

Чертов О.Р., Тавров, Ден. Групповая анонимность данных в микрофайлах

На сегодняшний день обнаружение первичных данных стало обычным способом обеспечения информацией разнообразных исследований. В то же время первичные данные должны быть защищены, чтобы предупредить нежелательную утечку информации и нарушение ее конфиденциальности. Были разработаны различные алгоритмы и подходы к обеспечению приватности для индивидуальных респондентов. Однако, задача обеспечения анонимности для группы респондентов все еще не решена. В данной работе мы рассматриваем практические аспекты обеспечения групповой анонимности данных. Мы предлагаем формальное определение такой задачи, анализируем ее основные особенности и виды, а также изучаем вопросы предотвращения утраты полезности данных. В частности, мы обращаемся к задаче минимизации искажения данных во время изменения микрофайла. Для ее решения мы вводим специальную метрику и предлагаем стратегию использования последней для получения оптимального маскирования микроданных.

Ключевые слова: групповая анонимность, статистический контроль за раскрытием информации, защита информации от несанкционированного доступа, микрофайл, полезность данных.

УДК 681.324

Аспір. У.В. Поліщук¹;

проф. Р.О. Ткаченко, д-р техн. наук – НУ "Львівська політехніка"

**МЕТОДИ І ЗАСОБИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ
НЕЙРОМЕРЕЖЕВОГО УЩІЛЬНЕННЯ ДАНИХ
НА ПРИКЛАДІ ЗОБРАЖЕНЬ**

Описано особливості ущільнення та відновлення даних (на прикладі зображень) з використанням автоасоціативних нейроподібних структур геометричних перетворень для реалізації процедури виділення головних компонентів та відновлення початкових даних на їхній основі. Проаналізовано базовий алгоритм автоасоціативних нейроподібних структур у режимах компресії-декомпресії зображень та вміст архіву скомпресованих даних. Визначено варіанти вдосконалення базового нейромережевого методу шляхом застосування додаткових засобів ущільнення, що можуть бути використані для файлів зображень, мовних повідомлень і текстових даних.

Вступ. Методи та засоби зменшення обсягу "громіздких" даних, зокрема статичних растрових зображень, спрямовані на забезпечення ефективного їхнього представлення під час зберігання у пам'яті комп'ютера чи передачі каналами зв'язку. Серед багатьох методів ущільнення даних виділяється підхід, який передбачає відображення даних із простору великої розмірності у простір зменшеної розмірності за допомогою лінійного перетворення Карунена-Лоєва, базисні функції якого є власними векторами коваріаційної матриці вхідного сигналу. Таке перетворення є оптимальним згідно з критерієм досягнення декореляції вхідного сигналу, оскільки на перший базисний вектор, у середньому, припадає найбільший обсяг інформації про зображення, на другий – максимум із тої, що залишилась, і т. ін., іншими словами, здійснюється лінійний аналіз головних компонент (ГК). Проте, враховуючи, що базисні вектори перетворення Карунена-Лоєва обчислюються на основі статистичних властивостей зображень, що перетворюються, виникає необхідність

¹ Наук. керівник: проф. Ю.М. Рашкевич, д-р техн. наук – НУ "Львівська Політехніка"

записувати ці вектори в ущільнений файл для використання декодером [5]. Окрім того, немає швидких алгоритмів їхнього обчислення та здійснення зворотних перетворень, що робить цей метод суто теоретичним.

Основи базового нейромережного методу. Справді, ефективний метод аналізу головних компонент, який набув практичного застосування, базується на використанні штучних багат шарових нейромереж прямого поширення автоасоціативного типу, зокрема побудованих на основі моделі геометричних перетворень (ГП) [6]. Особливості розробленого методу компресії для структур цього типу, що полягають у представленні скомпресованих даних у форматі чисел з фіксованою комою, дають змогу доповнити базовий нейромережний метод додатковими засобами, які підвищують коефіцієнт ущільнення та покращують якість відтворюваних даних [7].

