УДК 681.3.06

Т. В. Грищук, к.т.н.

ОТРИМАННЯ ХАРАКТЕРИСТИЧНОЇ ОБСЕРВАЦІЇ ПРИХОВАНОЇ МАРКОВСЬКОЇ МОДЕЛІ

У статті розглядається процес отримання на основі натренованої прихованої моделі характеристичної обсервації голосової команди. Ця інформація може використовуватися для підвищення ефективності процесу розпізнавання мови. Введено поняття оціночної функції, знаходження максимуму якої дає характеристичну обсервацію. Подано рекомендації з вибору початкового наближення.

Ключові слова: *характеристична обсервація, прихована марковська модель, розпізнавання мови, оціночна функція.*

Зі стрімким розвитком комп’ютерної техніки все більшу популярність здобувають голосові інтерфейси, які дозволяють зробити процес спілкування з комп’ютером більш природнім для користувача і значно збільшити коло споживачів програмного забезпечення з можливістю голосового спілкування. Потрібно зазначити, що подібне програмне забезпечення з розряду прикладного поступово переходить у розряд системного програмного забезпечення, основними вимогами до якого є висока швидкодія та невеликий обсяг пам’яті для зберігання [1].

Основною задачею, що вирішується в системах мовного спілкування, є задача визначення імовірності приналежності вхідного сигналу до однієї з команд голосового інтерфейсу. На рис. 1 показано спрощену схему автоматичного класифікатора.

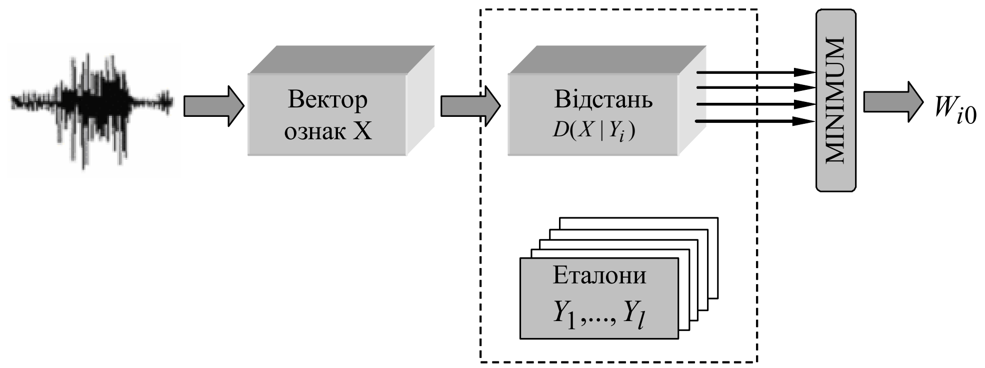


Рис. 1. Структура автоматичного класифікатора

Класичним та загальновизнаним на сьогодні підходом до класифікації мовних команд є підхід, що ґрунтується на використанні математичного апарату прихованих марковських мереж (ПММ). Цей підхід передбачає зберігання в пам’яті не самих еталонів мовних сигналів, а наборів параметрів прихованих марковських моделей, які за своєю суттю є “генераторами” векторів спостереження сигналу (обсервацій). Неформально, обсервація – це послідовність векторів, координати яких відповідають часовим та частотним ознакам мовного сигналу у кожному вікні аналізу (приблизно 256 відліків). Задача класифікації зводиться до задачі визначення ймовірностей генерації вхідної обсервації моделями, що відповідають фразами вхідної командної мови [2, 3].

Першим етапом побудови системи розпізнавання є етап навчання прихованих марковських моделей, тобто розрахунок параметрів закону розподілу в кожній вершині моделі. У цій статті розглядається зворотній до навчання процес – отримання на основі навченої моделі еталонної (тобто характеристичної) обсервації голосової команди. Отримання характеристичної обсервації моделі дозволило б розв’язати не тільки задачу спрощення обчислювальної складності процесу класифікації мовних команд, але й надати додаткові можливості для розв’язання задач адаптації та побудови моделей із шумом.

Постановка задачі

У статті розглядається задача зворотня до задачі побудови ПММ за алгоритмом Баума-Велха [3].

> Дана марковська модель. Необхідно знайти обсервацію, яка дає максимальну оцінку в даній моделі по Бауму-Велху.

Більш формально, нехай M – прихована марковська модель. Необхідно знайти обсервацію x', таку що:

$$x' = \arg\max_{x} P\left(x \mid M\right). \tag{1}$$

Метод знаходження характеристичної обсервації

Для початку введемо необхідні позначення і дамо визначення основних понять, які будуть використані в цій статті.

Вектори обсервацій будемо позначати так:

$$\overline{x} = x_1, x_2, ..., x_n,$$

де n – кількість координат (ознак) векторів обсервацій.

