

УДК 37.015:007.002.645:519.722

Лисогор В.М.

(Вінницький національний аграрний університет)

**КОНЦЕПЦІЇ ПОБУДОВИ МОДЕЛЕЙ НАВЧАННЯ ІНФОРМАЦІЙНИХ
ТЕХНІЧНИХ СИСТЕМ АГРАРНОГО СЕКТОРУ**

Выполнен анализ источников энтропийных и информационных моделей систем обучения. Предложены концепции построения информационно-энтропийных систем, разработаны новые подходы исследования их в области образования на примере повышения эффективности функционирования аграрного высшего учебного заведения (АВУЗа)

Analyzed sources entropijnyh and informacijnyh models Proposed the concept of learning systems: information systems, developed entropijnyh new approaches to their research in the field of pedagogy: the case of cyber-improve the functioning of higher education institution (HEI and).

Обґрунтування напрямків дослідження

Проблема аналізу невизначеності поведінки об'єктів у різних галузях досліджень: техніки, економіки, організаційних систем, стоять достатньо гостро. Встає нескінченна кількість задач, питань, рішення яких можуть покращити ефективність функціонування системи, забезпечити умови мінімізації функцій вигоди. Подібна проблема також постає перед новою галуззю аграрного сектору – моделей навчання. Витоки з побудови моделей інформаційних систем можна вбачати у публікації [2] К.Шеннона, де автор розкрив основні підходи, виникнення, інтерпретації, формування основних теоретичних і практичних положень цієї науки. Вхідження в нову область математики – теорію інформації, що тісно пов'язана з кібернетикою, і яка має ряд використань в техніко-економічному управлінні, техніці зв'язку, лінгвістиці, біології, були викладені в [3]. Довгий час існувала точка зору, що ентропія та інформація представляє не самостійну наукову галузь, а є розділом теорії ймовірностей і математичної статистики [4,7,8,9]. Достатньо фундаментальні результати по теорії інформації отримані науковою школою Кузьміна І.В. [5,6], який як ректор політехнічного інституту використовував свої нароби в навчальному процесі. Значну увагу проблемі підвищення ефективності управління Вінницьким національним технічним університетом в рамках своєї наукової школи приділив і нинішній ректор Мокін Б.І. [10,11].

Незважаючи на наявні успіхи отриманих результатів у різних галузях знань, проблема дослідження аналізу моделей ентропії та інформації в педагогічній кібернетичній знаходиться на початковому етапі розвитку і є достатньо актуальною проблемою, яка чекає свого повного чи часткового вирішення.

Мета

Виконати аналіз джерел витоків, запропонувати концепції побудови інформаційно-ентропійних систем, з розробкою нових підходів дослідження у галузі кібернетичної педагогіки.

Основний результат

З самого початку зародження теорії інформації [2] вона будувалась на основі ймовірного числення з використанням логарифмічних методів досліджень, для яких показник функції невизначеності при $K=1$ повинен дорівнювати нулю. далі, при зростанні невизначеності K – функція ризику повинна зростати. Автори цієї публікації для прозорості аналізу запропонували особливу увагу приділити до значень $K = 2; 3; \dots; 10$; $K = \pi = 3,1416$, $K = e = 2,7183$. Використавши вказані цифрові дані змогли дослідити, як окремі величини, так і зв'язок між ними. При використанні $K = 2$ автори запропонували свою методику дослідження і звели аналіз до отримання відомого результату – двійкової одиниці невизначеності – біта. Взавши $K = 10$ з припущенням рівномірного ймовірнісного розподілу, ми отримали цікаві результати для аналізу з основою логарифмів двійкової, десяткової, натуральної невизначеності системи. Отже, двійкова одиниця (цифра – герб будь-

якої металічної грошової одиниці, два шари в урні чорного та білого кольорів) впливала назва двійкової одиниці невизначеності – біта. Десяткова одиниця ($K = 10$) невизначеності більше невизначеності двійкової на $\log_2 K = \log_2 10 \approx 3,32 \approx 3\frac{1}{3}$ рази. Коли K будь-яке число, то залежність між наслідками спроб та їх ймовірностями можна було представити так (табл.1, табл.2)

