

УДК 506:510

Т.В. ШУЛЬКЕВИЧ, І.В. БАКЛАН
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського"

ГІБРИДНИЙ ЛІНГВІСТИЧНИЙ ПІДХІД ДО МОДЕЛЮВАННЯ ЧАСОВИХ РЯДІВ

В статті розглядається використання гібридного лінгвістичного підходу до моделювання та прогнозування часових рядів.

Гібридний лінгвістичний підхід передбачає процес побудови лінгвістичних моделей (ЛМ) з використанням прихованих Марковських моделей (ПММ). Одним з кроків побудови ЛМ є застосування інтервального підходу для розбиття множин значень часового ряду з використанням елементів інтервальної математики та різноманітних розподілів ймовірностей. Будується морфізм перетворення чисельних образів до символного вигляду, який можна використовувати для вирішення певних задач.

Запропоновані методи дають якісні результати на короткостроковому прогнозі, що не відрізняються від прогнозів профільними методами, при використанні менших обчислювальних ресурсів.

Ключові слова: лінгвістизація, лінгвістична модель, інтервалізація, інтервальна математика, розподіли.

Т.В. ШУЛЬКЕВИЧ, И.В. БАКЛАН
Национальный технический университет Украины
"Киевский политехнический институт имени Игоря Сикорского"

ГИБРИДНЫЙ ЛИНГВИСТИЧЕСКИЙ ПОДХОД К МОДЕЛИРОВАНИЮ ВРЕМЕННЫХ РЯДОВ

В статье рассматривается использование гибридного лингвистического подхода к моделированию и прогнозированию временных рядов.

Гибридный лингвистический подход предполагает процесс построения лингвистических моделей (ЛМ) с использованием скрытых Марковских моделей (СММ). Одним из шагов построения ЛМ является применение интервального подхода для разбиения множеств значений временного ряда с использованием элементов интервальной математики и различных распределений вероятностей. Строится морфизм преобразования численных образов в символный вид, который можно использовать для решения определенных задач.

Предложенные методы дают качественные результаты на краткосрочном прогнозе, которые не отличаются от прогнозов профильными методами, при использовании меньших вычислительных ресурсов.

Ключевые слова: лінгвістизація, лінгвістическа модель, інтервалізація, інтервальна математика, розподілення.

T.V. SHULKEVYCH, I.V. BAKLAN
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

HYBRID LINGUISTIC APPROACH TO MODELING OF TIME RANGES

The article deals with the use of a hybrid linguistic approach to analysis, modeling and forecasting of time series.

For the implementation of the hybrid linguistic approach mathematical formalisms were used: interval mathematics, probability theory, structural (syntactic, linguistic) approach, hidden Markov models (HMM) and formal probability grammars.

The hybrid linguistic approach involves the process of constructing linguistic models (LMs) using hidden Markov models (HMM). One of the steps in constructing LM is the use of an interval approach for splitting sets of time series values using elements of interval mathematics and various probability distributions. The morphism of converting numerical images to a symbolic form, which can be used to solve certain problems, is under construction.

The proposed methods provide qualitative results in the short-term forecast, which do not differ from the forecasts by profile methods when using less computational resources.

During the conduct of numerical experiments, the quality of forecasting of time series of diverse nature at various parameters was proved. Experimental way to find optimal parameters of the algorithm. The algorithm was applied to a variety of time series (social, medical, financial, and economic), calculated static indicators of accuracy of the forecast. Experiments have shown that the algorithm consistently performs the forecast of values in a time series of 3-4 steps ahead and forecasts the trend change by 3-5 steps.

The information system with the use of the original mathematical apparatus for analyzing and predicting nonlinear non-stationary processes is developed. Numerical experiments on prediction of time series on different types of intervals (probability distributions).

For intervalization, the Beta probability distribution was used; the quality of the time series forecast at various alphabet capacities (number of intervals) was investigated.

The tables show the results of the conducted numerical experiments.

To estimate the quality of forecasts, the indicator was used based on the Levenstein distance.

Keywords: linguisticization, linguistic model, intervalization, interval mathematics, distributions.

