

ЛІНГВІСТИЧНЕ МОДЕЛЮВАННЯ ЧАСОВИХ РЯДІВ РІЗНОЇ ПРИРОДИ З ФРАКТАЛЬНИМИ ВЛАСТИВОСТЯМИ

Анотація. Стаття присвячена особливостям побудови лінгвістичних моделей часових рядів різної природи, які мають фрактальні властивості. Приведені визначення ймовірнісних граматики, які є частиною лінгвістичних моделей часових рядів, етапи побудови лінгвістичної моделі. Наведені результати чисельних експериментів лінгвістичного моделювання різноманітних фрактальних послідовностей, таких як броунівський рух та класичні фрактали.

Ключові слова: лінгвістичне моделювання, лінгвістична модель, ймовірнісна граMATика, часовий ряд, фрактал.

Постановка проблеми

Широка поширеність фрактальних властивостей часових рядів дозволяє сподіватися на наявність єдиного універсального механізму, який приводить до аналізу фрактальності динамічних процесів, що виникають в цілком різноманітних природних та технічних системах. Пошук такого механізму є одним з найактуальніших завдань сучасної науки.

В статті пропонується у якості такого універсального інструменту побудова лінгвістичних моделей динамічних процесів з фрактальними властивостями.

Аналіз публікацій по темі дослідження

Напрямок лінгвістичного моделювання як спеціалізованого виду математичного моделювання базується з одного боку на роботах з розпізнавання образів Фу К.С.[1], з другого боку теорії формальних граматики Хомського Н.[2] Серед сучасних науковців можна згадати роботи Потапова А.В.[3], Сенкевича Ю.І. [4], Соколова О.Ю. [5], Старченко М.В. [6]. Однак усі наведені роботи не дають універсального механізму для моделювання динамічних процесів з урахуванням фрактальності отримуваних часових рядів. Початок уза-

гальненого підходу був викладений у роботах автора цієї статті [7,8,9].

Мета дослідження

Метою дослідження є пропозиція лінгвістичного моделювання у якості універсального механізму моделювання часових рядів динамічних процесів з наявністю фрактальних властивостей для подальшого аналізу та прогнозу.

Основна частина

Лінгвістична модель — побудована на основі лінгвістичного моделювання сукупність символічних (лінгвістичних) послідовностей за обраними параметрами лінгвістизації та відновлена на її основі 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.

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

Лінгвістичне моделювання повинно забезпечувати:

- обґрунтований вибір інтервалів для виконання задач лінгвістизації (інтервалізація);
- ефективне перетворення числових масивів даних до лінгвістичних ланцюжків;
- підходів вивчення впливу обраних параметрів лінгвістизації на кінцеві результати застосування лінгвістичного моделювання;
- відновлення за лінгвістичними ланцюжками формальних граматики, в тому числі на ймовірнісних граматики та використання апарату прихованих марковських моделей для побудови гібридних моделей на основі лінгвістичного моделювання;
- інтеграцію лінгвістичних моделей з іншими обчислювальними парадигмами та створенні на їх основі гібридних процедур для вирішення різноманітних практичних завдань.

Лінгвістизація - процес перетворення часових рядів до сукупності лінгвістичних послідовностей, на основі яких будується формальна граматики.

Необмежені граматики (або граматики типу 0) – мають у своїй основі правила вигляду $\alpha \rightarrow \beta$, де на ланцюжки $\alpha \in V^+, \beta \in V^*$ не накладається ніяких обмежень, де V^* - усі можливі ланцюжки символів алфавіту V , а $V^+ = V^* \setminus \{\Lambda\}$ [3].

Лінгвістичне моделювання припускає різномантні способи побудови лінгвістичних ланцюжків, але певні можуть використовуватися істотно частіше, ніж інші. У зв'язку з цим кожному правилу можна придати деяку ймовірність, з якою це правило застосовується. Тим самим ми приходимо до поняття ймовірнісних (або стохастичних граматик) [1].

Ймовірнісною граматикою називається формальна трійка $G = \langle A, P, S \rangle$, де A – алфавіт (в загальному випадку він може бути поданий, як об'єднання термінального та не термінального алфавітів), S – початковий символ, а P – множина правил, які мають наступний вигляд $\alpha \xrightarrow{P(\beta|\alpha)} \beta$, де $P(\beta|\alpha)$ – ймовірність заміщення ланцюжка α на ланцюжок β .