Визначення 1 (Forward-імовірності):

$$\alpha_1(1) = 1$$

$$\alpha_j(1) = a_{1j}, \quad \text{для j = 2,...,N – 1}$$

$$\alpha_j(t) = \sum_{i=2}^{N-1} \alpha_i(t-1)\, G_i(x_{t-1}) \cdot a_{ij}, \quad \text{для j = 2,...,N – 1, t = 2,...,} T_\chi$$

$$\alpha_N(T_\chi) = \sum_{j=2}^{N-1} \alpha_j(T_\chi) G_j(x_{T_\chi}) \cdot a_{jN}.$$

де a_{ij} – транзитна імовірність переходу з вершини i у вершину j;

N – кількість вершин моделі;

$G_j(\overline{x})$ – функція розподілу, асоційована з емісійною вершиною j;

T_χ – математичне сподівання довжин обсервацій, що генеруються моделлю M:

$$T_\chi = \sum_{T=1}^{\infty} P(\chi(T) = N) \cdot T \text{ або } T_\chi = \sum_{T=1}^{\infty} \alpha_N(T) \cdot T.$$

Визначення 2 (Backward-імовірності):

$$\beta_N(T_\chi) = 1$$

$$\beta_j(T_\chi) = a_{jN}, \text{ для j = 2, ... ,N –1}$$

$$\beta_i(t) = \sum_{i=2}^{N-1} a_{ij} G_j(x_{t+1}) \beta_j(t+1), \text{ для i = 2,...,N – 1, t = 1,...,} T_\chi \text{ -1,}$$

$$\beta_1(1) = \sum_{j=2}^{N-1} a_{1j} G_j(x_1) \cdot \beta_j(1).$$

Зараз ми маємо співвідношення, які висвітлюють основну ідею розв’язання задачі:

$$P\left(x \mid M\right) = \sum_{j=2}^{N-1} \alpha_j(t) G_j(x_t) \cdot \beta_j(t). \tag{2}$$

Співвідношення (2) справедливе для довільного t = 1, ... , T_χ .

Таким чином, ми можемо поетапно покращити оцінку Баума-Велха шляхом знаходження максимуму:

$$\arg\max_{x_t} P\left(x_t \mid M\right).$$

Для $t = 1$ отримаємо: $P\left(x \mid M\right) = \sum_{j=2}^{N-1} \alpha_j(1) G_j(x_1) \cdot \beta_j(1)$, а для $t = T_\chi$ отримаємо: $P\left(x \mid M\right) = \sum_{j=2}^{N-1} \alpha_j(T_\chi) G_j(T_\chi) \cdot a_{jN}$.

Для оцінки якості оцінки обсервації $x = x_1, x_2, ..., x_{T\chi}$ приймемо:

$$e(j,t) = \alpha_j(t) \cdot \beta_j(t),$$

де *j = 2, ..., N – 1, t = 1,..., T_χ* .

Оцінка Баума-Велха визначає найбільш імовірну вершину, в якій перебуває вектор обсервації в момент часу t [3]. На основі цієї оцінки ми введемо оціночну функцію у вигляді суми законів розподілу всіх емісійних вершин моделі:

$$E(x_t,t) = e(2,t) G_2(x_t) + ... + e(j,t) G_j(x_t) +$$
$$... + e(N-1,t) G_{N-1}(x_t)$$

Отже, для довільного *t*, *t = 1, 2, ...,* T_χ ми маємо оціночну функцію $E(x_t,t)$ для оптимізації оцінки по x_t .

На верхньому рівні деталізації алгоритм отримання характеристичної обсервації має вигляд:

```
Alg  Get Observation
     Get Init Observation                ‘ отримати характеристичний час Tχ,
                                           ініціальну обсервацію x1,…, xTχ
                                           і її оцінку Pnew по Бауму-Велху
     Repeat
         P ←— Pnew
         Get Backward Probabilities      ‘ отримати всі βj (t)
         Get First Forward Probabilities ‘ отримати αj (1) = a1j

         Loop ‘ по t від 1 до Tχ
             Get Estimation Function     ‘ отримати E (xt,t)
             Get New Observation Vector  ‘ отримати x’t
             Get New Forward Probabilities ‘ отримати αj (t+1), якщо t < Tχ
                                           або αN (Tχ), якщо t = Tχ
         end
         Pnew ←— αN (Tχ)
     until  | P – Pnew | < ε
```

Зупинимось детальніше на процедурах, що складають даний алгоритм.

Get Init Observation

Тут ми отримуємо характеристичний час, початкову характеристичну обсервацію і її оцінку за Баумом-Велхом.

Get Backward Probabilities

Розраховуємо для даної моделі величини $\beta_j(t)$.

Get First Forward Probabilities

Знаходимо для всіх емісійних вершин величину $\alpha_j(1) = a_{1j}$.