Постановка проблеми

Лінгвістичний (структурний, синтаксичний) підхід відомий ще за роботами американського науковця китайського походження Кінг Сунь Фу [1, 2]. Саме він висловив можливість використання ймовірнісних формальних граматики для опису складних образів. Стоїть задача використання гібридного лінгвістичного підходу для моделювання числових образів у вигляді часових рядів з використанням ймовірнісних граматики на основі прихованих часових рядів.

Аналіз останніх досліджень і публікацій

Елементи лінгвістичного підходу були викладені в роботах Соколова О.Ю. [3], Сенкевича Ю.І. [4], Дрождина В.В. [5], Сулимовой В.В. [6], Oscar Cordon та Francisco Herrera [7], Gui-Wu Wei [8]. Гібридний лінгвістичний підхід запропонований у роботах Баклана І.В. [9–11].

Мета дослідження

Метою чинного дослідження є емпіричне доведення придатності використання гібридного лінгвістичного підходу для прогнозування часових рядів.

Викладення основного матеріалу дослідження

Лінгвістичне моделювання – комплекс методів, методик та алгоритмів, які використовують процес перетворення числових масивів інформації до лінгвістичних послідовностей, на основі яких відновлюється формальна граматики.

Лінгвістична модель динамічного процесу складається з наступних елементів
 $\langle D, I, L, G \rangle$

де D – сукупність часових рядів динамічного процесу та рядів, похідних від вхідних даних,

I – спосіб та правила інтервалізації,

L – морфізм відображення інтервального представлення ряду на певний алфавіт,

G – відновлена граматики динамічного процесу.

Кроки (підзадачі) побудови лінгвістичної моделі на основі чисельного часового ряду динамічного процесу:

- підзадача отримання різницевого ряду;
- підзадача інтервалізації;
- підзадача лінгвістизації;
- підзадача відновлення формальної граматики.

На першому кроці отримуємо різниці часового ряду:

Часовий ряд $X = (x_i)_{(i=1...N)}$;

Перша різниця $X^1 = \{x_{(i+1)} - x_i\}_{(i=1...N)} = \{x_i\}_{(i=1...N)}$;

Друга різниця $X^2 = \{(x_{i+1} - x_i)\}_{(i=1...N)} = \{x^2_i\}_{(i=1...N)}$.

На рис. 1 та 2 представлені оригінальний часовий ряд, який взято з книги авторів Vox G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M. [12, С. 305-306] та інтернет-джерела [13], та його перша різниця.

Для прогнозування використовувалися такі розподіли:

- Рівномірний;
- Лапласа;
- Нормальний;
- Бета;
- Ст'юдента.

Використовувалися алфавіти різних розмірностей: 5, 10, 15, 20, 25 символів.

Нехай X та Y – частково впорядковані множини. Кожну з цих множин вважатимемо

умовно повними структурами $S(X), S(Y)$. Якщо $a, b \in S(X)$ та $a \leq b$, то множину $I(a, b) = [a, b] = \{x \in X, a \leq x \leq b\}$ будемо називати інтервалом на $S(X)$.

З основ інтервальної арифметики маємо: $X \subseteq Is(x) \subseteq B(X), Y \subseteq Is(y) \subseteq B(Y)$, де $Is(y), Is(x)$ – множини всіх інтервалів. Якщо множина X утворює поле, то в $Is()$ можна ввести правила інтервальної арифметики:

