

Олійник О. В.,

*кандидат фізико-математичних наук, ст. викладач кафедри математики,
інформатики та методики навчання
ДВНЗ «Переяслав-Хмельницький державний педагогічний університет
імені Григорія Сковороди»*

ОСОБЛИВОСТІ ВИВЧЕННЯ ТЕМИ «ОСНОВНІ АЛГОРИТМИ СТИСКАННЯ ІНФОРМАЦІЇ БЕЗ ВТРАТ» КУРСУ «ТЕОРІЯ ІНФОРМАЦІЇ»

Ми живемо в інформаційному суспільстві, яке характеризується стрімким розвитком інформаційно-комунікаційних технологій, що проникають в усі галузі людської діяльності. Різні задачі обробки інформації потребують відповідної підготовки, інформаційної культури усіх членів суспільства. Кількість інформації швидко росте, людина виявляється зануреною в океан інформації. Щоб «не потонути» в цьому океані, необхідно мати теоретичні знання в області інформаційно-комунікаційних технологій, вміти ефективно використовувати персональні комп'ютери і телекомунікаційні засоби.

Для майбутніх вчителів інформатики та математики досить важливо знати, що в основі дисципліни «Теорія інформації та кодування» лежить розуміння різних підходів до обґрунтувань науки, розкриття сучасних наукових концепцій, понять, методів математичного моделювання та дослідження. Крім того, в процесі вивчення основних напрямків досліджень в теорії інформації та кодування в інформаційних системах, студенти мають оволодіти методами теорії кодування, обробки та захисту інформації при наявності завад, управління потоками інформації в інформаційних мережах, проводити оцінку ефективності різних методів кодування. Крім того, курс базується на широкому використанні попередніх знань з предметів «Математичний аналіз», «Теорія ймовірностей і математична статистика», «Дискретна математика», «Фізика», «Основи електротехніки», «Програмування».

У педагогіці розроблено різні аспекти теорії і практики міжпредметних зв'язків (И.Д. Зверев, Н.М. Скаткин, В.А. Далингер, В.А. Байдак, О.Н. Лучко, Л.В. Смоліна, Е.А. Кашина та ін.): обґрунтована об'єктивна необхідність відображати реальні взаємозв'язки світу у навчанні, визначена світоглядна функція міжпредметних зв'язків, їх роль в

загальному розумовому розвитку учнів, виявлено їх позитивний вплив на формування системи знань, розроблена методика скоординованого викладання різних навчальних предметів.

Тут можна відмітити, що у роботах, присвячених міжпредметним зв'язкам інформатики з іншими науковими дисциплінами у вищій педагогічній освіті (А.Р. Гейн, С.Б. Поморцева та ін.) також не приділяється належної уваги теорії інформації, адже саме теорія інформації, завдяки універсальності її системоутворюючого поняття «інформація» та інформаційного підходу в науковому пізнанні, більш, ніж які-небудь інші наукові дисципліни, сприяє вивченню міжпредметних зв'язків, які найбільш ярко охарактеризував Й.Г. Песталоцці: «Приведи у своїй свідомості всі по суті взаємопов'язані між собою предмети в той самий зв'язок, в якому вони дійсно знаходяться в природі».

Внаслідок вивчення теми «Основні алгоритми стискання інформації без втрат» студенти повинні: знати алгоритми кодування для джерел повідомлень і передачі даних по каналам зв'язку, принципи побудови кодів та їх використання в сучасних комп'ютерних інформаційних системах; вміти використовувати основні принципи кодування інформації з метою підвищення ефективності коду.

Кодування в широкому сенсі – перетворення повідомлень у сигнал. Кодування у вузькому сенсі – представлення дискретних повідомлень визначеними поєднаннями символів. Кодування відбувається з одного боку для того, щоб забезпечити найкращого узгодження характеристик джерела повідомлень та каналу, з іншого боку, для підвищення достовірності передачі інформації при наявності перешкод. Крім того, при виборі системи кодування (представлення повідомлень) намагаються забезпечити простоту й надійність апаратної реалізації пристроїв. У процесі кодування повідомлень довга послідовність (наприклад, з N символів) зазвичай формується з кодових комбінацій, кожна з яких відповідає одному знаку (букві). Для того щоб кожен символ (наприклад, двійкового) коду ніс максимум інформації, символи кодової комбінації повинні бути незалежними та приймати значення (0 і 1) з рівними ймовірностями. Побудова ефективних кодів у випадку статистичної незалежності символів повідомлень спирається на методики Шеннона та Фано. Код будується так: знаки виписують в таблицю у порядку спадання їх ймовірностей у повідомленнях. Потім розділяють на дві групи так, щоб значення сум ймовірностей в кожній групі були близькими. Всі знаки однієї з груп у відповідному розряді кодуються, наприклад, одиницею, а знаки другої групи кодуються нулем. До кожної отриманої у процесі поділу групи застосовують дану операцію до тих пір, поки в результаті чергового поділу в кожній групі не залишиться по одному.

У наслідок того, що дана методика не має чітких рекомендацій щодо поділу на підгрупи, вона не завжди приводить до гідних результатів.