Get Estimation Function

В цій процедурі потрібно отримати оціночну функцію $E(x_t, t)$ для розрахунку вектора x_t. Ця оціночна функція задається за допомогою коефіцієнтів:

$$\alpha_j(t) \cdot \beta_j(t) = e(j,t),$$

$$E(x_t, t) = \sum_{j=2}^{N-1} e(j,t)\, G_j(x_t).$$

Get New Observation Vector

Призначенням цієї процедури є знаходження x_t',

$$x_t' = \arg\max_{x_t} E(x_t, t).$$

Знаходження максимуму такого типу функцій ми розглянемо нижче.

Get New Forward Probabilities

Ця процедура знаходить нові forward-імовірності з урахуванням того, що знайдено новий вектор обсервації x_t':

$$\alpha_j(t+1) = \sum_{i=2}^{N-1} \alpha_i(t)\, G_i(x_t') \cdot a_{ij}$$

для $j = 2, ..., N-1$, $t = 1, ..., T_\chi -1$;

$$\alpha_N(T_\chi) = \sum_{i=2}^{N-1} \alpha_i(T_\chi)\, G_i(T_\chi) \cdot a_{iN}.$$

Зазначимо, що forward-імовірності на цьому кроці циклу, що відповідає моменту часу t, розраховуються тільки для моменту часу $t+1$ (для $t < T_\chi$).

Якщо крок циклу відповідає моменту часу $t = T_\chi$, то розраховується тільки $\alpha_N(T_\chi)$.

Знаходження максимуму функції при різних положеннях складових

Для знаходження векторів обсервацій необхідно знаходити максимум функцій вигляду:

$$F(x) = \gamma_2\, G_2(x) + ... + \gamma_j\, G_j(x) + ... + \gamma_{N-1}\, G_{N-1}(x),$$

де $G_j(x) = G_j(x_1, ..., x_n)$ – функція розподілу, асоційована з вершиною j;

$\gamma_j = \alpha_j(t) \cdot \beta_j(t)$.

Нам необхідно знайти максимум функції, яка є сумою декількох нормальних законів розподілу і має декілька локальних екстремумів. На рис. 2 показано приклад такої функції.

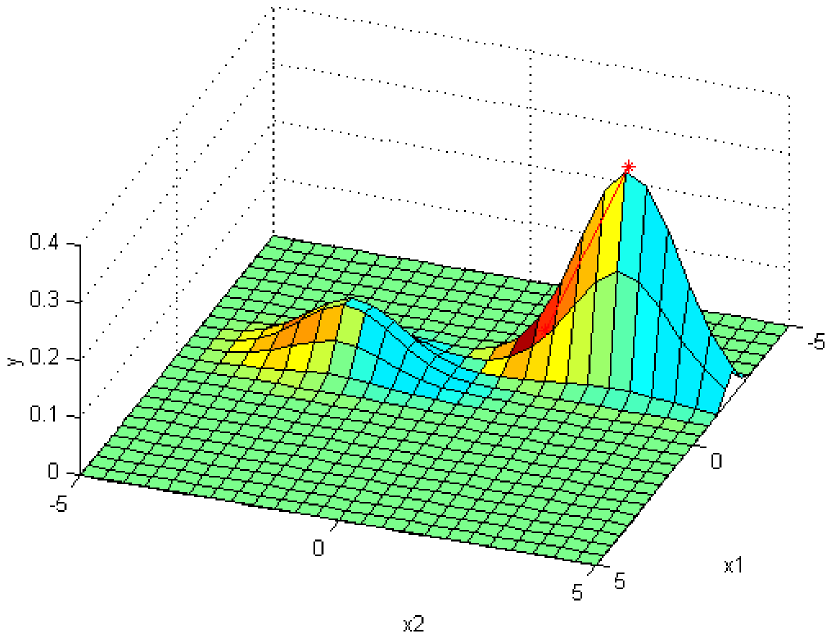


Рис. 2. Приклад функції розподілу моделі для 2-х вершин і 2-х координат

Особливістю даної задачі є те, що глобальний максимум досягається не завжди. Тому в якості початкового наближення виберемо обсервацію x^o:

$$x^o = \frac{1}{W}\left(W_2 \cdot c_2 + \ldots + W_{N-1} \cdot c_{N-1}\right),$$

де $W_j = F(c_j) = \sum_{i=2}^{N-1} \gamma_i \, G_i(c_{j1},\ldots,c_{jn}),$

$$W = \sum_{j=2}^{N-1} W_j.$$

Розглянемо систему рівнянь:

$$\begin{cases} \dfrac{\partial E(x_1,\ldots,x_n,t)}{\partial x_1} = 0 \\ ------------ \\ \dfrac{\partial E(x_1,\ldots,x_n,t)}{\partial x_j} = 0 \\ ------------ \\ \dfrac{\partial E(x_1,\ldots,x_n,t)}{\partial x_n} = 0 \end{cases}.$$