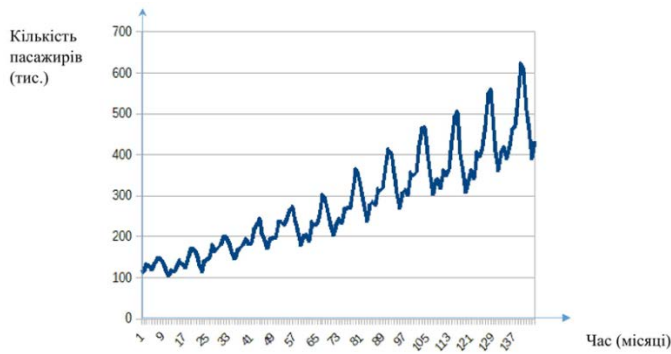


Рис. 1. Графік оригінального часового ряду

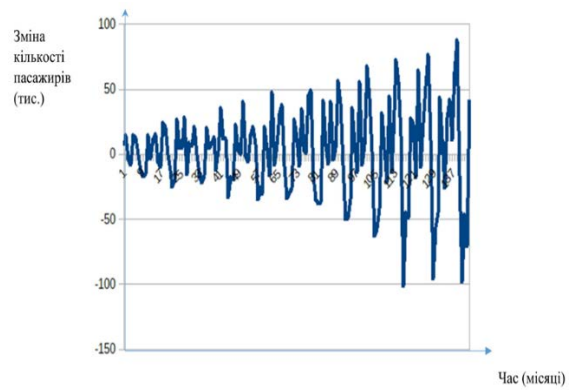


Рис. 2. Графік першої різниці оригінального часового ряду

$$[a, b] + [c, d] = [a + c, b + d],$$

$$[a, b] - [c, d] = [a - c, b - d],$$

$$[a, b] \times [c, d] = [\min(a \times c, a \times d, b \times c, b \times d), \max(a \times c, a \times d, b \times c, b \times d)],$$

$$[a, b]: [c, d] = [a, b] \times [1/d, 1/c], 0 \notin [c, d].$$

Суміжними інтервалами є інтервали вигляду: $[a, b], [c, d] \subseteq X, b=c, a < b, c < d$.

Нехай $X = \{x_i | x_i > 0, x_i < x_m\}, i = 1..M, x_1 \leq x_2 \leq \dots \leq x_m$.

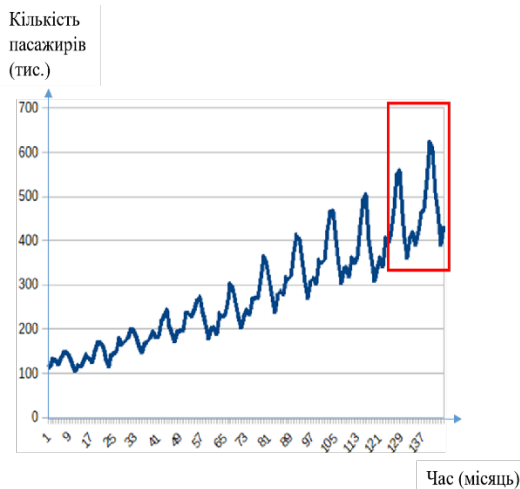


Рис. 3. Оригінальний часовий ряд

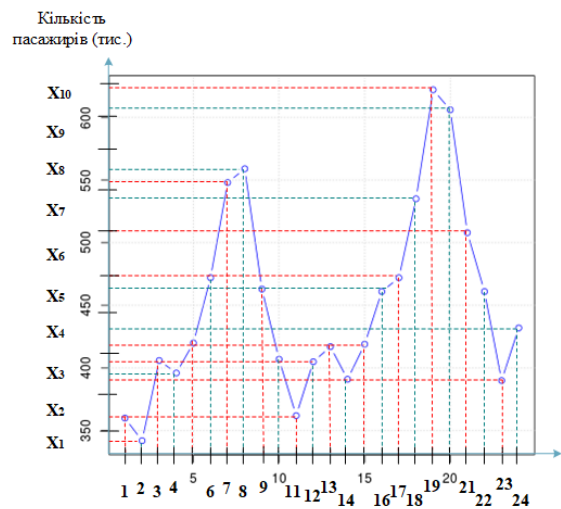


Рис. 4. Інтервалізація

При рівнозначній інтервалізації: $[0, X_1] = [X_1, X_2] = [X_2, X_3] = [X_3, X_4]$,

при рівномірній інтервалізації: $P[0, X_1] = P[X_1, X_2] = P[X_2, X_3] = P[X_3, X_4]$.

На кроці лінгвістизації будемо морфізм:

$$L: X \rightarrow A,$$

де X – множина значень часовий ряду;

A – алфавіт обраної потужності;

$A = \{a_1, \dots, a_M\}$, де M – потужність алфавіту;