Як відомо, автоасоціативна нейронна мережа є багат шаровим перцептроном прямого поширення, навченим відтворювати вхідні сигнали на виході. Кількість входів такої мережі збігається з кількістю виходів, а кількість нейронів прихованого шару є меншою, що дає змогу ущільнювати дані, зменшуючи їхню розмірність, одночасно реалізуючи стандартний алгоритм виділення головних компонент (Карунена-Лоєва) [1, 2] (рис. 1).

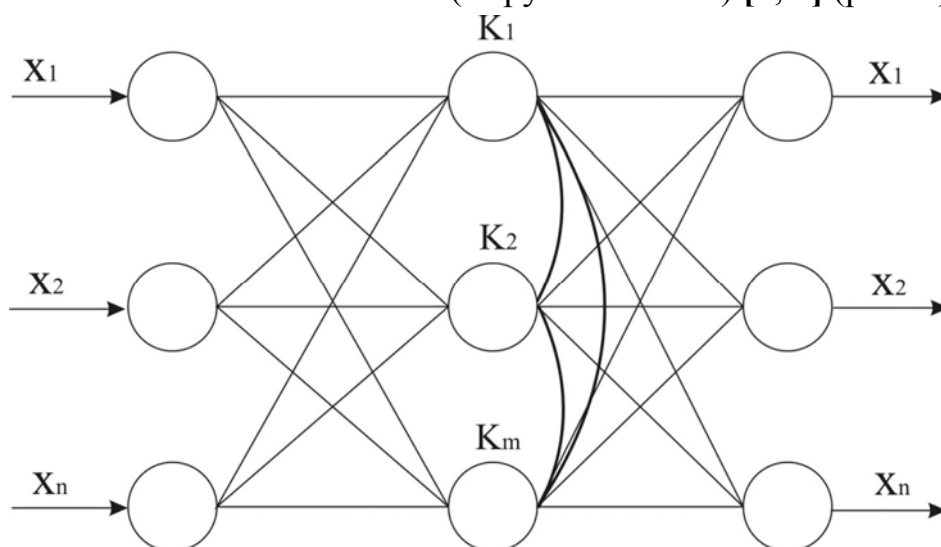


Рис. 1. Автоасоціативна нейронна мережа ГП

Застосування нейронної мережі для будь-яких завдань у загальному випадку передбачає два послідовні етапи навчання і, відповідно, використання навченої мережі. Однак, для задач ущільнення зображень можливими є два варіанти: навчання і застосування мережі на одному і тому самому зображенні; навчання на певному еталонному зображенні та застосування навченої раніше нейромережі в режимі використання для ущільнення і відновлення інших зображень.

Навчання компресуючої мережі ГП в автоасоціативному режимі застосування здійснюється на основі матриці реалізацій, елементами якої є яскравості пікселів $x_{i,j}$ у випадку такого кодування зображень. Кожен рядок матриці відповідає вектору-реалізацій, зокрема окремому блокові зображення розміром $n \times n$ пікселів (рис. 2).

