

МУРАШИНИЙ АЛГОРИТМ КЛАСТЕРИЗАЦІЇ

Горносталь Олександр, Дорогий Ярослав

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Анотація

В роботі розглянуто питання створення алгоритму кластеризації на базі мурашиних колоній для задачі розпізнавання мови. Створений алгоритм інтегрований в аудіовізуальну систему розпізнавання мови і протестований в шести режимах розпізнавання.

Abstract

The paper considers the problem of creating an algorithm for clustering based on ant colonies for the problem of speech recognition. The created algorithm is integrated into the audio-visual speech recognition system and tested in six modes of recognition.

Вступ

Даний алгоритм використовується для вирішення проблем оптимізації. Наприклад, його активно застосовують для вирішення:

- задачі комівояжера [1];
- задачі кластеризації даних [2-3];
- задачі текстового пошуку [4-5].

Мурашиний алгоритм кластеризації

Алгоритм можна пояснити наступними 4-ма правилами:

1. Мурахи шукають їжу і різні шляхи для її пошуку.
2. Кожна мураха виділяє феромон.
3. Мурах принадажує феромон, тому шляхи з більшою кількістю феромону є більш інтенсивними.
4. Феромон з часом випаровується. Відповідно, падає інтенсивність шляхів з меншою кількістю феромону.

Алгоритм складається з наступних кроків:

1. Визначається початкова кількість феромону для кожного елемента таблиці феромонів за формулою 1:

$$\tau_0 = \frac{f_{fit}(S_{best})}{N} \quad (1)$$

де

f_{fit} - фітнес-функція;

S_{best} - найкращий розв'язок;

N - розмірність задачі.

2. Обчислюється ймовірнісна функція вибору того чи іншого шляху пошуку (2):

$$P_k(r, s) = \left\{ \frac{[\tau(r, s)]^\alpha \cdot [\eta(r, s)]^\beta}{\sum_{\mu \in J_k(\tau)} [\tau(r, s)]^\alpha \cdot [\eta(r, s)]^\beta} \right\} \quad (2)$$

де

$P_k(r, s)$ - ймовірність пересування мурахи k з точки r в точку s ;

$\tau(r, s)$ - кількість феромону на шляху;

$\eta(r, s)$ - фітнес-функція пересування шляхом;

α, β - коефіцієнти, що залежать від типу конкретної задачі;

$J_k(r)$ - множина точок пересування та r - остання точка, яку відвідала мураха.

3. Обчислюються нові локальні значення таблиці феромонів, що запобігає руху мурах одними і тими ж шляхами (3):

$$\tau(r, s) = (1 - \sigma) \cdot \tau(r, s) + \sigma \cdot \tau_0 \quad (3)$$

де

σ - локальний коефіцієнт оновлення.

4. Обчислюються нові глобальні значення таблиці феромонів (4):

$$\tau(r, s) = (1 - \rho) \cdot \tau(r, s) + \rho \cdot \Delta\tau(r, s) \quad (4)$$

де

ρ - глобальний коефіцієнт оновлення;

$$\Delta\tau(r, s) = \begin{cases} f_{fit}(S_{best}) & \text{if } (r, s) \in S_{best} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Кластеризація – процес поділу об'єктів або даних на класи за якимось критерієм. В якості критерію виступає значення відстані між об'єктами, яке розраховується відповідно до обраного типу цієї відстані: Евклідова, Махаланобіса або відстань міських кварталів Мінковського. Центри кластерів є змінними, відносно яких приймається рішення і які можна отримати шляхом мінімізації Евклідової відстані між всіма елементами навчальної вибірки. Мета кластеризації – мінімізація суми відстаней міських кварталів за всіма N зразками тренувальної вибірки і віднесення їх до одного з K кластерів. Тобто, маємо (6):

$$D(x_i, c_k) = \sum_k \sum_{i=1}^N (x_i - c_k) \quad (6)$$

де

N - кількість зразків мови;

K - кількість кластерів.