$A \xrightarrow{0,13} F$	$C \xrightarrow{0,17} E$	$D \xrightarrow{0,14} A$	$F \xrightarrow{0,04} I$	$G \xrightarrow{0,13} G$	$I \xrightarrow{0,09} G$
$A \xrightarrow{0,25} H$	$C \xrightarrow{0,17} G$	$D \xrightarrow{0,24} E$	$F \xrightarrow{0,04} A$	$G \xrightarrow{0,19} F$	$I \xrightarrow{0,09} F$
$A \xrightarrow{0,25} G$	$C \xrightarrow{0,17} H$	$E \xrightarrow{0,05} H$	$F \xrightarrow{0,09} B$	$G \xrightarrow{0,19} A$	$I \xrightarrow{0,09} H$
$A \xrightarrow{0,38} I$	$C \xrightarrow{0,17} D$	$E \xrightarrow{0,05} I$	$F \xrightarrow{0,09} G$	$G \xrightarrow{0,25} E$	$I \xrightarrow{0,09} E$
$B \xrightarrow{0,08} I$	$D \xrightarrow{0,05} F$	$E \xrightarrow{0,05} A$	$F \xrightarrow{0,13} D$	$H \xrightarrow{0,07} I$	$I \xrightarrow{0,09} J$
$B \xrightarrow{0,08} H$	$D \xrightarrow{0,05} G$	$E \xrightarrow{0,09} F$	$F \xrightarrow{0,13} H$	$H \xrightarrow{0,07} J$	$I \xrightarrow{0,09} I$
$B \xrightarrow{0,17} D$	$D \xrightarrow{0,05} C$	$E \xrightarrow{0,09} B$	$F \xrightarrow{0,22} F$	$H \xrightarrow{0,07} A$	$I \xrightarrow{0,09} B$
$B \xrightarrow{0,25} G$	$D \xrightarrow{0,05} F$	$E \xrightarrow{0,14} G$	$F \xrightarrow{0,22} E$	$H \xrightarrow{0,07} H$	$I \xrightarrow{0,18} D$
$B \xrightarrow{0,42} F$	$D \xrightarrow{0,1} H$	$E \xrightarrow{0,14} E$	$G \xrightarrow{0,06} D$	$H \xrightarrow{0,14} F$	$I \xrightarrow{0,18} C$
$C \xrightarrow{0,08} B$	$D \xrightarrow{0,1} D$	$E \xrightarrow{0,18} D$	$G \xrightarrow{0,06} B$	$H \xrightarrow{0,14} E$	$J \xrightarrow{0,5} B$
$C \xrightarrow{0,08} C$	$D \xrightarrow{0,1} B$	$E \xrightarrow{0,23} C$	$G \xrightarrow{0,06} C$	$H \xrightarrow{0,14} B$	$J \xrightarrow{0,5} C$
$C \xrightarrow{0,17} F$	$D \xrightarrow{0,14} I$	$F \xrightarrow{0,04} C$	$G \xrightarrow{0,06} H$	$H \xrightarrow{0,29} D$	

Авторами статті розроблено інформаційну систему з використанням оригінального математичного апарату для аналізу та прогнозування нелінійних нестационарних процесів.

Рис. 5. Фрагмент программы "Интервализация"

Рис. 6. Фрагмент программы "Построение правил"

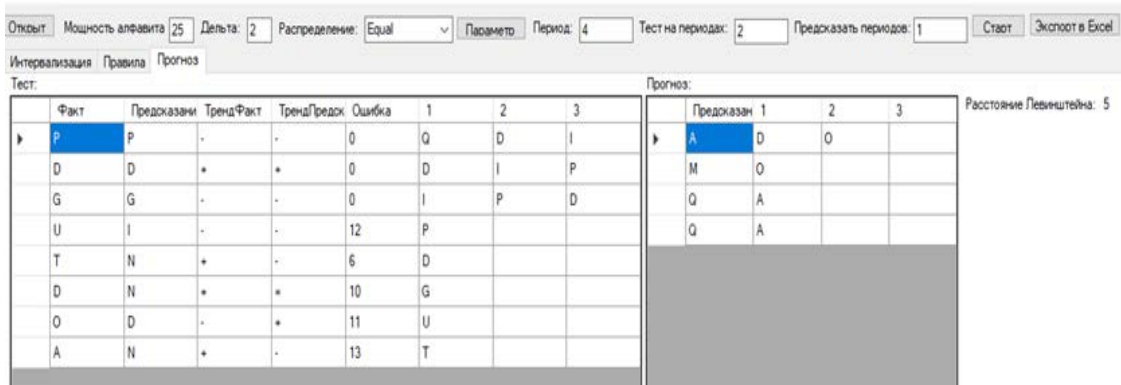


Рис. 7. Фрагмент програми "Прогноз"

Отримуємо лінгвістичну послідовність PDGUTDOA

Обчислювальні експерименти щодо прогнозування часових рядів були проведені на різних типах інтервалізації (розподілах ймовірностей).

