УДК 51-75:519.61

К. А. БАЗИЛЕВИЧ, М. С. МАЗОРЧУК, Н. С. БАКУМЕНКО

Национальный аэрокосмический университет им. Н. Е. Жуковского «ХАИ», Украина

АЛГОРИТМИЧЕСКАЯ МОДЕЛЬ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ МОДЕЛИ ПОПУЛЯЦИОННОЙ ДИНАМИКИ СТАРЕНИЯ ГОМПЕРЦА

Предложена алгоритмическая модель для решения задачи нахождения параметров модели популяционной динамики старения Гомперца. Предложен подход к решению поставленной задачи путем сведения ее к системе уравнений, полученных с помощью метода наименьших квадратов на основе анализа интенсивности смертности по прошлым периодам, что обеспечивает быстрое нахождение искомых параметров и дальнейшее моделирование процесса изменения популяционной динамики. Проведен сравнительный анализ результатов, полученных по разным методам. Оценена достоверность полученных результатов на основе статистических данных трех стран: Украины, России и Германии. Отклонения результатов моделирования от фактических данных минимальны и находятся в пределах нормы.

***Ключевые слова**: аналитические модели смертности, геронтология, популяционная динамика старения, кривая смертности, метод наименьших квадратов.*

Введение

Процесс изменения численности популяции населения является слабоизученным и изменчивым. Численность населения зависит от многих факторов: уровень и качество жизни, врожденные и приобретенные заболевания, влияние внешней среды и многое другое.

На данный момент существует целое множество подходов к моделированию процесса смертности, к изучению механизмов и причин старения, а также к количественному и качественному анализу этого процесса [1-3].

Данные подходы используются в разных областях: биологии, зоологии, в различных сферах медицины и генетики. Также, процесс смертности, а точнее, его численные характеристики изучаются математиками, актуариями при расчетах тарифов страхования жизни и медицинского страхования.

Главной задачей биологов, зоологов и медиков является изучение характера старения человеческого организма, выявление тех факторов, которые имеют непосредственное влияния на этот процесс, а также возможность управления этими факторами.

Главной задачей математиков и актуариев является получение численных характеристик процесса смертности. Исходя из полученных данных можно рассчитать вероятности смерти или дожития в определенном возрасте, что является основой для определения финансово-экономических показателей процесса страхования.

Результаты, полученные в ходе подобного моделирования, имеют большое значение в тех случаях, когда нет возможности использовать реальные статистические выборки.

Постановка задачи исследования

Вычисление параметров моделирования является важнейшей задачей. Описание моделей аналитических законов представлено в целом ряде зарубежных и отечественных работ [4-6], однако отсутствуют алгоритмические модели и методики расчета параметров аналитических моделей, что затрудняет использование данных результатов в практике страхования.

Целью исследования является разработка алгоритмической модели оценки параметров аналитических моделей популяции старения на основе численных методов.

В данной работе рассматривается моделирование процесса старения популяции с помощью аналитического закона Гомперца. Преимущество такого подхода заключается в том, что для моделирования необходимо найти всего лишь два параметра, которые непосредственно влияют на вид кривой смертей. Для сравнения результатов, полученных при моделировании с реальными данными была выбрана статистическая информация Всемирной Организации Здоровья о смертности населения в Украине, России и Германии [7].

Следует отметить, что данные носили точечный характер, т.е. количество людей, оставшихся в живых из 100000 новорожденных было известно лишь в определенных, ключевых точках, с разницей в 5 лет (рис. 1). Для того чтобы иметь полную картину данного процесса была проведена линейная интерполяция по ключевым точкам на основе данных, представленных в [7]. Полученные результаты для каждой из стран можно увидеть на рисунке 1.

В модели Гомперца функция выживания имеет вид:

$$s(x) = \exp[-\int_0^x \mu_u du] = \exp[-\int_0^x Be^{\alpha u} du] =$$
$$= \exp[-B(e^{\alpha x} - 1)/\alpha],$$

где $\alpha > B > 0$ – некоторые параметры, которые задают вид кривой смертей.