Таблиця 1

Залежність між подіями і рівномірними значеннями K

Наслідки спроб	A_1	A_2	A_k
Ймовірності $1/K$	$1/K_1$	$1/K_2$	$1/K_k$

Таблиця 2

Залежності між подіями та їх ймовірностями у загальному виді

Наслідки спроб	A_1	A_2	A_k
Ймовірності	$P(A_1)$	$P(A_2)$	$P(A_k)$

Якщо врахувати, що міра невизначеності повинна бути від'ємною величиною, тоді:

$$\frac{I}{K} \log K = -\frac{I}{K} \log \frac{I}{K}, \quad (1)$$

Звернувши увагу появи знаку мінусу перед рівнянням (1) ми отримали результат, який є основою невизначеності будь-якої інформаційної системи. Означимо міру невизначеності і довільної інформаційної системи через $H(\alpha)$, а її найкраще значення через $H(\alpha_0)$. Використавши значення табл.1, табл.2 отримаємо аналітичний вид функцій невизначеності $H(\alpha)$, $H(\alpha_0)$ у вигляді:

$$H(\alpha) = -P(A_1) - P(A_2) \log P(A_2) - \dots - P(A_k) \log P(A_k) \leq \log K;$$

$$-\frac{I}{K} \log \frac{I}{K} - \frac{I}{K} \log \frac{I}{K} \dots - \frac{I}{K} \log \frac{I}{K} = H(\alpha_0); \quad (2)$$

причому рівність досягається тільки у тому випадку коли:

$$P(A_1) = P(A_2) = \dots = P(A_k) = \frac{1}{K}, \quad (3)$$

Поки що не будемо займатися доведенням твердження (2), а продемонструємо пояснення на прикладі, двійкової одиниці - біта

$$-P(A_1) \log P(A_1) - P(A_2) \log P(A_2) \leq \log 2 \quad (4)$$

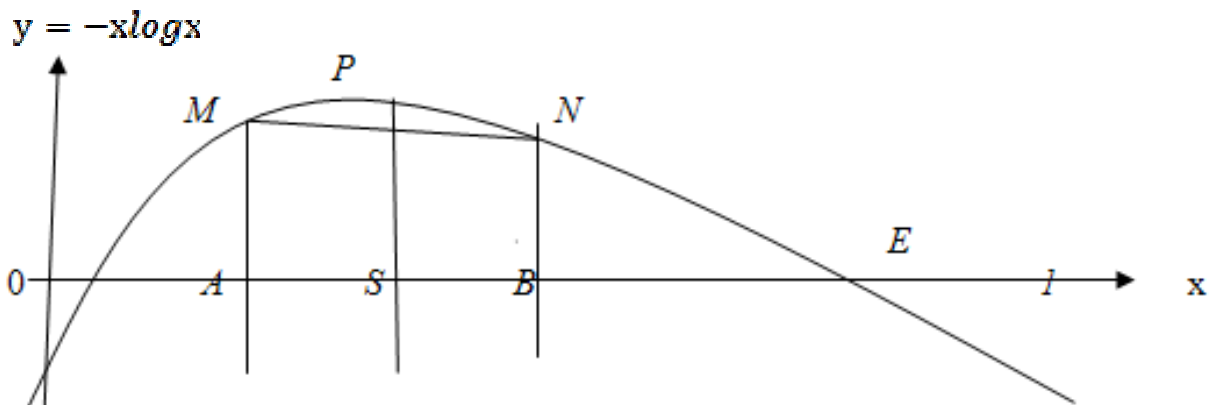


Рис.1 - Обґрунтування невизначеності бінарної системи.

Аналізуючи (рис.1) бачимо, що значення функції $y = -x \log x$ при $x \rightarrow 0$ збігається до нуля. З іншої сторони, при $x = 1$ значення її також дорівнює нулю, а при $0 \leq x \leq 1$ ця функція позитивна, бо у цьому випадку $\log x$ від'ємний (рис.1) при $x > 1$ функція $-x \log x$ від'ємна. Графік функції, що розглядається на рис.1, де $OE = 1, OA = P(A_1), OB = P(A_2)$, відрізок BN зображає $-P(A_2) \log P(A_2)$. По причині того, що $OA + OB = P(A_1) + P(A_2) = 1 = OE$, бо, з одної сторони вони утворюють повну групу подій, а з іншої сторони OE безпосередньо дорівнює одиниці.

Відстань OS від початку до середини S відрізка AB дорівнює $1/2$, а тому відрізок SP дорівнює $-1/2 \log 1/2 = \frac{1}{2} \log 2$. Але напівсума відрізків AM та BN дорівнює середній лінії трапеції $ABNM$, яка не перевищує SP , а тому

$$\frac{1}{2}[-P(A_1) \log(A_1) - P(A_2) \log P(A_2)] \leq \frac{1}{2} \log 2, \\ -P(A_1) \log(A_1) - P(A_2) \log P(A_2) \leq \log 2 \quad (5)$$

де в (5) рівність може бути у тому випадку, коли $OA = OB = OS$. Ми довели, що функція невизначеності:

$$H(P) = -P \log P - (1-P) \log(1-P) \quad (6)$$

Визначають ентропію спроби з двома наслідками P та $1-P$ приймає найбільше значення $\log 2 \approx 0,30$ при $P = 1/2$.