Для інтервалізації був використаний Бета розподіл ймовірностей, було досліджено якість прогнозу часового ряду на різних потужностях алфавіту (кількостях інтервалів).

В таблицях наведено результати проведених обчислювальних експериментів на даних з джерела [13].

Для оцінки якості прогнозів був використаний показник на основі відстані Левенштейна.

Нехай M та N – довжини вхідних рядків S_1 та S_2 , тоді для всіх i від 0 до M та для всіх j від 0 до N підраховуємо $D(i, j)$, та повертаємо $D(M, N)$ де:

$$D(i, j) = \begin{cases} 0; i = 0, j = 0 \\ i; j = 0, i > 0 \\ j; i = 0, j > 0 \\ D(i-1, j-1); S_1[i] = S_2[j] \\ \min \begin{pmatrix} D(i, j-1) + insertcount \\ D(i-1, j) + deletcount \\ D(i-1, j-1) + replacecount \end{pmatrix}; j > 0, i > 0, S_1[i] \neq S_2[j] \end{cases}$$

Функція $\min(a, b, c)$ повертає найменший з аргументів. Ціну вставки, видалення та заміни вважатимемо однаковою та рівною одиниці.

Далі в таблицях використовується позначення:

0 – це точне попадання в інтервал при прогнозуванні.

З обчислювальних експериментів, які наведено в табл. 1–4 можемо зробити висновки.

В табл. 1 найкращим розподілом часового ряду перельоту пасажирів є рівномірний розподіл. При використанні бета-розподілу зробити прогноз виявляється неможливим.

Дані з табл. 2 показують дозволяють зробити більш точні прогнози (на великих розмірностях алфавіту), оскільки прогнози робляться в медичній сфері.

Таблиця 1

Помилка тестового прогнозу часового ряду перельоту пасажирів
(за відстанню Левенштейна)

Розподіли	Розмір алфавіту	Прогноз (кроків)				
		1	2	3	4	5
Рівно-мірний	5	0	0	0	1	1
	10	0	0	0	1	1
	15	0	0	0	8	5
	20	0	0	0	10	5
	25	0	0	0	12	6
Лапласа	5	0	0	0	0	0
	10	0	0	0	0	0
	15	1	1	1	1	1
	20	0	0	1	0	0
	25	1	1	0	0	0
Нормальний	5	0	0	0	0	0
	10	0	0	0	0	0
	15	0	1	1	1	1
	20	1	0	0	0	1
	25	0	1	0	0	0
Бета	5	0	0	0	0	0
	10	0	0	0	0	0
	15	0	0	0	0	0
	20	0	0	0	0	0
	25	0	0	0	0	0
Ст'юдента	5	1	1	1	1	1
	10	2	0	1	0	0
	15	0	0	0	1	0
	20	0	1	0	0	1
	25	0	0	0	1	1

Таблиця 2

Помилка тестового прогнозу часового ряду вимірювання рівня глюкози в крові
(за відстанню Левенштейна)

Розподіли	Розмір алфавіту	Прогноз (кроків)				
		1	2	3	4	5
1	2	3	4	5	6	7
Рівно-мірний	5	0	0	1	0	1
	10	0	0	1	1	1
	15	2	3	1	1	0
	20	3	0	2	2	1
	25	3	0	3	3	3

Продовження табл. 2

1	2	3	4	5	6	7
Лапласа	5	0	4	0	4	0
	10	0	7	0	0	7
	15	0	9	7	0	1
	20	1	2	7	1	1
	25	0	0	22	0	0
Нормальний	5	2	0	0	0	3
	10	0	1	0	0	0
	15	0	0	1	2	1
	20	0	0	0	0	1
	25	2	3	0	19	0
Бета	5	0	0	1	2	1
	10	2	4	0	3	0
	15	1	4	0	4	0
	20	4	0	4	2	1
	25	0	2	0	4	1
Ст'юдента	5	4	0	0	0	4
	10	0	0	1	9	0
	15	1	0	0	3	1
	20	14	0	0	16	0
	25	0	1	0	1	23

Таблиця 3

Помилка тестового прогнозу часового ряду ціни Скандинавського ринку електроенергії (за відстанню Левенштейна)

