модель и происходит расчет скрытых вероятностей перехода (рис. 4). На рис. 5 наглядно представлен алгоритм Баума – Велша, который рассчитывает скрытые вероятности «нового» инновационного проекта и продолжает данные итерации до схождения.

В ходе выполнения итераций наблюдаем, что все параметры сходятся примерно на двадцатом шаге реализации алгоритма Баума-Велша.

Определение «выживаемости» позволит дать вероятностный прогноз относительно перспектив реализации инновационного

проекта, также позволит оценить финансовые риски и другие показатели, связанные с их реализацией.

#### Список литературы

1. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы: Учебное пособие. – 2-е изд. – М: Физматлит, 2006. – С. 320.
2. Марковские процессы / Портенко Н.И., Скороход А.В., Шуренков В.М. – 1989. – 765 с.
3. Основы инновационного менеджмента. Учебное пособие / Под. ред. проф. Завалина П.Н. М.: Экономика, 2000. – 475 с.
4. Программа «СТАРТ». [Электронный ресурс]. URL: <http://www.fasie.ru/programs/programma-start/> (дата обращения: 09.02.2016).
5. Программа «УМНИК на СТАРТ». [Электронный ресурс]. URL: <http://www.nauka-nov.ru/fsrmpnts/665/> (дата обращения: 09.02.2016).
6. Программа «УМНИК». [Электронный ресурс]. URL: <http://www.fasie.ru/programs/programma-umnik/> (дата обращения: 09.02.2016).
7. Суровцев И.С., Клюкин В.И., Пивоварова Р.П. Нейронные сети. – Воронеж: ВГУ, 1994. – 224 с.

8. Geoffrey McLachlan and Thriyambakam Krishnan. The EM Algorithm and Extensions. John Wiley & Sons, New York, 1996.

#### References

1. Gladkov L.A., Kurejchik V.V., Kurejchik V.M. Geneticheskie algoritmy: Uchebnoe posobie. 2-e izd. M: Fizmatlit, 2006. pp. 320.
2. Markovskie processy / Portenko N.I., Skorohod A.V., Shurenkov V.M. 1989. 765 p.
3. Osnovy innovacionnogo menedzhmenta. Uchebnoe posobie / Pod. red. prof. Zavalina P.N. M.: Jekonomika, 2000. 475 p.
4. Programma «START». [Jelektronnyj resurs]. URL: <http://www.fasie.ru/programs/programma-start/> (data obrashhenija: 09.02.2016).
5. Programma «UMNIK na START». [Jelektronnyj resurs]. URL: <http://www.nauka-nov.ru/fsrmpnts/665/> (data obrashhenija: 09.02.2016).
6. Programma «UMNIK». [Jelektronnyj resurs]. URL: <http://www.fasie.ru/programs/programma-umnik/> (data obrashhenija: 09.02.2016).
7. Surovcev I.S., Kljukin V.I., Pivovarova R.P. Nejrionnye seti. Voronezh: VGU, 1994. 224 p.
8. Geoffrey McLachlan and Thriyambakam Krishnan. The EM Algorithm and Extensions. John Wiley & Sons, New York, 1996.