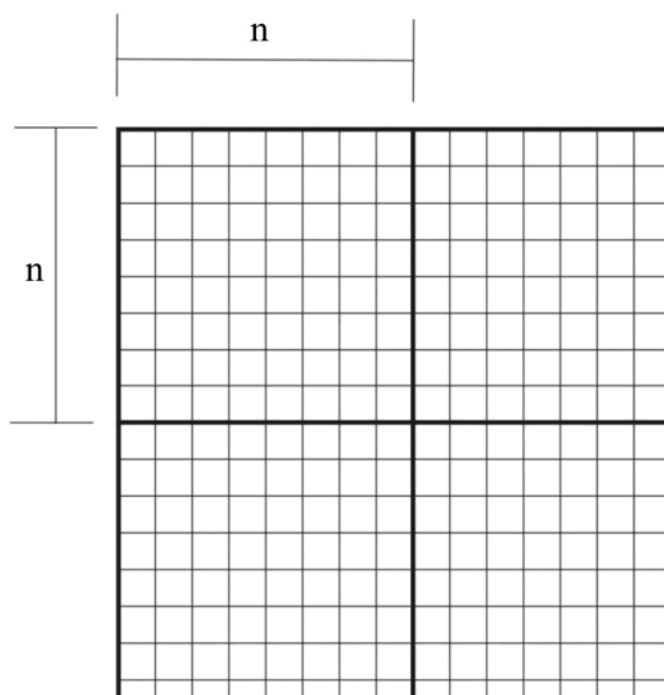


Рис. 2. Вибір блоку зображення

Загальне кількість векторів (кількість блоків зображення) дорівнює M , а кількість ознак-пікселів векторів (стовпців матриці) дорівнює N .

$$x_{i,j} = \begin{Bmatrix} x_1 & x_2 & \dots & x_N \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \end{Bmatrix}, \text{ де } i = \overline{1, M}, j = \overline{1, N} \quad (1)$$

З метою істотного підвищення коефіцієнта компресії запропонований метод застосування еталонного зображення для навчання нейронної мережі з подальшим її використанням на робочих зображеннях.

На вхід нейронної мережі для навчання подається певне еталонне зображення підвищеної складності. Питання вибору еталону докладно розглянуто у [4].

Параметри нейромережі, отримані в процесі навчання еталонного зображення, зберігаються в пам'яті комп'ютера та надалі використовуються для виділення головних компонент рухомих зображень. Кількість перетворень (кількість нейронів прихованого шару) під час навчання на еталонному зображенні k дорівнює кількості кроків, необхідних для виділення ГК тестового зображення.

Для здійснення операцій ущільнення-відновлення зображення необхідно виконати такі перетворення.

Знаходимо базовий рядок для центрованих елементів масиву:

$$xB_j^{(k)} \approx \max_i \left\{ \sum_{j=1}^n (x1_{i,j}^{(k)})^2 \right\}. \quad (2)$$

Виконуємо нормування базового рядка до одиничного радіуса:

$$XB_j^{(k)'} = \frac{x B_j^{(k)'}}{D^{(k)'}} \text{ для } j = \overline{1, N}, \quad (3)$$

де $D^{(k)'}$ – віддаль від початку координат до базової точки.

Обчислюємо коефіцієнт $z_j^{(k)'}$ – віддаль від i -ї точки до нормальної гіперплощини на k -му кроці ортогональних перетворень:

$$z_i^{(k)'} = \frac{1}{D^{(k)'}} \sum_{j=1}^n (x l_{i,j}^{(k)'} \times x B_j^{(k)'}). \quad (4)$$

4. Виконуємо k -й крок ортогональних перетворень – проектування усіх точок-реалізацій на нормальну гіперплощину вимірності $N-1$:

$$x_{i,j}^{(k+1)'} = x_{i,j}^{(k)'} - X B_j^{(k)'} \times z_i^{(k)'}, \text{ для } i = \overline{1, M}, j = \overline{1, N}. \quad (5)$$

У пам'яті комп'ютера для подальшого застосування запам'ятовуємо величини:

1. $x c_j'$, для $i = \overline{1, M}, j = \overline{1, N}$;
2. $x B_j^{(k)'}$ та $D^{(k)'}$ для $k = \overline{1, k_{\max}}$.

Виділення ГК біжучого зображення здійснюється на основі коефіцієнтів, отриманих внаслідок перетворень (1)-(4):

$$z_i^{(k)} = \frac{1}{D^{(k)'}} \sum_{j=1}^n \left(x l_{i,j}^{(k)} \times x B_j^{(k)'} \right), \quad (6)$$

де $x l_{i,j}^{(k)}$ – елементи центрованої матриці тестового зображення.

Відновлення первинного зображення.