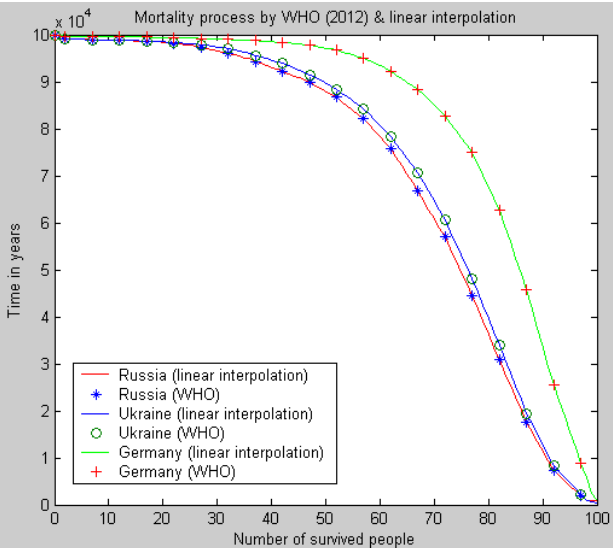

Рис. 1. Популяционная динамика на 100000 новорожденных, Украина, Россия, Германия

Кривая смертей определяется функцией:

$$f(x) = \mu_x s(x) = B \exp\left[\alpha x - B(e^{\alpha x} - 1)/\alpha\right] [5].$$

Рассмотрим нахождение параметров данной аналитической модели с использованием различных численных методов.

Определение параметров модели на основе данных о максимуме и квартили кривой смертей

Рассмотрим подход к определению параметров аналитической модели на основе системы уравнений, приведенной в [6].

Для этого на основе статистических данных по Украине [7] найдем возраст, в котором смертность максимальна – это возраст $m = 82$ лет. Квартиль смертности на основе той же статистической выборки составляет 27 лет.

Составим систему уравнений, с помощью которой можно будет найти необходимые параметры:

$$\begin{cases} \dfrac{\ln(\alpha) - \ln(B)}{\alpha} = 82 \\ 1 - e^{\frac{-B(e^{\alpha 27} - 1)}{\alpha}} = 0.25 \end{cases}$$

Вектор-функция примет вид:

$$f(\alpha, B) = \begin{cases} \dfrac{\ln(\alpha) - \ln(B)}{\alpha} - 82 \\ 1 - e^{\frac{-B(e^{\alpha 27} - 1)}{\alpha}} - 0.25 \end{cases}$$

Решим данную систему уравнений с помощью метода Ньютона в математическом пакете MathCad. Метод Ньютона чувствителен к начальному приближению, поэтому при разных начальных условиях решения будут различны.

Полученное решение системы уравнений и найденные параметры: $\alpha = 0.0318917$, B=0.0023295.

Проведем моделирование на основе полученных значений параметров. Для автоматизации расчетов и представления графических результатов была разработана программа «АММ», графическое представление результата в программе можно увидеть на рисунке 2.

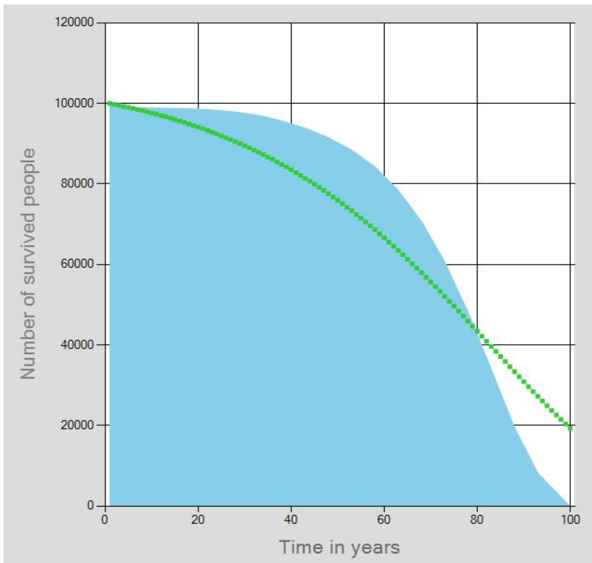

Рис. 2. Кривая смертности, полученная с помощью моделирования по аналитическому закону Гомперца и реальная смертность, Украина (параметры рассчитаны на основе данных о максимуме и квартили распределения)

Среднюю ошибку моделирования найдем по формуле:

$$\mathrm{Dev} = \sum_{x=1}^{\gamma} \frac{\left[\frac{|l_x\_Stat - l_x\_Mod|}{l_0} \cdot 100\right]}{\gamma}, \quad (1)$$

где l_x_Stat - количество людей, оставшихся в живых в возрасте x (по статистическим данным);

l_x_Mod - количество людей, оставшихся в живых в возрасте x (с помощью моделирования);

$l_0 = 100000$ - начальная популяция (количество новорожденных);

x - возраст;

γ – максимальный возраст популяции.