Розподіли	Розмір алфавіту	Прогноз (кроків)				
		1	2	3	4	5
Рівно-ймовірний	5	0	0	0	0	0
	10	2	0	0	0	0
	15	1	0	0	0	0
	20	0	0	0	0	0
	25	1	0	0	0	0
Лапласа	5	0	1	0	0	0
	10	0	8	0	5	0
	15	0	1	2	5	0
	20	0	0	0	0	17
	25	0	7	18	0	22
Нормальний	5	0	0	0	4	0
	10	8	0	0	0	1
	15	0	5	0	2	9
	20	1	19	9	1	0
	25	2	20	0	7	1

Продовження табл. 3

1	2	3	4	5	6	7
Бета	5	0	0	0	0	0
	10	3	0	0	0	2
	15	0	4	3	0	1
	20	3	1	1	8	0
	25	3	3	0	1	4
Ст'юдента	5	0	0	4	0	0
	10	0	0	0	0	0
	15	14	0	1	14	5
	20	1	0	0	0	0
	25	0	0	23	0	1

Таблиця 4

Помилка тестового прогнозу часового ряду ціни природного газу на нью-йоркській товарній біржі NYMEX (за відстанню Левенштейна)

Розподіли	Розмір алфавіту	Прогноз (кроків)				
		1	2	3	4	5
Рівномірний	5	0	0	1	0	1
	10	1	0	8	0	4
	15	1	0	0	2	13
	20	0	16	0	13	18
	25	6	0	0	5	12
Лапласа	5	0	0	0	0	0
	10	9	9	0	9	9
	15	0	14	14	14	14
	20	0	0	19	19	19
	25	0	0	0	1	0
Нормальний	5	0	0	0	4	4
	10	0	0	0	3	3
	15	14	14	5	0	0
	20	0	0	19	0	0
	25	0	0	0	0	0
Бета	5	0	0	0	0	0
	10	1	1	1	1	1
	15	0	0	0	0	1
	20	0	0	0	5	8
	25	0	0	0	0	1
Ст'юдента	5	0	0	0	0	0
	10	0	9	9	9	0
	15	0	0	0	0	14
	20	0	0	19	19	19
	25	0	0	0	24	0

В табл. 4 з короткостроковим прогнозом отримані гарні результати при використанні всіх розподілів. При використанні нормального закону розподілу при

розмірі алфавіту 15 символів виявилось неможливим отримати вірний прогноз на 1, 2, 3 кроки.

Висновки

У статті досліджено використання методів лінгвістичного моделювання для задачі прогнозування часових рядів різної природи. У роботі приведений авторський алгоритм прогнозування на основі лінгвістичного моделювання та прихованих Марковських моделей. Описано використаний математичний апарат для кожного етапу алгоритму, застосування прихованих Марковських моделей, базові характеристики використання апарату формальних граматики для відновлення лінгвістичної моделі.

Для проведення обчислювальних експериментів створена програмна реалізація алгоритму.

Під час проведення обчислювальних експериментів доведена якість прогнозування часових рядів різноманітної природи при різних параметрах. Експериментальним шляхом виявлені оптимальні параметри алгоритму. Алгоритм застосовується до різноманітних часових рядів (соціальних, медичних, фінансових та економічних), враховані статистичні показники точності прогнозу. Експерименти показали, що алгоритм стабільно виконує прогноз значень ряду на 3-4 кроки вперед та прогноз зміни тренду на 3-5 кроків.

