Златко Зографски

Електротехнички факултет

Универзитет "Кирил и Методиј"

Карпош II, 91000 Скопје

Proglašen za najbolji rad u komisiji.

ЕДНА НОВА МЕТОДА НА АВТОМАТНО УЧЕЊЕ И НЕЈЗИНА ПРИМЕНА ВО МОДЕЛИРАЊЕ И СИМУЛАЦИЈА НА ДИНАМИЧКИ СИСТЕМИ

Резиме - Во трудот е изложена една нова метода на автоматно учење и опишана е нејзината примена во доменот на сложени динамички системи. Со алгоритмот се конструира репрезентација на динамичките равенки на системот како нов вид на асоцијативна меморија заснована на локална интерполациска схема и брз алгоритам за пребарување. Експериментите во примена на методата при учење на динамичкиот модел за системот на инвертирано нишало и при употребувањето на здобиеното знаење во предикција на еволуцијата на системот ја демонстрираат ефикасноста на алгоритмот и ги верифицираат теоретски изведените поставки за неговата временска и просторна комплексност.

A NOVEL MACHINE LEARNING METHOD AND ITS USE IN THE MODELLING AND SIMULATION OF DYNAMICAL SYSTEMS

Abstract - In this paper a novel method of machine learning is presented, and its application in the domain of complex dynamical systems is described. The learning algorithm constructs a representation of the model of system's dynamics in a new form of accosiative memory which features a local interpolation scheme for approximating nonlinear mappings and fast geometric searching algorithms. Experiments with the acquisition of the dynamics model of the well-known pole balancing system demonstrate the efficiency of the learning method and verify theoretical results with regard to its time and space complexity.

1. УВОД

Со трудот на Michie и Chambers [6] започнува применувањето на методите на вештачка интелигенција во управувањето со физички процеси. Методите се интересни затоа што класичната теорија на управување е понекогаш неприменлива за определен процес, поради непознавањето на детален модел, неговите параметри и влијанието од околината. Постојат мислења дека методологијата на автоматско учење од примери може да биде полезна во овие случаи; таквиот пристап резултирал со некои интересни резултати во врска со невроморфно ([2], [3], [10]) и индуктивно учење [8].

Сличен е и проблемот на предикција на еволуцијата на системот при зададена почетна состојба. Клучен концепт, што ги поврзува двата проблеми, е мемориски модел на динамиката на системот, како апроксимација на егзактните равенки (кои често пати е практично невозможно да се добијат). Доволно точен модел, здобиен преку процес на учење од примери на состојбите на системот и неговите одзиви, би можел да се употреби за определување на следната состојба на системот (доколку се познава сегашната состојба и влезниот сигнал), и со тоа ефективно да се изврши интеграција на неговите равенки на движење како нивната математичка форма е непозната.

Во овој труд прво даваме кус преглед на една нова метода за автоматско учење ([5], [6], [7]). Потоа ја опишуваме нејзината примена во проблемот на аквизиција на динамиката на инвертирано нишало на подвижна количка. Се дискутираат перформансите на методата, а се даваат и сугестии за понатамошно истражување.

2. МОДЕЛ ЗА УЧЕЊЕ НА НЕЛИНЕАРНИ ДИНАМИЧКИ ПРЕСЛИКУВАЊА

Разгледуваме динамички систем:

$$dy/dt = F(y) \tag{1}$$

каде секој $y = (y_1, y_2, ...)$ претставува вектор на состојба на системот и може да се набљудува како точка во соодветен фазен простор S ($S \subseteq R^n$). Така векторското поле $F(y)$ е нелинеарен оператор врз точки од S (ако $S \subseteq R^n$, тогаш $F: R^n \rightarrow R^n$); од y_t во момент t може да се определи y_{t+1} во $t + dt$:

$$y_{t+1} = y_t + F(y) \cdot \Delta t \tag{2}$$

Често пати аналитичкиот израз за $F(y)$ не е познат; меѓутоа, позната е низа на измерени величини на векторот на состојба и неговите изводи $\{(dy/dt)\}_t = f_t$. Проблемот на предикција на однесувањето на системот F така се сведува на проблем на мултидимензионална интерполација, и тоа кон конструкција на функција $\Phi: R^n \rightarrow R^n$, така што $\Phi(y_t) = f_t$ за сите t во секвенцата на набљудувања.