Для данной методики средняя ошибка моделирования по формуле (1) составляет 9.46 %, что является довольно большой величиной. Это можно увидеть на рисунке 2.

Также были построены кривые смертности для таких стран, как Германия и Россия. Результаты дали большие расхождения по сравнению с эмпирическими данными.

Полученные параметры не являются оптимальными для этой модели, так как кривая старения имеет слишком плавный спуск.

Определение параметров модели на основе анализа функции интенсивности смертности

Интенсивность смертности для модели Гомперца рассчитывается по формуле:

$$\mu_x = \frac{f(x)}{s(x)} = Be^{\alpha x}.$$

Определим функцию:

$$N(\alpha, B, x) = Be^{\alpha x}. \quad (2)$$

Прологарифмируем данную функцию:

$$\ln(N(\alpha, B, x)) = \ln(B) + \alpha \cdot x,$$

где обозначим слагаемые:

$$K = \alpha,\ T = \ln(B), \quad G(x) = \ln(N(\alpha, B, x)) = Kx + T. \quad (3)$$

Для нахождения параметров K и T будем использовать метод наименьших квадратов [8] и метод Крамера.

На основании утверждения из [6] будем приближать интенсивность смертности по формуле:

$$\mu_x \approx q_x = \frac{d_x\_Stat}{l_x\_Stat}, \quad (4)$$

где q_x - вероятность смерти в возрасте x, не дожив до возраста x+1,

d_x_Stat - количество людей, которые умерли при переходе из возраста x в возраст x+1,

l_x_Stat - количество людей, оставшихся в живых в возрасте x.

Данные могут быть найдены на основе статистики за прошлые периоды.

Составим систему уравнений для нахождения искомых параметров, используя метод наименьших квадратов:

$$\begin{cases} K \cdot \sum_{i=1}^{n} x_i^2 + T \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i \cdot G(x_i) \\ K \cdot \sum_{i=1}^{n} x_i + n \cdot T = \sum_{i=1}^{n} G(x_i), \end{cases} \quad (5)$$

где n – количество лет наблюдения (количество моментов времени x_i).

Искомые параметры могут бытть найдены по формуле:

$$\alpha = K,\ B = e^T. \quad (6)$$

На основе статистических данных по Украине [7] найдем μ_x в ключевых точках и прологарифмируем полученные значения (табл. 1).

Найдем вспомогательные параметры для дальнейшего расчета методом наименьших квадратов:

$$\sum_{i=1}^{20} x_i = 990,\ \sum_{i=1}^{20} G(x_i) = -96.7345,$$

$$\sum_{i=1}^{20} x_i \cdot G(x_i) = -3514.988,\ \sum_{i=1}^{20} x_i^2 = 65630.$$

Подставим численные значения:

$$\begin{cases} K \cdot 65630 + T \cdot 990 = -3514.988 \\ K \cdot 990 + 20 \cdot T = -96.7345. \end{cases}$$

Таблица 1

Значения функции $G(x_i) = \ln(\mu_{x_i})$ при различных значениях x_i

Моменты времени x_i	μ_{x_i}	$G(x_i) = \ln(\mu_{x_i})$
2	0.0003	-8.0379
7	0.0002	-8.4293
12	0.0003	-8.2821
17	0.0006	-7.4496
22	0.0011	-6.8562
27	0.0017	-6.3545
32	0.0029	-5.8566
37	0.0037	-5.5861
42	0.0050	-5.2946
47	0.0067	-5.0030
52	0.0095	-4.6605
57	0.0137	-4.2888
62	0.0198	-3.9234
67	0.0285	-3.5586
72	0.0412	-3.1891
77	0.0590	-2.8295
82	0.0853	-2.4617
87	0.1155	-2.1585
92	0.1449	-1.9315
97	0.5583	-0.5829

Данную систему линейных алгебраических уравнений можно решить с помощью метода обратной матрицы, методом Крамера либо последовательных исключений Гаусса. Параметры K и T были получены в программе MathCad с помощью функции Isolve(). Результат можно увидеть на рисунке 3.

$$a := \begin{pmatrix} 65630 & 990 \\ 990 & 20 \end{pmatrix} \quad b := \begin{pmatrix} -3514.988 \\ -96.7345 \end{pmatrix}$$

$$\text{lsolve}(a, b) = \begin{pmatrix} 0.0765937 \\ -8.6281116 \end{pmatrix}$$

Рис. 3. Расчет параметров аналитического закона Гомперца в пакете MathCad

Таким образом, искомые параметры аналитического закона будут равны:

$$\alpha = K = 0.07659, \; B = e^T = e^{-8.628} = 0.000179.$$

Графическое представление результата можно увидеть на рисунке 4.