1. Приймаємо

$$x l_{i,j}^{(k_{\max}+1)} = 0, \text{ для } i = \overline{1, M}, j = \overline{1, N}. \quad (7)$$

2. Відновлюємо центровану матрицю $x l_{i,j}^{(k)}$.

Для $k = k_{\max}$ кроки зворотніх перетворень здійснюються послідовно, починаючи від кроку $k-1$ до 0; для $i = 1$ – крок від 1 до M ; для $j = 1$ – крок від 1 до N

$$x l_{i,j}^{(k)} = x l_{i,j}^{(k+1)} + X B_j^{(k)'} \times z_i^{(k)}. \quad (8)$$

3. Відновлюємо початковий масив даних $x_{i,j}$.

Для $i = 1$ – крок від 1 до M , для $j = 1$ – крок від 1 до N , виконуємо:

$$x_{i,j} = x l_{i,j}^{(0)} + x c_j'. \quad (9)$$

Таким чином, архів збережених даних містить масив значень коефіцієнтів $z_j^{(k)'}$, отриманих унаслідок виконання алгоритму (дані коефіцієнти безпосередньо відображають сигнали головних компонент) та масив пара-

метрів нейромережі – елементів базових рядків та параметрів структури, де всі зазначені елементи архіву подаються у форматі чисел з фіксованою комою. Зазначена особливість, а також висока швидкодія нейромереж ГП надає змогу застосувати деякі додаткові засоби підвищення якості перетворень.

Додаткове застосування засобів ущільнення, що використовуються для текстових даних. Враховуючи, що після нейромережного ущільнення дані про зображення подаються у форматі цілих чисел із фіксованою комою, виникає додаткова можливість застосувати до архіву одну з програм архівації, наприклад 7Zip, як це зазвичай використовується для архівації текстових повідомлень. Як виявилось, це забезпечує помітне збільшення загального коефіцієнта компресії, без внесення жодних спотворень у відновлювані зображення. Результати виконаних експериментів подано в табл. 1, де результати додаткового компресування відображені колонками 7-8.

**Табл. 1. Результати кодування зображення "Mandrill_gray",
розмір блоку зображення 8×8**

Розмір файлу зображення, КБ	К-ть НЕ прихованого шару	К-ть видалених молодших бітів	PSNR	SSIM	Розмір.рас файлу, КБ	Метод додаткової архівації	Розмір.7z файлу, КБ	CR
257	37	1	30	0,982	186	PPMd	168	1,5
257	37	1	30	0,982	186	LZMA	166	1,5
257	37	1	30	0,982	186	BZip2	172	1,5
257	19	3	25	0,943	80,7	PPMd	68,4	3,7
257	19	3	25	0,943	80,7	LZMA	66,1	3,8
257	19	3	25	0,943	80,7	BZip2	70,5	3,6
257	50	4	32,42	0,990	170	PPMd	134	1,9
257	50	4	32,42	0,990	170	LZMA	133	1,9
257	50	4	32,42	0,990	170	BZip2	141	1,8
257	64	4	35,21	0,994	210	PPMd	161	1,6
257	64	4	35,21	0,994	210	LZMA	160	1,6
257	64	4	35,21	0,994	210	BZip2	169	1,5
257	64	5	28,95	0,979	178	PPMd	125	2
257	64	5	28,95	0,979	178	LZMA	123	2
257	64	5	28,95	0,979	178	BZip2	131	2

Очевидно, що описаний підхід виявляється ефективним також для задач ущільнення з даними іншого походження.

Метод комітету нейронних мереж. Комітет нейронних мереж розглядають як ефективний засіб розширення можливостей нейромережних методів для різноманітних сфер застосування: під час прогнозування, класифікації, розв'язання оптимізаційних задач. Комітети нейронних мереж ГП, які мають швидкі алгоритми навчання, надають додаткові можливості для побудови комітетів достатньо великих розмірів. На рис. 3 зображено розроблену структуру комітету нейронних мереж ГП, що попередньо навчаються на відповідних еталонах і функціонують на основі відомого в нейрокомп'ютерній техніці принципу WTA (переможець забирає все).