Проблемот може да се формулира и како парадигма на автоматно учење: од дадено множество на примери L на релации влез/излез $\{(y_t, f_t)\}$, $t = 1,2,....N$ (каде y_t и f_t се реални n-вектори, да се конструира меморискн модел M на нелинеарното пресликување Φ кое ги вклучува паровите од множеството на примери. Откако моделот M е конструиран, од него се добива излезниот вектор f_t при индексирање со соодветниот влезен вектор y_t. Во овој контекст, генерализацијата се однесува на способност да се произведе "прифатлив" одзив φ_τ и на влезови y_τ кои не биле презентирани во учењето; овој поим може да се прецизира преку дефиницијата на нормализираната грешка ε:

$$\varepsilon = \left(\sum_{i=1}^{P_1} \|\varphi_i - \varphi_n\|^2 / \sum_{i=1}^{P_1} \|\varphi_i - \langle\varphi_n\rangle\|^2\right)^{1/2} \tag{3}$$

каде (y_i, φ_i), $i=1,2,...,P_1$ е колекција на тест-примери (неупотребени во учењето). За секој вектор y_i, излезниот вектор φ_i што се добива од M се споредува со очекуваниот (точен) вектор φ_n; $\langle\varphi\rangle$ е средниот излезен вектор (над сите тест-вектори).

3. АСОЦИЈАТИВНА ИНТЕРПОЛИРАЧКА МЕМОРИЈА: МОДЕЛ И АЛГОРИТАМ ЗА УЧЕЊЕ

Репрезентацијата на множеството на примери во вид на множество на влезно-излезни парови на вектори може да се смета и како нов тип на асоцијативна меморија [12], наречена AICAM. Излезниот вектор за пример од пресликувањето може да се добие со пребарување и делумно поклопување помеѓу влезниот вектор (употребен како клуч за адресирање) и меморираните примери. Задачата на учење се состои во алгоритмичка конструкција на оваа мемориска репрезентација за бараното пресликување од множеството на примери. Подетален опис на методата и алгоритмот, скицирани подолу, може да се најде во [11] и [12].

За множество на N точки $\{y\}$ со придружени вредности на функцијата $\{f\}$, дефинираме функција од две променливи (експозицијата на методата е за 2-D случај; генерализацијата за повеќедимензионални случаи е едноставна) $F \in C(R)$ таква што $F(y_i) = f_i$ за $i = 1,2,...,N$. Функцијата F е дефинирана со сумата:

$$F(y) = \sum_{k=1}^{N} W_k(y) \cdot Q_k(y) \,/ \sum_{i=1}^{N} W_i(y) \tag{4}$$

каде Q_k се функции за кои важи $Q_k(y_k) = f_k$, додека тежинските коефициенти W_k се определуваат со функцијата на инверзно растојание $W_k(y) = (R_w - d_k)_+^2 / R_w^2 d_k^2$, каде $(R_w - d_k)_+ = (R_w - d_k)$ ако $d_k < R_w$ и $(R_w - d_k)_+ = 0$ аго $d_k \leq R_w$, d_k е Евклидското растојание помеѓу интерполирачката точка (y) и меморирана точка (y_k), и R_w е радиусот на влијание околу точката (y_k). Јасно е оти податокот f_k влијае само врз интерполирачки вредности во радиус R_w. Функциите Q_k се дадени со изразот:

$$Q_k(y)=\alpha_{k1}(y_1-y_{k1})^2+\alpha_{k2}(y_1-y_{k1})(y_2-y_{k2})+\alpha_{k3}(y_2-y_{k2})^2+\alpha_{k4}(y_1-y_{k1})+\alpha_{k5}(y_2-y_{k2})+f_k \qquad (5)$$