Для данного закона, при составлении системы уравнений с использованием метода наименьших квадратов, средняя ошибка моделирования составит 1,67% по формуле (1), что говорит о правильно подобранных параметрах и положительных результатах моделирования.

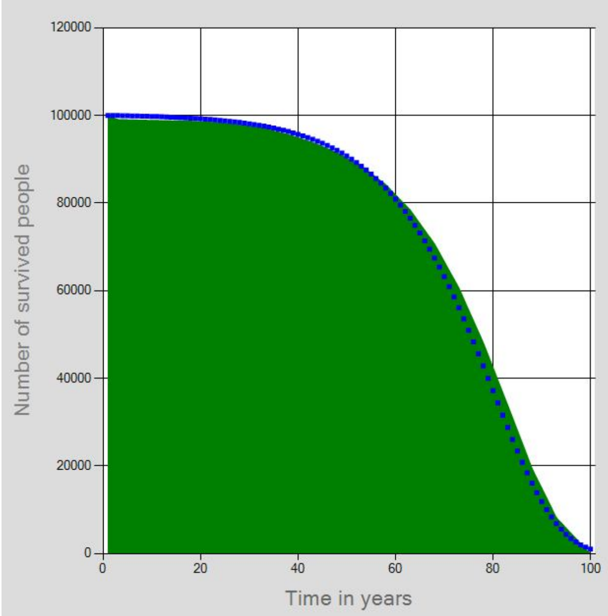

Рис. 4. Кривая смертности, полученная с помощью моделирования по аналитическому закону Гомперца и реальная смертность, Украина (параметры рассчитаны на основе функции интенсивности смертности)

Найдем параметры для аналитической модели Гомперца и проведем моделирование на примерах статистики по России и Германии.

Для России параметры аналитической модели кривой смертности равны:

$$\alpha = 0.07296956692, \; B = 0.00024862.$$

Графическое представление результата можно увидеть на рисунке 5. Средняя ошибка моделирования по формуле (1) не превышает 1.7 %.

Для Германии параметры моделирования равны: $\alpha = 0.086329$, B=0.00004412.

Графическое представление результата можно увидеть на рисунке 6. Средняя ошибка моделирования по формуле (1) не превышает 2.07 %.

Можно сделать вывод о том, что расчет параметров модели Гомперца на основе анализа функции интенсивности смертей с использованием метода наименьших квадратов дает более точные результаты.

Алгоритм нахождения параметров для аналитического закона Гомперца при помощи метода наименьших квадратов и метода Крамера

При наличии статистических данных об интенсивности смертности за прошлые периоды можно предложить следующий алгоритм проведения расчетов нахождения параметров кривой смертности.

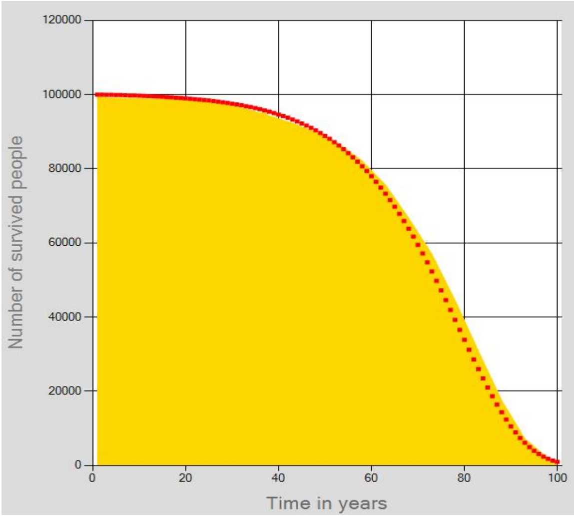

Рис. 5. Кривая смирности, полученная с помощью моделирования по аналитическому закону Гомперца и реальная смертность, Россия (параметры рассчитаны на основе функции интенсивности смертности)