Розглянувши конкретний приклад, використавши методи математичної індукції, авторам публікації вдалось вийти на загальнотеоретичні положення теорії ентропії та інформації. У зв'язку з цим можна зробити декілька пунктів проміжних висновків: якісні міркування відносно табл.1, табл.2 та рис.1, дали можливість отримати кількісну оцінку ентропії, яка стверджує: інформація також має кількісну оцінку;

-залежність між кількістю інформації і кількістю комбінацій, що складені з алфавіту табл.1, табл.2, рис.1 – логоріфмічна;

-загально прийнята одиниця кількості інформації – двійкова, вона отримана при виборі одного з двох рівномірних символів (цифра – герб, чорний – білий колір шарів) і дорівнює 1 біт;

-кількість інформації залежить від усуненої невизначеності – ентропії про невідомий до випробування факт, що передається;

-ентропія представляє собою математичне сподівання величини $-\log p_i$, а тому невизначеність появи кодового слова на виході джерела повідомлень дорівнює ентропії цього джерела.

У першій частині публікації ми отримали декілька результатів, про що відмічено у проміжному висновку, але система кібернетичної педагогіки представляє собою об'єкт зі значно складнішою структурою і параметрами контролю та управління. Ускладнимо деякі позиції об'єкта аналізу. Вичерпною характеристикою координат об'єкта дослідження в кожний заданий момент часу є розподіл ймовірностей цих величин.

Для безперервних координат таких, як вищий навчальний заклад (ВНЗ), що приймають довільні значення у деяких інтервалах, пропонуємо розглянути щільність ймовірностей $p(x_1, x_2, \dots, x_n)$ у якої функція $p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$ виражає ймовірність одночасного знаходження координат в інтегралах $x_i \leq x_i \leq x_i + dx_i; i = 1, 2, \dots, n$.

Для аналізу теорії ймовірностей і теорії інформації будемо використовувати два поняття ентропії: внутрішню ентропію, зовнішню або інформаційну ентропію з однаковою формулою запису

$$H = - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1, x_2, \cdots x_n) \log p(x_1, x_2, \cdots x_n) dx_1, \cdots dx_n \quad (7)$$

При необхідності будемо робити відповідні пояснення, інколи будемо пояснювати ентропію так $H = H(x_1, x_2, \cdots x_n)$.

Визначимо кількість інформації, що циркує у ВНЗ. Інформацією називаємо будь-які повідомлення, первинним джерелом яких є спроба та спостереження.

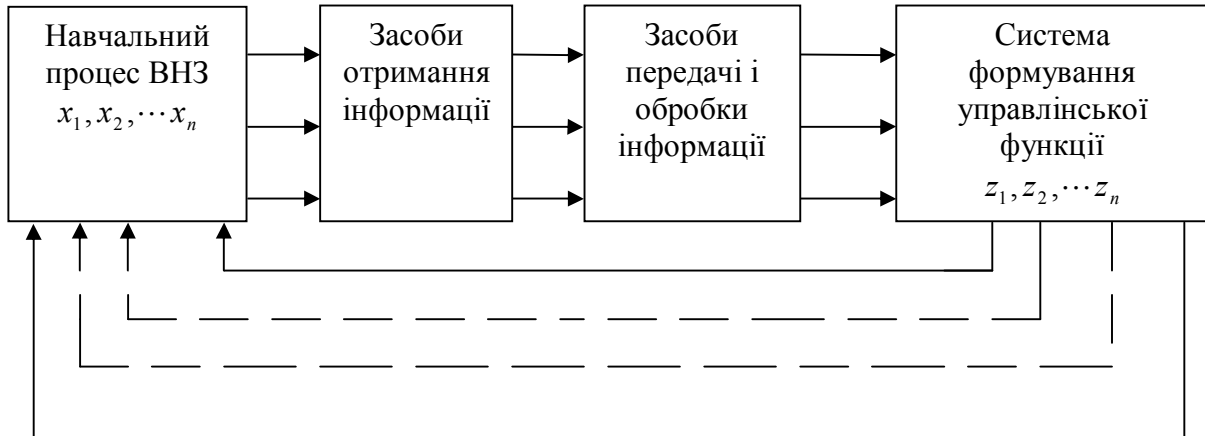


Рис. 2 - ВНЗ, засоби отримання, передачі, обробки, управління навчальним процесом

Середня кількість інформації блоку засобів передачі і обробки інформації у скалярному випадку ($x \rightarrow y$) виразимо такою моделлю:

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \log \frac{p_y(x)}{p(x)} dx dy, \quad (8)$$

Тут, $p(x, y)$ взаємна щільність засобів передачі і обробки інформації, величина $p(x, y) dx dy$ виражає ймовірність того, що величина переданої інформації знаходиться в інтервалі $x - (x + dx)$, а прийнята інформація знаходиться в інтервалі $y - (y + dy)$.