Список використаної літератури

1. Фу, К. Структурные методы в распознавании образов / Под ред. М.А. Айзермана. Москва: Мир, 1977. 319 с.
2. Fu, K.S. Structural Methods in Pattern Recognition. New York: Academic Press, 1969. 227 p.
3. Соколов, О. Ю. Алгебраїчні моделі та методи аналізу і синтезу систем керування слабо формалізованими процесами: дис. ... д-ра техн. наук. Харків, 2002. 393 с.
4. Сенкевич, Ю. И. Теоретические основы разработки автоматизированных инструментальных средств для телемедицинских систем полярных зон: дисс. ... д-ра техн. наук. Санкт-Петербург, 2008. 271 с.
5. Дрождин, В. В., Баканов, А. Б. Грамматика описания домена фамилий. *Вопросы радиоэлектроники. Серия: Электронная вычислительная техника*. 2007. Вып. 1. С. 77-82.
6. Моттль, В. В., Сулимова, В. В., Татарчук, А. И. Автоматический выбор наиболее информативных фрагментов в задачах распознавания сигналов разной длительности. *Таврический вестник математики и информатики*. 2006. № 1. С. 109-115.
7. Cordón, Oscar, Herrera, Francisco. A Proposal for Improving the Accuracy of Linguistic Modeling. *IEEE Transactions on Fuzzy Systems*, VOL. 8 (3). (Spaine, Granada, June 3, 2000). Granada: University of Granada, 2000. P. 335-344.
8. Gui-Wu Wei. Uncertain Linguistic Hybrid Geometric Mean Operator and Its Application to Group Decision Making Under Uncertain Linguistic Environment. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. 2009. Vol. 17, № 2. P. 251-267.
9. Баклан, И. В. Лингвистическое моделирование: основы, методы, некоторые прикладные аспекты. *Системные технологии*. 2011. Вып. 3 (74). С. 10-19.
10. Баклан, И. В. Интервальный підхід до побудови лінгвістичної моделі. *Системні технології*. 2013. Вип. 3 (86). С.3-8.
11. Баклан, И. В. Лінгвістичне моделювання часових рядів різної природи з фрактальними властивостями. *Системні технології*. 2016. Вип. 3 (104). С.110-118.

12. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. Time Series Analysis: Forecasting and Control. 5th Edition. New Jersey: Wiley, 2015. 712 p.
13. База статистичних даних. URL: <http://www-eio.upc.edu/~pau/cms/rdata/datasets.html>

References

1. Fu, K. (1977) *Strukturnyie metodyi v raspoznavanii obrazov* (eds. M.A. Ayzermana). Moscow: Mir.
2. Fu, K.S. (1969) *Structural Methods in Pattern Recognition*. New York: Academic Press.
3. Sokolov, O. Yu. (2002) *Algebraichni modeli ta metody analizu i syntezu system keruvannia slabo formalizovanyimi protsesami* (Doctor's thesis), Kharkiv.
4. Senkevich, Yu. I. (2008) *Teoreticheskie osnovy razrabotki avtomatizirovannykh instrumentalnykh sredstv dlya telemeditsynskikh sistem polyarnykh zon* (Doctor's thesis), Sankt-Peterburg.
5. Drozhdin, V. V., & Bakanov, A. B. (2007) *Grammatika opisaniya domena familii. Voprosy radioelektroniki. Seriya: Elektronnaya vychislitel'naya tekhnika*. **1**, 77-82.
6. Mottl, V. V., Sulimova, V. V., & Tatarchuk, A. I. (2006) *Avtomatiicheskiy vyibor naibolee informativnykh fragmentov v zadachah raspoznavaniya signalov raznoy dlitelnosti. Tavricheskiy vestnik matematiki i informatiki*. **1**, 109-115.
7. Cordón, Oscar, & Herrera, Francisco. (2000) *A Proposal for Improving the Accuracy of Linguistic Modeling. IEEE Transactions on Fuzzy Systems*, VOL. 8 (3). (Spane, Granada, June 3, 2000), Granada: University of Granada, pp. 335-344.
8. Gui-Wu Wei. (2011) *Uncertain Linguistic Hybrid Geometric Mean Operator and Its Application to Group Decision Making Under Uncertain Linguistic Environment. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. **17**, 2, 251-267.
9. Baklan, I. V. (2011) *Lingvisticheskoe modelirovanie: osnovy, metodyi, nekotoryie prikladnyie aspektyi. Sistemnyie tekhnologii*. **3** (74), 10-19.
10. Baklan, I. V. (2013) *Intervalnyi pidkhid do pobudovy linhvistychnoi modeli. Sistemni tekhnologii*. **3** (86), 3-8.
11. Baklan, I. V. (2016) *Linhvistychno modeliuвання chasovykh riadiv riznoi pryrody z fraktalnymi vlastyivostyamy. Sistemni tekhnologii*. **3** (104), 110-118.
12. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015) *Time Series Analysis: Forecasting and Control*. 5th Edition. New Jersey: Wiley.
13. Database of statistical data. (2018) Retrieved from <http://www-eio.upc.edu/~pau/cms/rdata/datasets.html>