У випадку побудови лінгвістичної моделі правила вигляду $\alpha \xrightarrow{P(\beta|\alpha)} \beta$ фактично будуть правилами передування ланцюжка (символа) α перед ланцюжком (символом) β .

Звичайно, повинна виконуватися аксіоматика теорії ймовірності, яка у випадку лінгвістичного моделювання має наступні особливості. Тобто, сума ймовірностей правил передування певного ланцюжка (символа) перед іншими можливими β повинна дорівнювати одиниці:

$$\sum_{\beta} P(\beta|\alpha) = 1.$$

Сукупність правил передування може бути задана матрицею передування наступного вигляду:

	z	...	a	A	...	Z
z						
...						
a						
A						
...						
Z						

де на перетині відповідних символів (ланцюжків) повинні стояти ймовірності передування.

Розглянута вище аксіоматика стосується кожного рядка матриці.

Окремою, важливою областю вивчення фрактальних властивостей є аналіз часових рядів: послідовностей вимірювання різноманітних величин, упорядкованих за часом. Як правило, інформація про поведінку складних систем виходить у вигляді саме таких експериментальних даних. На сьогоднішній день добре відомо, що фрактальні властивості мають графіки реалізацій надзвичайно різних процесів, як стохастичних (наприклад, броунівський рух), так і детермінованих (наприклад, реалізація рішень логістичного рівняння при певних значеннях параметра). Фрактальні часові ряди виникають, зокрема, при вимірах різних природних процесів: сонячної активності, рівня розливів річок, шумів електронних приладів, геофізичної та геомагнітної активності, фізіологічних характеристик організму людини, наприклад, у вигляді кардіограм й таке ін.

Розглянемо приклади застосування лінгвістичного моделювання для побудови моделі динамічних процесів, які можуть мати фрактальні властивості [8].

Сам процес побудови лінгвістичної моделі на основі часового ряду динамічного процесу повинен мати в собі наступні кроки:

1) Підготовчий крок, який має на увазі на основі вхідного ряду отримання різницевого ряду.

2) Сортування різницевого ряду для оцінювання області значень.

3) Інтервалізація – розбиття області значень на інтервали за певною схемою [9].

4) Формування лінгвістичного ланцюжка на основі різницевого чисельного ряду для обраного алфавіту.

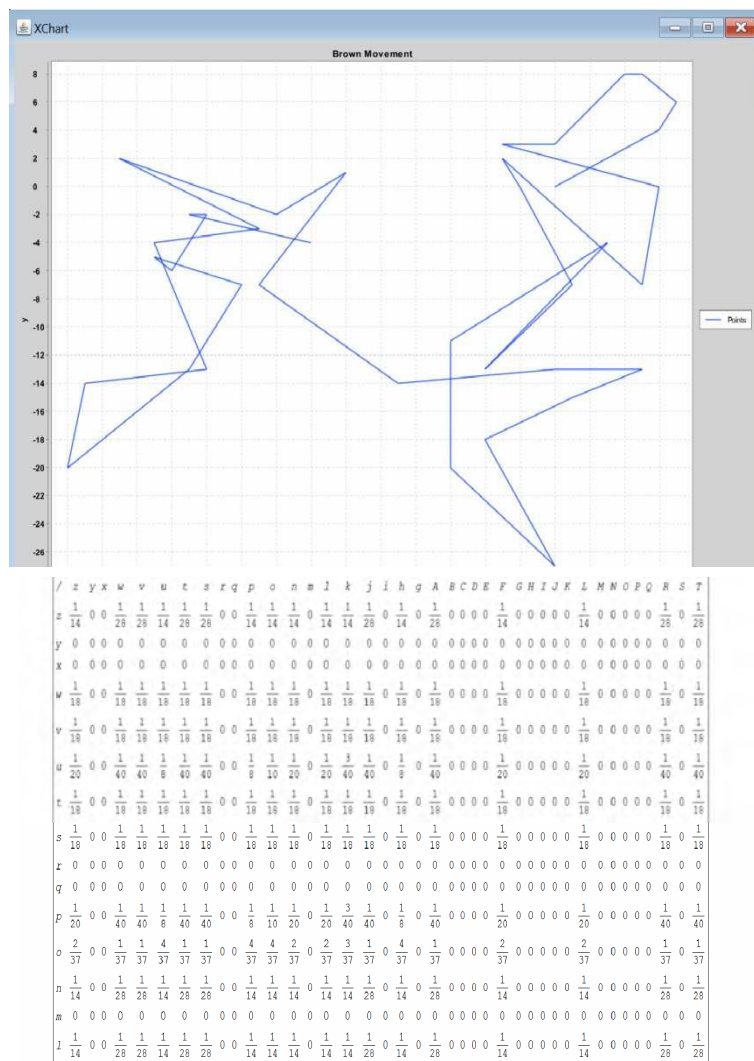
5) Визначення матриці передування.