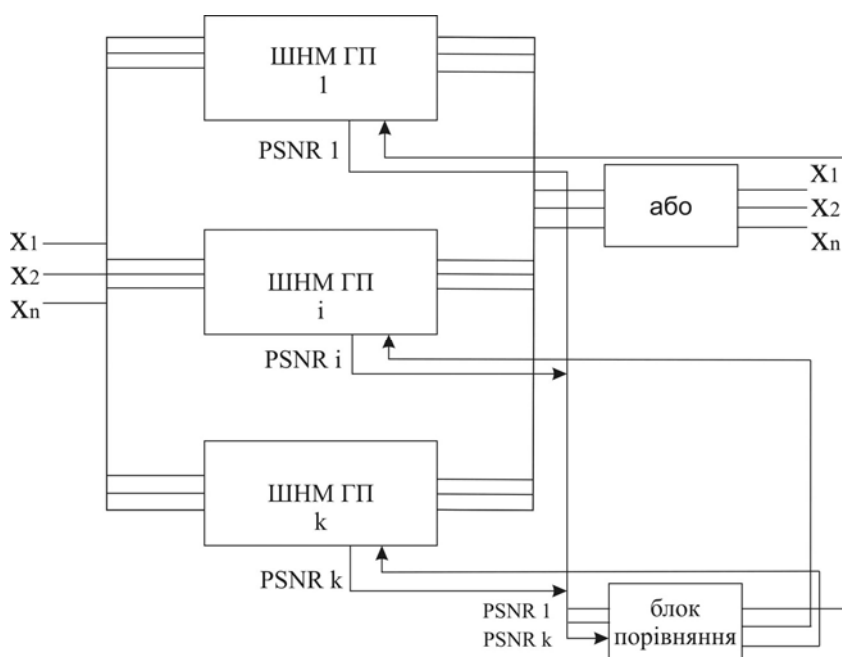


Рис. 3. Структура комітету нейронних мереж

Таким чином k ШНМ ГП, що формують комітет, попередньо навчаються на k зображеннях-еталонах здійснювати компресію-декомпресію зображень, де коефіцієнт ущільнення встановлюється заздалегідь. Еталони повинні бути достатньо різноманітними в плані текстури та композиції рисунку. Параметри кожної з ШНМ отримують шляхом попереднього навчання, не змінюються в процесі компресування даних, а отже, не входять до архіву скомпресованих даних.

Зображення, що підлягає ущільненню, проходить режим послідовної компресії-декомпресії на кожній з k ШНМ, оцінюються коефіцієнти якості PSNR та SSIM, що характеризує міру структурної схожості зображень, та обирається варіант із найкращими показниками якості. Тобто реалізується принцип WTA (переможець обирає все), а вибір переможця здійснюється за обраними показниками якості зображень, що порівнюються.

Найефективнішим варіантом побудови цієї структури є варіант, коли здійснюється пофреймове порівняння зображень. Тобто зображення пофреймово подається на k компресуючих мереж, а для окремих фреймів обирається своя компресуюча мережа. Помітне підвищення коефіцієнта компресії та якості відновлюваних зображень досягається ціною збільшення відповідних часових затримок. У цьому випадку варто взяти до уваги, що навчання всіх k ШНМ здійснювалось попередньо. Отже, для малих k реально забезпечується навіть пришвидшення процедур компресії-декомпресії порівняно з варіантом одиначної ШНМ, для якої передбачене попереднє навчання. Однак, для великих значень k часові затримки для комітету нейромереж стають помітними.

Вибір числа k нейромереж, які формують комітет, набору k еталонних зображень показників якості відтворюваних зображень на підставі яких реалізується принцип WTA залежатиме певною мірою від класів зображень, що підлягають компресуванню і обираються шляхом виконання певних експериментів.

