

СИСТЕМНА АДАПТАЦІЯ НЕЙРОННИХ МЕРЕЖ

УДК 004.75:004.272.26:004.032.26:519.1

КАТЕРИНИЧ Лариса Олександрівна

к.ф.-м.н., асистент кафедри інформаційних систем Київського національного університету імені Тараса Шевченка.

Наукові інтереси: нейронні мережі, методи об'єктно-орієнтованого програмування, веб-програмування.

e-mail: katerinich@rambler.ru

ВСТУП

Глобальна адаптація нейронної мережі (НМ) на предмет ефективного функціонування достатньо складна задача, оскільки вимагає одночасного розв'язання великого класу задач [1,2]. Тому більш ефективним способом є розв'язання більш локальних задач адаптації.

ПОСТАНОВКА ЗАДАЧІ

Необхідно розглянути системну адаптацію НМ, оскільки розв'язання і оптимізація роботи на цьому етапі сприяє більш ефективній роботі усієї НМ у подальшому.

Мета роботи полягає у тому, щоб показати, що для підвищення ефективності функціонування НМ при різних умовах експлуатації необхідне витрачання ресурсу системного забезпечення для роботи НМ та відповідно його адаптація.

ОСНОВНИЙ МАТЕРІАЛ

Нехай P - умови експлуатації НМ, U - керуючий інструмент НМ, який може змінюватися в процесі її роботи, а

$$q = q(P, U) \quad (1)$$

- критерій функціонування НМ, який характеризує ефективність її роботи в умовах P при управлінні U . Вираз (1) задається алгоритмічно, тобто у вигляді правила оцінки ефективності q реальної (а не модельної) НМ при різних P і U . Це обставина виключає можливість обчислення інтегрального критерію

$$Q(U) = \int q(P, U) p(P) dP. \quad (2)$$

необхідного для синтезу оптимального управління U^* , що розв'язує задачу

$$Q(U) \rightarrow \min_{U \in S} \Rightarrow U^*. \quad (3)$$

Тут $p(P)$ - статистичні властивості умов експлуатації НМ (щільність розподілу цих умов), які зазвичай невідомі. Тому завдання підтримки НМ в оптимальному стані слід вирішувати методами адаптації, тобто змінювати керований фактор U , керуючись тільки спостереженнями значень критерію (1).

Розглянемо конкретну задачу – задачу адаптивної сегментації програмного забезпечення НМ.

Нехай $c_l = c(l=1, \dots, m)$ де c - частина об'єму пам'яті, що виділяється для інформації, що надходить із зовнішньої пам'яті, і обмеження мають детермінований характер. Імовірнісні властивості середовища X тут визначають звертання до переходу від одного блоку інформації до іншого. Ці властивості в простому випадку визначаються стохастичною матрицею

$$P = \|p_{lk}\|_{m \times m}. \quad (4)$$

де p_{lk} - імовірність переходу від l -го блоку інформації до k -го. Маючи таку інформацію будемо вважати, що, тобто матриця зв'язків блоків інформації збігається зі стохастичною матрицею (4), і немає необхідності робити осереднення, так як воно вже зроблено при визначенні матриці (4). Тоді при відомій матриці переходів (4) задача оптимальної сегментації

приймає детермінований характер. Вирішення цієї завдання методами адаптації буде розглянуто нижче.

АДАПТАЦІЯ РОЗТАШУВАННЯ БЛОКІВ ІНФОРМАЦІЇ НА ЖОРСТОКОМУ ДИСКУ

Задача полягає у мінімізації часу необхідного для механічного руху головок, що считують від одного циліндру до іншого. При проектуванні експертної системи Н-Гомеопат [3,4,5], на етапі навчання нейронної мережі на вибірці більше 5 тисячі елементів спостерігається досить активне звернення до жорсткого диску, що в свою чергу певним чином впливає на швидкість навчання НМ. Таким чином актуальність задачі очевидна.

В поставленій задачі мінімізується середній час руху магнітних головок запису-зчитування. Нехай m - число циліндрів магнітного барабана; a_l - інформація, що розташовується на l -му циліндрі ($a_l < c$, де c - інформаційний обсяг циліндра); p_l - ймовірність звернення до інформації, розташованої на l -му циліндрі. Середній час переміщення головки від одного циліндра до іншого (5)

$$Q = \sum_{l,k=1}^m p_l p_k \varphi(l,k) \quad (5)$$

де $\varphi(l,k)$ - час переміщення головки від l -го циліндра до k -го (наприклад, в простому випадку $\varphi(l,k) = w|l-k|$, де w - коефіцієнт розмірності).

Позначимо літерою U варіант розташування інформації, проутвориться з вихідного за допомогою перестановки:

$$\begin{pmatrix} 1 & 1 & \dots & m \\ u_1 & u_2 & \dots & u_m \end{pmatrix} \quad (6)$$

де $u_i \in \{1, \dots, m\}$ і $u_l \neq u_k$ ($l \neq k$).