6) Отримання системи правил передування символів (ланцюжків).

Тим самим ми отримуємо усі компоненти лінгвістичної моделі динамічного процесу.

Чисельні експерименти по лінгвістичному моделюванню

Надалі наведемо отримані результати при лінгвістичному моделюванні броунівського руху в середовищі системи MAXIMA .



За матрицею передування були отримані правила для відновленої ймовірнісної граматики (нижче приведені деякі з них):

```

Rule=[ z -> w,  $\frac{1}{28}$  ]
Rule=[ z -> v,  $\frac{1}{28}$  ]
Rule=[ z -> u,  $\frac{1}{14}$  ]
Rule=[ z -> t,  $\frac{1}{28}$  ]
Rule=[ z -> s,  $\frac{1}{28}$  ]
Rule=[ z -> p,  $\frac{1}{14}$  ]
Rule=[ z -> o,  $\frac{1}{14}$  ]
Rule=[ z -> n,  $\frac{1}{14}$  ]
Rule=[ z -> l,  $\frac{1}{14}$  ]
Rule=[ z -> k,  $\frac{1}{14}$  ]
Rule=[ z -> j,  $\frac{1}{28}$  ]

```

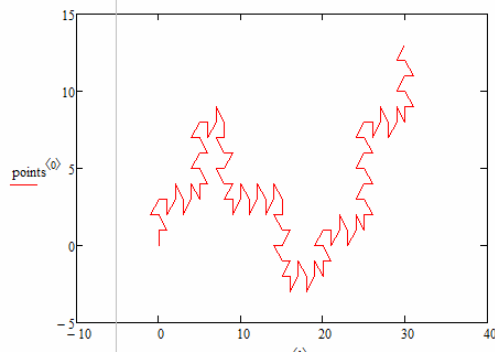
Другий приклад стосується лінгвістичної моделі процесу побудови класичного фракталу «Бур`ян», яка фактично подана у вигляді двох моделей - координат руху x та y . Вхідні дані є «виродженим» випадком для побудови лінгвістичного ряду, оскільки різницеві ряди мають невелику кількість значень, які постійно повторюються (кількість таких значень залежить від зсуву при обчисленні різницевого ряду та дорівнює $2n+1$, де n – розмір зсуву різницевого ряду). Розрахунки були виконані в системі MathCAD.

Points

points := ar(word)

	0	1
0	0	0
1	1	0
2	1	1
3	2	0
4	2	-1
5	3	0
6	3	1
7	2	1
8	3	2
9	4	2
10	3	3
11	3	3
12	2	3
13	3	4
14	4	4
15	3	...

Difference rows:



Обраний в цьому випадку алфавіт – $A=\{i, h, g, f, e, d, c, b, a, A, B, C, D, E, F, G, H, I\}$. При цьому маленькі латинські літери використовуються для відображення від'ємних значень, а великі – для позитивних.