Висновки. 1. Нейроподібні структури автоасоціативного типу треба розглядати як ефективні засоби реалізації ущільнення даних, зокрема зображень.

2. Зберігання архіву даних для компресуючих структур геометричних перетворень у форматі цілих чисел із фіксованою комою дає змогу застосувати додаткові засоби компресії без втрати якості.

3. Принципова здатність здійснювати навчання автоасоціативних структур геометричних перетворень на еталонах без істотної втрати якості перетворень забезпечує можливість створення комітету нейромереж із високими показниками коефіцієнтів ущільнення та якості.

4. Головні компоненти, отримані за допомогою ущільнюючих нейромереж геометричних перетворень, можна безпосередньо застосувати для виконання операцій розпізнавання у системах технічного зору.

Література

1. **Gaetan Kerschen.** Feature extraction using auto-associative neural networks / Gaetan Kerschen, Jean-Claude Golinval // INSTITUTE OF PHYSICS PUBLISHING SMART MATERIALS AND STRUCTURES. Smart Mater. Struct. – 13 (2004). – 211-219 pp.

2. **Nandakishore Kambhatla,** Todd K. Leen. Dimension Reduction by Local Principal Component Analysis / Department of Computer Science and Engineering, Oregon Graduate Institute of Science and Technology // Portland, Oregon 97291-1000, U.S.A.

3. **Simon Haykin,** Robert D. Dony. Neural Network Approaches to Image Compression / PROCEEDINGS OF THE IEEE // FEBRUARY, 1995. – Vol. 83, No. 2. – 288-301pp.

4. **Шлезингер М.И.** О построении эталонов для корреляционных читающих автоматов / М.И. Шлезингер, Л.А. Святогор // III Всесоюзная конференция по информационно поисковым системам и автоматизированной обработке научно-технической информации. – М. : Всесоюз. ин-т научной и техн. информации. – 1967. – Т. 3. – С. 129-139.

5. **Сэломон Д.** Сжатие данных, изображений и звука. – М. : Изд-во "Техносфера". – 2004. – 236 с.

6. **Ткаченко Р.О.** Нова парадигма штучних нейронних мереж прямого поширення // Вісник Державного університету "Львівська політехніка". – Сер.: Комп'ютерна інженерія та інформаційні технології. – 1999. – № 386. – С. 43-54.

7. **Поліщук У.** Ущільнення зображень за допомогою нейроподібних структур моделі геометричних перетворень / У. Поліщук, О. Ткаченко, Ю. Цимбал // Вісник Національного університету "Львівська політехніка". – Сер.: Комп'ютерні системи та інформаційні технології. – 2010. – № 663. – С. 275-280.

Поліщук У.В., Ткаченко Р.О. Методы и средства повышения эффективности нейросетевого сжатия данных на примере изображений

Описаны особенности сжатия и восстановления данных (на примере изображений) с использованием автоассоциативных нейроподобных структур геометрических преобразований для реализации процедуры выделения главных компонентов и восстановления исходных данных на их основе. Проанализированы базовый алгоритм автоассоциативных нейроподобных структур в режимах компрессии-декомпрессии изображений и содержимое архива сжатых данных. Показаны варианты усовершенствования базового нейросетевого метода путем применения дополнительных средств сжатия, которые могут быть использованы для файлов изображений, речевых сообщений и текстовых данных.

Polishchuk U.V. Methods and means of increasing the efficiency of the neural data compression by the example of images

The features of compression and decompression data (for example images) using auto-associative neural structures of geometric transformations to implement procedures of the principal components analyzing and restore the original data based on them is been vi-

ewed. Analyzed the basic algorithm of auto-associative neural structures in the compression-decompression modes of images and contents of the compressed data archive. The following options to improve the basic neural network method by applying additional compression software that can be used for image files, voice and text data.

Keywords: neural colour image compression, image-pattern, neurolike Geometrical Transformation Machine structure.