Виконавши необхідні перетворення, отримаємо таку систему рівнянь для знаходження максимум функції:

$$x_i = \frac{1}{\sum\limits_{j=2}^{N-1} e_{ji}(x)} \cdot \sum_{j=2}^{N-1} e_{ji}(x) \cdot c_{ji} \ ,$$

де $i = 1, 2, ..., n$.

$$e_{ji}(x) = \frac{\gamma_j}{\sigma_{ji}^2} \cdot G_j(x) \, ,$$

де σ_{ji} – середньоквадратичне відхилення для *i*-ї координати функції розподілу $G_j(x)$, що асоційована з вершиною *j*.

c_{ji} – математичне сподівання для *i*-ї координати функції розподілу $G_j(x)$.

Вирішити дану систему рівнянь можна будь-яким ітераційним методом чисельної математики.

Експериментальні дослідження

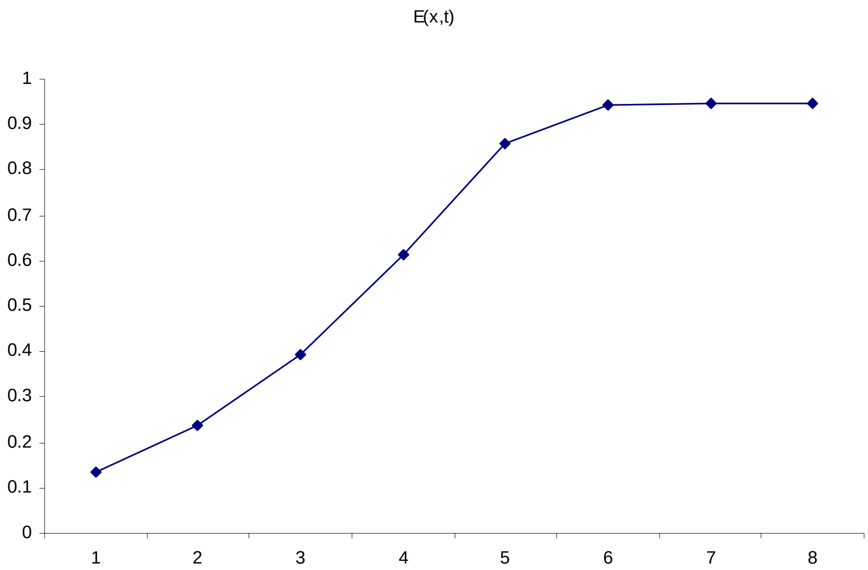
Нами проводились експериментальні дослідження розробленого алгоритму. У якості ілюстрації наведемо приклад. У табл. представлено параметри прихованої марковської моделі з трьома емісійними вершинами, для якої потрібно було знайти характеристичну обсервацію з 2-х координат довжиною $T_\chi = 9$.

Таблиця

Параметри ПММ

Математичні сподівання	СКВ	Транзитні імовірності
0 0	0 0	0 1 0 0 0
3.56 1.2	0.6275 0.2115	0 0.6 0.4 0 0
3.5 1.23	0.6169 0.2168	0 0 0.5 0.5 0
2.499 1.189	0.4405 0.2096	0 0 0 0.8 0.2
0 0	0 0	

На рис. 3 показана динаміка зростання оціночної функції і значення результуючої характеристичної обсервації.



Результат

3.5600 1.2000
3.5137 1.2227
2.7548 1.2039
2.5408 1.1919
2.5055 1.1895
2.5000 1.1891
2.4991 1.1890
2.4990 1.1890
2.4990 1.1890

Рис. 3. Значення цільової функції на кожній ітерації

Висновки

Основний науковий результат, викладений у цій статті, полягає в розробці методу знаходження характеристичної обсервації прихованої марковської моделі. Проведені експериментальні дослідження підтвердили адекватність розробленого методу. Подальші дослідження цього питання можуть полягати в дослідженні можливості використання характеристичних обсервацій для розв'язання задач підвищення швидкодії процесу розпізнавання, адаптації та побудови моделей з шумом.

СПИСОК ЛІТЕРАТУРИ

1. Борисов М. ПО для распознавания речи // Мир ПК. – 1999. – №3. – С. 124 – 129.
2. Normandin Y. Hidden Markov Models, Maximum Mutual Information Training, and the Speech Recognition Problem. – Montreal, Canada: Punlished by the Department of Electrical Engineering, McCall University, 1991. – 266 p.
3. Young S., Kershaw D., Odell J., Ollason D., Valthev V., Woodland P. The HTK Book (for HTK Version 3.0). – 2000. – 278 p.

Грищук Тетяна Вікторівна – старший викладач кафедри комп'ютерних систем управління.
Вінницький національний технічний університет