Правила передування ймовірнісної граматики цьому випадку будуть мати наступний вигляд:

Правила руху координати x	Правила руху координати y
a - (0.5) -> a	e - (0.5) -> e
a - (0.25) -> A	e - (0.25) -> a
a - (0.25) -> E	e - (0.17) -> A
A - (0.5) -> a	a - (1.0) -> e
A - (0.25) -> A	A - (0.6) -> e
A - (0.25) -> E	A - (0.4) -> E
E - (0.11) -> A	E - (0.27) -> A
E - (0.83) -> E	E - (0.64) -> E
	E - (0.09) -> I
	I - (1.0) -> E

На тому самому алфавіті А для послідовності дій при побудові фракталів «Острова та озера» та «Острів Коха» отримуємо для координати x наступну сукупність правил передування для ймовірнісної граматики лінгвістичної моделі процесу:

Фрактальна послідовність «Острова та озера»	Фрактальна послідовність «Острів Коха»
i - (0.53) -> g	i - (0.38) -> i
i - (0.47) -> C	i - (0.25) -> f
g - (0.08) -> i	i - (0.13) -> A
g - (0.42) -> g	i - (0.13) -> F
g - (0.14) -> e	g - (0.74) -> c
g - (0.08) -> c	g - (0.26) -> A
g - (0.09) -> A	f - (0.29) -> i
g - (0.18) -> C	f - (0.71) -> c
e - (0.81) -> g	c - (0.03) -> i
e - (0.19) -> e	c - (0.32) -> g
c - (0.28) -> g	c - (0.08) -> f
c - (0.17) -> e	c - (0.45) -> c
c - (0.22) -> c	c - (0.03) -> a
c - (0.33) -> a	c - (0.08) -> A
a - (1.0) -> c	a - (1.0) -> c
A - (0.48) -> g	A - (0.13) -> g
A - (0.52) -> C	A - (0.13) -> c
C - (0.07) -> i	A - (0.46) -> A
C - (0.19) -> g	A - (0.17) -> D
C - (0.1) -> A	A - (0.1) -> G
C - (0.42) -> C	C - (1.0) -> F
C - (0.14) -> E	D - (0.27) -> A
C - (0.08) -> G	D - (0.73) -> G
E - (0.81) -> C	F - (0.33) -> i
E - (0.19) -> E	F - (0.33) -> D
G - (0.28) -> C	F - (0.33) -> G
G - (0.17) -> E	G - (0.11) -> A
G - (0.24) -> G	G - (0.02) -> C
G - (0.32) -> I	G - (0.41) -> D
I - (1.0) -> G	G - (0.02) -> F
	G - (0.45) -> G

Легко бачити, що отримані сукупності для двох фрактальних послідовностей на одному й тому ж алфавіті та способі інтервалізації відрізняються складом правил.

Взагалі були проведені чисельні експерименти по лінгвістичному моделюванню побудови цілої низки фрактальних послідовностей

ЛІТЕРАТУРА

1. Фу К.С. - Структурные методы в распознавании образов. – М.:Мир, 1977.
2. Хомский Н., Миллер Дж. Введение в формальный анализ естественных языков // Кибернетический сборник / Под ред. А.А.Ляпунова и О.Б.Лупанова. — М.: Мир, 1965.
3. Потапов А.С. Распознавание образов и машинное восприятие: Общий поход на основе принципа минимальной длины описания. – СПб.: Политехника, 2007. – 548 с.
4. Сенкевич Ю.И. Алгоритмы лингвистического анализа динамических систем // Вопросы радиоэлектроники. Серия: Электронно-вычислительная техника. – 2007. – Вып.4. – С.88-96.
5. Соколов А.Ю. Синтез оптимального управления лингвистическими динамическими системами // Праці п'ятої української конференції з автоматичного управління (АВТОМАТИКА 98). – Ч. 1. - Київ: КПІ. – 1998. – С. 356-363.
6. Старченко Н.В. Локальный фрактальный анализ в физических приложениях. // Препринт № 006-2005 М.: МИФИ, 2005.
7. Баклан І. В. Лінгвістичне моделювання: основи, методи, деякі прикладні аспекти / І. В. Баклан // Систем. технології. — 2011. — № 3. — С. 10-19.
8. Баклан І.В. Фрактальні властивості при лінгвістичному моделюванні динамічних процесів // Вестник Херсонского национального техн. ун-та. - Херсон: ХНТУ, 2015. - Вып. 3 (54).– С.210-213.
9. Баклан І.В. Інтервальний підхід до побудови лінгвістичної моделі // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 3(86). – Дніпропетровськ, 2013. – С.3-8.