Розглянемо методику кодування Хаффмана, яка не має цього недоліку.

Знаки, що кодуються, так само, як при використанні методу Шеннона-Фано, потрібно розташувати у порядку спадання ймовірностей. Далі, на кожному етапі дві наступні позиції списку замінюються однією, та їй приписують ймовірність, що дорівнює сумі ймовірностей позицій, які замінюються. Після того відбувається пересортування списку за спаданням ймовірностей, із збереженням інформації про те, які саме знаки об'єднувалися на кожному етапі. Процес продовжується до тих пір, доки не залишиться одна єдина позиція з ймовірністю, що дорівнює 1.

Чудовою властивістю кодів, які побудовано за допомогою методик Шеннона-Фано або Хаффмана, є префіксність. Вона полягає в тому, що жодна комбінація коду не є початком іншої, більш довгої комбінації. Це дозволяє при відсутності помилок здійснювати однозначне декодування ряду кодових комбінацій, які йдуть одна за одною, між якими відсутні роздільні символи.

У матеріалах Луцького національного технічного університету <http://elib.lutsk-ntu.com.ua/book/knit/ki/2013/13-34/page12.html> представлено наступні матеріали стосовно розв'язування практичних задач за методами Шеннона-Фано та Хаффмана. Побудуємо таблицю кодів за методом *Шеннона-Фано* для повідомлень, заданих розподілом ймовірностей. За методом Шеннона-Фано:

Таблиця 1.

Буква, x_i	Ймовірність, p_i	Код	Довжина коду, l_i	$p_i l_i$
А	0,6	1	1	0,6
Б	0,2	0 1	2	0,4
В	0,1	0 0 1	3	0,3
Г	0,04	0 0 0 1	4	0,16
Д	0,025	0 0 0 0 1	5	0,125
Е	0,015	0 0 0 0 0 1	6	0,09
Ж	0,01	0 0 0 0 0 0 1	7	0,07
З	0,01	0 0 0 0 0 0 0	7	0,07
	=1			=1,815

За методом Хаффмана:

Таблиця 2.

Буква, x_i	Ймовірність, p_i	Код	Довжина коду, l_i	$p_i l_i$
А	0,6	0	1	0,6
Б	0,2	1 1	2	0,4
В	0,1	1 0 1	3	0,3
Г	0,04	1 0 0 1	4	0,16
Д	0,025	1 0 0 0 1	5	0,125
Е	0,015	1 0 0 0 0 1	6	0,09
Ж	0,01	1 0 0 0 0 0 1	7	0,07
З	0,01	1 0 0 0 0 0 0	7	0,07
	=1			=1,815

Середня довжина коду визначена в обох випадках і має значення

$$\bar{L} = \sum_{i=1}^8 l_i p_i = 1,815 .$$

Недоліки алгоритмів Шеннона-Фано і Хаффмена:

1. Необхідність побудови таблиці ймовірностей для кожного типу даних, що стискаються.
2. Необхідність зберігання таблиці кодів разом із стисненим повідомленням, що знижує ефект стиснення.
3. Мінімальна довжина кодових слів не може бути менше 1, тоді як ентропія повідомлень може бути близькою до нуля.
4. Найкращий ефект досягається, коли ймовірності є цілими від'ємними степенями двійки. На практиці такі ситуації створюються штучно, що викликає ускладнення алгоритму.

При вивченні курсу «Теорія інформації та кодування» вирішується ряд актуальних задач професійного становлення майбутніх вчителів інформатики. Серед них: узагальнення теоретичних знань, формування цілісного уявлення про сутність та специфіку науки, що вивчається; засвоєння методології інформаційного підходу в дослідницькій діяльності; побудова, оцінка і корекція моделей майбутньої педагогічної діяльності, орієнтованої на виховання і розвиток учнів у процесі навчання їх інформатиці. Дуже важливим при вивченні цього курсу показати міжпредметні зв'язки між дисциплінами, такими як, теорія ймовірності, програмування та теорія інформації.

Список використаної літератури

1. Жалдак М.І., Морзе Н.В. Методика ознайомлення учнів з поняттям інформації / М.І. Жалдак, Н.В. Морзе //Комп'ютер у школі та сім'ї. – 2000. – №4. – С.11– 16.
2. Морзе Н.В. О новом курсе «Методика преподавания информатики в средних учебных заведениях» в педагогических вузах» //Использование информационной технологии в учебном процессе. – Киев: КГПИ, 1990. – С.15 – 21.
3. Морзе Н.В. Концепція методичної підготовки майбутніх вчителів інформатики в умовах багаторівневої освіти. //Наука і сучасність: Зб. наук. праць. – К.: НПУ ім. М.П. Драгоманова. – К.: Логос, 2001. – Т. XXVI – С. 57–65.
4. Пугач В.И., Добудько Т.В. Методика преподавания информатики. – Самара: СамГПИ, 1993. – 250 с.
5. Швецкий М.В. Методическая система фундаментальной подготовки будущих учителей информатики в педагогическом ВУЗе в условиях двухступенчатого образования. Автореф.дисс. докт. пед. наук.- СПб., 1994.- 36 с.