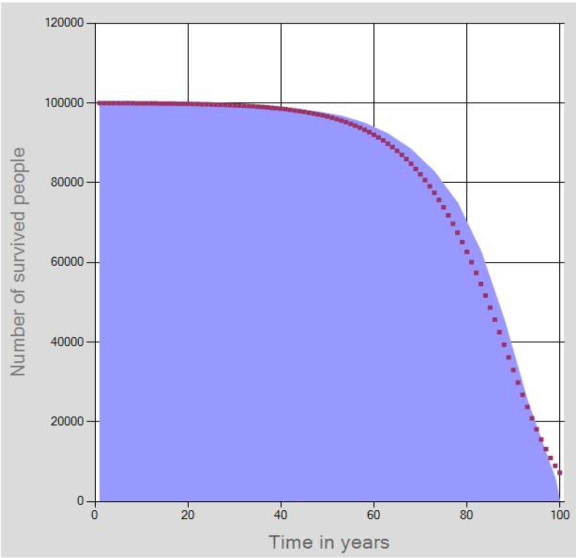

Рис. 6. Кривая смертности, полученная с помощью моделирования по аналитическому закону Гомперца и реальная смертность, Германия (параметры рассчитаны на основе функции интенсивности смертности)

Структура входных данных представлена в таблице 2.

Рассмотрим алгоритм расчета.

Этап 1. Введем исходные данные, представленные в таблице 2.

Этап 2. Рассчитаем значения функции μ_x по формуле (4).

Этап 3. Рассчитаем значения функции $G(x_i) = \ln(\mu_{x_i})$.

Этап 4. Найдем параметр K, используя метод Крамера:

$$K = \frac{\left[\sum_{i=1}^{n} x_i \cdot G(x_i)\right] \cdot n - \sum_{i=1}^{n} G(x_i) \cdot \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 \cdot n - (\sum_{i=1}^{n} x_i)^2}.$$

где n – количество лет наблюдения (количество моментов времени x_i).

Этап 5. Найдем параметр T, используя метод Крамера:

$$T = \frac{\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} G(x_i) - \sum_{i=1}^{n} x_i \cdot \left[\sum_{i=1}^{n} x_i \cdot G(x_i)\right]}{\sum_{i=1}^{n} x_i^2 \cdot n - (\sum_{i=1}^{n} x_i)^2}.$$

Этап 6. Найдем параметры α, B по формуле (6).

Этап 7. Проведем моделирование по закону Гомперца с полученными параметрами α, B:

a) зададим значение начальной популяции l_0.

b) рассчитаем значения функции выживания в моменты времени $x_i : s(x) = \exp[-B(e^{\alpha x} - 1)/\alpha]$.

c) определим количество людей, оставшихся в живых в моменты времени x_i: $l_{x\_Mod} = l_0 s(x)$.

Этап 8. Оценим среднюю ошибку моделирования по формуле (1).

Таблица 2

Входные данные

Название параметра	Назначение и интервал
. x_i .	Моменты времени, которые выбираются для моделирования (см. табл. 1) $x_i \in [0 \ldots \gamma]$, где γ – максимальный возраст популяции, i – число моментов времени, $i \in [1 \ldots n]$.
l_0	Начальное значение популяции. Параметр может принимать любое значение, при условии $l_0 > 0$.
d_x _Stat	Количество людей, которые умерли при переходе из возраста x в возраст $x+1$ на основании статистики.
l_x _Mod	Количество людей, оставшихся в живых в возрасте x на основании статистики. Параметр находится в интервале $[0 \ldots l_0]$.

Заключение

Таким образом, для оценки параметров аналитической модели популяционной динамики старения Гомперца (кривой смертности Гомперца) целесообразно использовать рассмотренный алгоритм, основанный на анализе интенсивности смертности с применением метода наименьших квадратов и метода Крамера.

Предложенная алгоритмическая модель позволит на практике проводить расчеты оперативно и качественно. Ошибка по результатам моделирования составляет не более 3%.