Аналогічно для векторного випадку передачі n величин ($x_1, x_2, \cdots x_n \rightarrow y_1, y_2, \cdots y_n$) математичне сподівання кількості інформації на одне повідомлення дорівнюватиме:

$$I = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, x_2, \cdots x_n, y_1, y_2, \cdots y_n) \log \frac{p_y(x_1 \cdots x_n)}{p(x_1, \cdots x_n)} dx_1, \cdots dx_n, \quad (9)$$

Тут, $p(x_1, x_2, \cdots x_n, y_1, y_2, \cdots y_n) dx_1 \cdots dx_n dy_1 \cdots dy_n$ виражає ймовірність одночасного знаходження переданих величин в інтервалах

$$x_i - (x_i + dx_i), i = 1, 2, \cdots, n \quad (10)$$

$$y_i - (y_i + dy_i), i = 1, 2, \cdots, n \quad (11)$$

Моделі (8) та (9) як наголошувалось вище, можна представити у декількох еквівалентних видах. Для скалярної невизначеності

$$\int_{-\infty}^{\infty} p(x, y) dy = p(x) \quad (12)$$

Для векторної величини

$$p(x, y) = p_y(x)p(y) = p(x)p_x(y) \quad (13)$$

Замінюючи у(8) $p_y(x)$ на $p(x, y)/p(y)$, отримаємо симетричну формулу для середньої кількості інформації при передачі повідомлення ($x \rightarrow y$)

$$I = \int_{-\infty}^{\infty} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (14)$$

Після низки проміжних перетворень ми можемо звести результат отримання середньої кількості інформації при передачі і обробці ($x \rightarrow y$) до виду:

$$I = H(x) - M_y[H_y(x)] \quad (15)$$

Тобто середня кількість інформації при передачі і обробці ($x \rightarrow y$) дорівнює (15) різниці ентропії розподілення ймовірностей вхідної величини X до отримання сигналу та усередненій ентропії розподілення ймовірностей усієї величини після отримання сигналу. Величина $M_y[H_y(x)]$ в (15) є усередненою характеристикою ефекту дії завад(шумів), а також запізнення при передачі сигналів як кібернетичної моделі ВНЗ. Уважно аналізуючи знаменник (14) бачимо, що найкращий результат ми отримаємо при незалежних значеннях $p(x)$ і $p(y)$, як це спостерігалось у першій частині публікації.

Останній результат зводиться до аналізу $p(x, y)$ важливої компоненти моделі (14). Тут бачимо, що кількість інформації, що передається падає при збільшенні завад (шумів) та запізньень в передачі сигналів. Проілюструємо нашу гіпотезу на прикладі нормального закону розподілення ймовірностей вхідної і вихідної величин (14), отримаємо

$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2}\right] - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right\} \quad (16)$$

Тут m_x, m_y - математичні сподівання величин (x, y); r - коефіцієнт кореляції цих величин, що дорівнює:

$$r = \frac{1}{\sigma_x\sigma_y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-m_x)(y-m_y)p(x, y) dx dy \quad (17)$$

Дисперсії величин (x, y)

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x-m_x)^2 p(x, y) dx dy = \int_{-\infty}^{\infty} (x-m_x)^2 p(x) dx \quad (18)$$

$$\sigma_y^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y-m_y)^2 p(x, y) dx dy = \int_{-\infty}^{\infty} (y-m_y)^2 p(y) dy \quad (19)$$

На кінець щільності нормального розподілення кожної з величин (x, y) дорівнюють:

$$p(x) = 1/\sigma_x\sqrt{2\pi} \exp\left[-\frac{(x-m_x)^2}{2\sigma_x^2}\right] \quad (20)$$

$$p(y) = 1/\sigma_y\sqrt{2\pi} \exp\left[-\frac{(y-m_y)^2}{2\sigma_y^2}\right]$$

Таким чином відношення $p(x, y)/p(x)p(y)$ дорівнюватиме

$$\frac{p(x,y)}{p(x)p(y)} = \frac{1}{\sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right] + \frac{(x-m_x)^2}{2\sigma_x^2} + \frac{(y-m_y)^2}{2\sigma_y^2} \right\} = \frac{1}{\sqrt{1-r^2}} \exp \left\{ -\frac{r^2}{2(1-r^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2(x-m_x)(y-m_y)}{r\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right] \right\} \quad (21)$$