Центри кластерів обчислюються за формулою (7):

$$c_k = \sum_{i \in c_k} \frac{x_i}{n_k} \quad (7)$$

Цільова функція для зразка i обчислюється за формулою (8):

$$G_i = \frac{1}{S_{ir}} \sum_{j=1}^{S_{ir}} d(x_j, m_i) \quad (8)$$

де

S_{ir} - кількість зразків мови в навчальній вибірці;

m_i - клас, до якого належить зразок вибірки.

Експериментальні дослідження

Запропонований алгоритм використаний для покращення роботи аудіовізуальної система розпізнавання, представленої в [6]. Побудована система була протестована на базі цифр від 0 до 9. Кожна цифра в базі даних повторюється десять разів кожним з десяти мовців. Для кожного оратора дев'ять прикладів кожної цифри використано для навчання, а приклад, що залишився, використано для тестування.

Усереднені результати розпізнавання для шести режимів розпізнавання представлені в таблиці 1. Режими розпізнавання наступні:

- тільки аудіо;
- тільки аудіо з використанням мурашиного алгоритму кластеризації;
- тільки відео;
- тільки відео з використанням мурашиного алгоритму кластеризації;
- аудіовізуальне розпізнавання;

– аудіовізуальне розпізнавання з використанням мурашиного алгоритму кластеризації.

Таблиця 1 – Усереднені результати розпізнавання

Класифікатор	Результат розпізнавання
ПММ (тільки аудіо)	92%
ПММ (тільки аудіо) + МАК	93%
ПММ (тільки відео)	51%
ПММ (тільки відео) + МАК	57%
ЗПММ (аудіовізуальна система)	98%
ЗПММ + МАК	98,5%

Висновки

Для режиму розпізнавання мови «лише аудіо» вектори акустичного спостереження (15 MFCC коефіцієнтів, які визначені з вікна 20 мс) моделюються з використанням ПММ. Для розпізнавання в режимі «аудіо-відео» використана ЗПММ з п'ятьма станами для зв'язаних вузлів як в аудіо-, так і в відеопотоках, без зворотних переходів і трьома зміщуваннями на кожний стан. Експериментальні результати показали, що рівень розпізнавання аудіовізуальної мови на основі ЗПММ з використанням МАК збільшується на 43,5% щодо розпізнавання мови в режимі «тільки аудіо». У порівнянні з багатопотоковою ПММ [7] запропонований варіант ЗПММ для аудіовізуальної системи розпізнавання показує кращі результати.

Список використаних джерел:

1. Coello C.A. Evolutionary algorithms for solving multi-objective problems [Текст] / C.A.Coello and others // Springer, 2007.
2. Lumer E.D. Diversity and adaptation in populations of clustering ants [Текст] / E. D. Lumer and B. Faieta // Proceedings of the third international conference on Simulation of adaptive behavior: from animals to animats 3: from animals to animats 3. – MIT Press, 1994.
3. Tsai C.F. ACODF: a novel data clustering approach for data mining in large databases [Текст] / C. F. Tsai, et al. // Journal of Systems and Software. – 2004. – Vol. 73, Is. 1. – pp. 133-145.
4. Handl J. Improved ant-based clustering and sorting in a document retrieval interface [Текст] / J. Handl and B. Meyer // Parallel Problem Solving from Nature—PPSN VII. Springer Berlin Heidelberg, 2002. – pp. 913-923.
5. Hone K.M. Homogeneous ants for web document similarity modeling and categorization [Текст] / K. M. Hone, et al. // Ant Algorithms, Springer Berlin Heidelberg, 2002. – pp. 256-261.
6. Горносталь А.Н. Розробка аудіовізуальної системи розпізнавання мови [Текст] / А.Н.Горносталь, Я.Ю.Дорогий // Scientific Journal «Science Rise». – 2017. - №12(41). – с. 42-47.
7. Neti C. Audio-visual speech recognition [Текст] : Final Workshop 2000 Report, Center for Language and Speech Processing / C. Neti, G. Potamianos, J. Luetttin et al. // The Johns Hopkins University, Baltimore. – 2000.