Литература

1. *Анисимов, В. Н. Молекулярные и физиологические механизмы старения [Текст] / В. Н. Анисимов. – СПб. : Наука, 2003. – 468 с.*
2. *Дильман, В. М. Большие биологические часы [Текст] / В. М. Дильман. – М. : Знание, 1986. – 256 с.*
3. *Фролькис, В. В. Старение, эволюция и продление жизни [Текст] / В. В. Фролькис, Х. К. Мурадян. – Киев : Наук. думка, 1992. – 336 с.*
4. *Gompertz, B. On the Nature of the Function Expressive of the Law of Human Mortality [Text] / B. Gompertz // Phil. Trans. Royal Soc. London. 1825. – 60 p.*
5. *Актуарная математика [Текст] : моногр. / Н. Бауэрс, Х. Гербер, Д. Джонс, С. Несбитт, Дж. Хикман. – М. : Янус-К, 2001. – 621 с.*
6. *Кошкин, Г. М. Основы актуарной математики [Текст] : учеб. пособие / Г. М. Кошкин ; Мин-во образования и науки России, Томск. гос. ун-т. – Томск : Изд-во Томск. гос. ун-та, 2002. – С. 14, 19-20.*
7. *Country life tables by Global Health Observatory of World Health Organization [Electronic resource]. – An access mode: http://apps.who.int/gho/data/?theme=main&vid=61740. – 3.09.2014.*
8. *Айвазян, С. А. Прикладная статистика. Том 2: Основы эконометрики [Текст] / С. А. Айвазян. – М. : Юнити-Дана, 2001. – 432 с.*

Поступила в редакцию 3.06.2014, рассмотрена на редколлегии 11.06.2014

Рецензент: д-р техн. наук, проф., зав. каф. охраны труда, стандартизации и сертификации Р. М. Трищ, Украинская инженерно-педагогическая академия, Харьков.

АЛГОРИТМІЧНА МОДЕЛЬ ВИЗНАЧЕННЯ ПАРАМЕТРІВ МОДЕЛІ ПОПУЛЯЦІЙНОЇ ДИНАМІКИ СТАРІННЯ ГОМПЕРЦА

К. О. Базілевич, М. С. Мазорчук, Н. С. Бакуменко

Запропоновано алгоритмічну модель для вирішення задачі знаходження параметрів моделі популяційної динаміки старіння Гомперца. Запропоновано підхід до вирішення поставленого завдання шляхом побудови системи рівнянь, отриманих за допомогою методу найменших квадратів на основі аналізу інтенсивності смертності за минулі періоди, що забезпечує швидке знаходження шуканих параметрів і подальше моделювання процесу зміни популяційної динаміки. Проведено порівняльний аналіз результатів, отриманих за різними методами. Оцінено достовірність отриманих результатів на основі статистичних даних трьох країн: України, Росії та Німеччини. Відхилення результатів моделювання від фактичних даних мінімальні і знаходяться в межах норми.

Ключові слова: аналітичні моделі смертності, геронтологія, популяційна динаміка старіння, крива смертності, метод найменших квадратів.

ALGORITHMIC MODEL FOR DETERMINING THE PARAMETERS OF THE MODEL POPULATION DYNAMICS BY GOMPERZ

K. A. Bazilevich, M. C. Mazorchuk, N. C. Bakumenko

Proposed algorithmic model to solve the problem of finding the model parameters population dynamics by Gomperz. An approach to construction the problem base on system of equations obtained by the least squares method by analyzing the intensity of mortality from the past, providing quick location of desired parameters and further modeling of the process of population dynamics change. A comparative analysis of results obtained by different methods. The estimated reliability of the results from statistic data from three countries: Ukraine, Russia and Germany. Deviation of simulation results from factual data are minimal and within normal limits.

Keywords: analytical model of mortality, gerontology, aging population dynamics, mortality curve, least-squares method.

Базилевич Ксения Алексеевна – аспирант каф. информатики, Национальный аэрокосмический университет им. Н. Е. Жуковского «Харьковский авиационный институт», Харьков, Украина, e-mail: ksenia.bazilevich@gmail.com.

Мазорчук Мария Сергеевна – канд. техн. наук, доцент, доцент каф. информатики, Национальный аэрокосмический университет им. Н. Е. Жуковского «Харьковский авиационный институт», Харьков, Украина, e-mail: mazorchuk_mary@inbox.ru.

Бакуменко Нина Станиславовна - канд. техн. наук, доцент каф. информатики, Национальный аэрокосмический университет им. Н.Е. Жуковского «Харьковский авиационный институт», Харьков, Украина, e-mail: nina@bigline.net.