Прологарифмувавши відношення $p(x,y)/p(x)p(y)$ отримаємо:

$$\log \frac{p(x,y)}{p(x)p(y)} = \log \frac{1}{\sqrt{1-r^2}} - \frac{r^2 \log e}{2(1-r^2)} \left[\frac{(x-m_x)^2}{\sigma_x^2} - \frac{2(x-m_x)(y-m_y)}{r\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right] \quad (22)$$

Матричне сподівання виразу в квадратних дужках дорівнює нулю.

Підставивши (22) в (14) отримаємо кількість інформації навчального процесу ВНЗ :

$$I = \log \frac{1}{\sqrt{1-r^2}} = -\frac{1}{2} \log(1-r^2) \quad (23)$$

Таким чином, для самого розповсюдженого нормального розподілення ймовірностей вхідних і вихідних величин ВНЗ середня кількість інформації у блоці засобів передачі і обробки даних (рис.2) $X \rightarrow Y$ достатньо просто виражається через коефіцієнт кореляції \mathcal{C} вхідних і вихідних величин (x,y) . У моделі кількості інформації (23) величина \mathcal{C} означає коефіцієнт кореляції, який дорівнює:

$$r = \frac{\sigma_x^2}{\sigma_x \sqrt{\sigma_x^2 + \sigma_u^2}} = \frac{1}{\sqrt{1 + \sigma_u^2 / \sigma_x^2}}, \quad (24)$$

де індекс u представляє вектор дисперсії σ_u^2 управляючої функції ВНЗ. Остаточно середня кількість інформації в одній посліці блоку засобів передані обробці інформації (рис.2) дорівнюватиме:

$$I = -\frac{1}{2} \log(1-r^2) = -\frac{1}{2} \log \left(1 - \frac{1}{1 + \sigma_u^2 / \sigma_x^2} \right) = \frac{1}{2} \log \frac{\sigma_x^2 + \sigma_u^2}{\sigma_u^2} \quad (25)$$

Підводячи підсумки з викладених двох частин публікації, можна стверджувати, що загальне завдання аналізу витоків і концепції побудови моделей інформаційних систем у кібернетичній педагогіці на прикладі ВНЗ виконано. Незважаючи на значну кількість моделей формульного типу, матеріал публікації виглядає достатньо прозоро і представляє собою своєрідний розгалужений алгоритм. Використовуючи сучасні математичні пакети зможемо виконати числові розрахунки в середовищі Excel.

Висновки

Виконаний аналіз джерел витоків ентропійних та інформаційних моделей систем навчання, запропоновані концепції побудови інформаційно-ентропійних систем, розроблені нові підходи дослідження їх у галузі кібернетичної педагогіки на прикладі підвищення ефективності функціонування ВНЗ.

Література

1. Шеннон К. Работы по теории информации и кибернетике. - М.: Интересная литература, -1963. - 829 с.
2. Яглом А.М., Яглом И.М. Вероятность и информация. - М.: - Физико-математическая литература, -1960. - 315с.
3. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. - М.: - Физико-математическая литература, - 1960. - 883 с.
4. Кузьмин И.В., Кедров В.А. Основы теории информации и кодирования. - К.: Высшая школа, -1977. - 280 с.
5. Кузьмин И.В. Оценка эффективности и оптимизация автоматических систем контроля и управления. - М.: - Советское радио, - 1971, - 295 с.
6. Цымбал В.П. Теория информации и кодирование. - К.: Вища школа, - 1982. - 304 с.
7. Красовский А.А. Пospelов Г.С. Основы автоматики и технической кибернетике. - М. - Л.: Госэнергоиздат. - 1962, - 600 с.
8. Красовский А.А. Фазовое пространство и статистическая теория динамических систем - М.: Наука. - 1974, - 232 с.
9. Хайкин Саймон. Нейронные сети: полный курс: Пер. с англ. - М.: ООО «И.Д. Вильямс», - 2006. - 1104с. ISBN 5-8459-0890-6.
10. Мокін Б.І., Маримончик Ю.В. Методологічні основи матеріального стимулювання ефективності діяльності персоналу ВНЗ із підготовки науково-педагогічних кадрів та створення якісної наукової продукції // Вісник Вінницького політехнічного інституту. - 2004. - №2. - с.26 - 31.