УДК 621.086.065

Проф. Б.М. Гевко, д-р техн. наук; доц. О.Л. Ляшук, канд. техн. наук – Тернопільський НТУ; проф. М.П. Мартинців, д-р техн. наук; доц. О.М. Удовиський, канд. техн. наук – НЛТУ України, м. Львів

РОЗРОБЛЕННЯ МАТЕМАТИЧНОЇ МОДЕЛІ ТА ОБҐРУНТУВАННЯ ПАРАМЕТРІВ ГНУЧКИХ КАНАТНИХ КОНВЕЄРІВ

Запропоновано математичну модель гнучких канатних конвеєрів для транспортування контейнерів зі сипкими матеріалами. Отримано залежності для вибору натягу канату. На основі аналізу наведених залежностей побудовано графіки зміни зусиль залежно від умов роботи конвеєра, що дає змогу вибрати його оптимальні параметри. Встановлено, що при транспортуванні контейнерів та їхньої зупинки для навантаження і розвантаження в ланцюгових конвеєрах важко забезпечити плавність руху. При зупинках і початку руху, внаслідок зазорів між ланками ланцюга та зірочками, виникають ударні динамічні навантаження, що призводить до руйнування ланок ланцюгів та зірочок, а також до розсипання сипких матеріалів з контейнерів.

Для транспортування сипких матеріалів широко використовують конвеєри з гнучким тяговим елементом [1-4]. У відомих конвеєрах таким гнучким елементом слугують ланцюги. Однак під час транспортування контейнерів та їхньої зупинки для навантаження і розвантаження в ланцюгових конвеєрах важко забезпечити плавність руху. При зупинках і початку руху, внаслідок зазорів між ланками ланцюга та зірочками, виникають ударні динамічні навантаження, що призводить до руйнування ланок ланцюгів та зірочок, а також розсипання сипких матеріалів з контейнерів.

Тому доцільно як тяговий елемент використати канати [5, 6]. Розрахункову схему контейнера з канатною тягою представлено на рис. 1.

Контейнери, які переміщує конвеєр, опира-

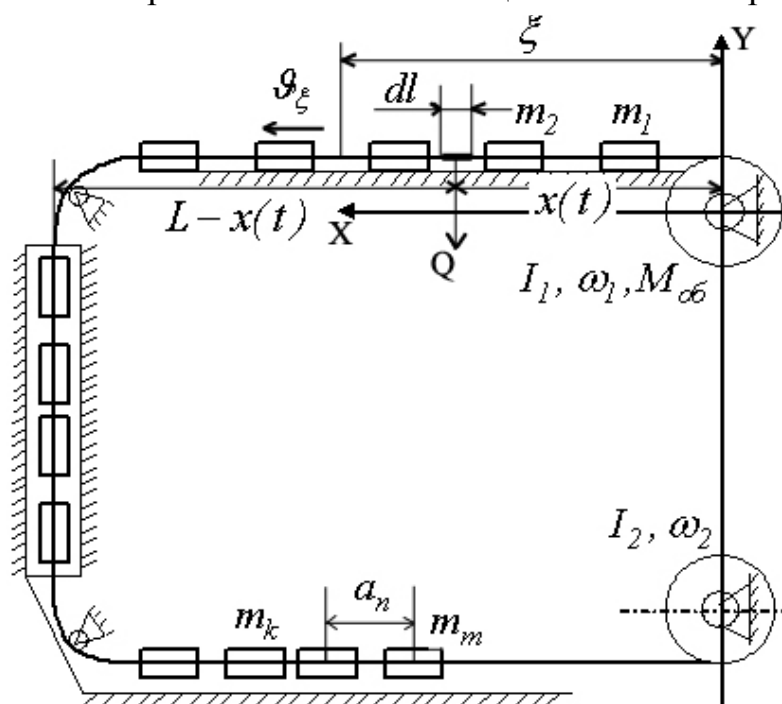


Рис. 1. Розрахункова схема гнучкого канатного конвеєра