Вираз (6) означає, що інформація a_l з l -го циліндра переводиться на k -й. Таким чином, адаптація як процес керування в даному випадку однозначно визначається наступним цілочисельним вектором:

$$U = (u_1, \dots, u_m) \quad (7)$$

з різними компонентами з відрізка натурального ряду. Задачу синтезу оптимального розташування інформації на магнітних циліндрах тепер можна сформулювати як задачу цілочисельний мінімізації:

$$Q(U, P) = \sum_{i,k=1}^m p_{u_i} p_{u_k} \varphi(u_i, u_k) \rightarrow \min_{U \in S} \Rightarrow U^* \quad (8)$$

де

$$P = (p_1, \dots, p_m) \quad (9)$$

Вектор ймовірностей звертання до блоків інформації a_1, \dots, a_m ($\sum_{l=1}^m p_l = 1$)

$$S: \begin{cases} u_l \in \{1, \dots, m\} & (l=1, \dots, m) \\ u_l \neq u_k & \text{при } l \neq k \end{cases} \quad (10)$$

Очевидно, що оптимальне розташування U^* залежить від стану середовища – вектора P :

$$U^* = U^*(P) \quad (11)$$

При відомому векторі P задача адаптації, як видно, зводиться до задачі оптимізації (8), тобто є детермінованою. При невідомому або змінюваному P доводиться користуватися його оцінкою

$$P = (\hat{p}_1, \dots, \hat{p}_m) \quad (12)$$

на базі кінцевого часу спостереження за процесом роботи магнітного диска. Процес адаптації повинен бути побудований так, щоб послідовність одержуваних при цьому перестановок:

$$U_1, \dots, U_N \quad (13)$$

кожна з яких використовує останню оцінку (12), відбивала оптимальне розташування інформації на дисках за критерієм (8) у кожний конкретний момент часу.

Легко помітити, що отримана задача не є задачею о розрізі графа - тут нічого не «розрізається» і не агрегується. Ця задача ближче до так званої задачі про призначення.

Її можна сформулювати таким чином. Введемо двійкову змінну u_{lk} такого роду: $u_{lk} = 1$ позначає розташування l -го блоку інформації a_l на k -му циліндрі ($l, k = 1, \dots, m$) і $u_{lk} = 0$ в протилежному випадку. Тоді функціонал, що мінімізується (5) може бути записаний у наступній формі:

$$Q(U) = \sum_{i,j,l,k} u_{lk} u_{ij} p_i p_j \varphi(l,k) \quad (14)$$

При цьому повинні виконуватися наступні обмеження:

$$S: \begin{cases} \sum_{l=1}^m u_{lk} = 1 \\ \sum_{k=1}^m u_{lk} = 1 \\ u_{lk} \in \{0;1\} \quad (k, l = 1, \dots, m) \end{cases} \quad (15)$$

Перше з них виражає очевидну вимогу, щоб на кожному k -м циліндрі розташовувався тільки один інформаційний блок. Друге обмеження пов'язане з тим, що кожен l -й блок повинен розташовуватися тільки на одному циліндрі, а третє визначає бінарність шуканого вектора

$$U = (u_{11}, u_{12}, \dots, u_{mm}) \quad (16)$$

містить m^2 компонент.

Задача адаптації в цьому випадку має вигляд.

$$\sum_{i,j,l,k} u_{lk} u_{ij} p_i(X) p_l(X) \phi(l, k) \rightarrow \min_{U \in S} \Rightarrow \Rightarrow U^*(X) \quad (17)$$

В роботі було розглянуте одне із питань нейронних мереж, а саме системна адаптація з під задачею, що стосується адаптації розташування блоків інформації на жорсткому диску. Було показано, що задача зводиться до мінімізації середнього час руху магнітних головок запису-зчитування.

ЛІТЕРАТУРА:

1. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. – М: Горячая линия-Телеком, 2006. – 452 с.
2. Круглов В.В., Длин М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети: Учеб. пособие. – М.: Издательство физико-математической литературы, 2001. – 224 с.
3. Л. Катеринич, А. Проватар. Синтез нейронных сетей на основе информационных гранул //International Book Series "Information Science and Computing: Advanced Research in Artificial Intelligence". – Sofia, 2008. – V1. – P.179-182.
4. Катеринич Л.О. О некоторых формальных моделях разработки нейросетевых алгоритмов в системе Гомеопат //Компьютерная математика. – 2010. – Вып. 1. – С.102-109.
5. Проватарь А.И., Катеринич Л.А. Нейронечеткие модели диагностики в системе Н-Гомеопат //Проблеми програмування. – 2010. – Вип. 2-3. – С.636-641.
6. Bargiela, Andrzej and Pedrycz, Witold. Granular Computing: An introduction. – Kluwer Academic Publishers, 2003. – 5 p.
7. Jacek Leski. Systemy neuronowo-rozmyte. – Warszawa: Wydawnictwo Naukowo-Techniczne, 2008. – 686 s.