во кој коефициентите α_{kj} се определени со минимизација на изразот (6) според методата на најмали квадрати:

$$\sum_{i=1,\, i\neq k}^{N} \omega_i(y_k)\left[\alpha_{k1}(y_{i1}-x_{k1})^2+\ldots+f_k-f_i\right]^2 \qquad (6)$$

во кој

$$\omega_i(y)=(R_q-d_i)^2/R_q^2\, d_i^2 \qquad (7)$$

R_q е радиус на влијание околу точката (y_i), додека d_i е растојанието помеѓу (y_i) и интерполираната точка (y). Од (6) следи дека во определувањето на Q_k придонесуваат само оние точки чии радиуси го вклучуваат (y_k), дефинирајќи ја локално. Значи, определувањето на функциите на јазли Q_k поаѓа од определувањето на радиусите R_q следено од формирањето и решавањето на системот (6). Радиусите R_q и R_w се бираат така да опфаќаат N_q и N_w точки, респективно, за фиксни, емпириски определени вредности на N_q и N_w.

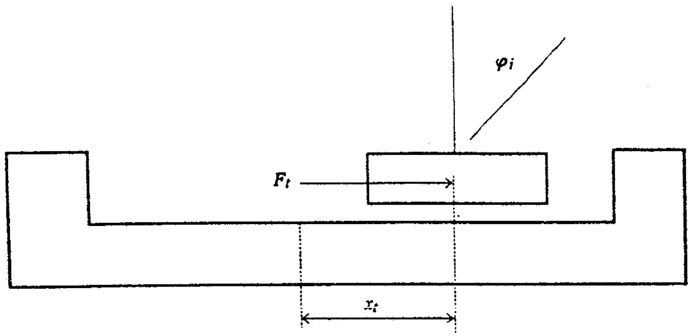
Значи, горе опишаниот алгоритам на учење врз основа на модифицираната Shepard-ова метода [11], се состои од три фази:

1. Фаза на "меморирање" на примерите, во која решетка од $O(N)$ просторни клетки се поставва на множеството од N податочни точки (примери), и секоја точка се придружува кон определена клетка.
2. Пребарување на околината на секоја податочна точка (y_k) за да се определи подредена низа од N_w најблиски соседи на секоја од нив, и според тоа радиусот на влијание R_w.
3. Определување на коефициентите α_{kj} за функциите Q_k според (6), со вклучување на најблиските N_q точки на секоја податочна точка (y_k) и решавање на соодветната регресиона матрица со декомпозиција по сингуларните вредности.

Што се однесува до просторната комплексност на методата, за 2-D случај се потребни $3N$ локации за сместување на (y_i) и f_i. За коефициентите α_{kj} за функциите на јазли Q_k и радиусите R_w за секоја податочна точка потребни се уште $6N$ локации. Кон ова треба да се додадат и $4N/3$ локации за сместување на решетчестата податочна структура. Така алгоритмот (фазите на конструкција и евалуација) бара $O(N)$ мемориски локации. Во [11] се покажува дека средното време на конструкцијата е линеарно ($O(N)$), додека на евалуацијата е константно ($O(1)$). Во [11] е исто така покажано дека методата лесно се обопштува за n-димензионален проблем.

4. АКВИЗИЦИЈА НА ДИНАМИЧКИОТ МОДЕЛ НА ИНВЕРТИРАНОТО НИШАЛО

Проблемот на балансирање на инвертирано нишало досега неколку пати бил разгледуван во врска со алгортми што учат закони на управување на динамички системи ([2], [3], [6], [8]). Системот е пример на инхерентно нестабилен мултидимензионален динамички систем, какви што се сретнуваат при управување со двоножно одење (како и ориентација на сателит). Системот се состои од нишало прицврстено за количка што може да се движи долж шина (Сл. 1); движењето на нишалото е ограничено на вертикалната рамнина. Задачата е да се балансира нишалото во определени граници, движејќи ја количката долж шината со аплицирање на секвенца на сили со фиксна величина во дискретни моменти.



Сл. 1. Системот на инвертирано нишало

Оваа задача на учење на стратегија на управување е сложена поради непостоење на *a priori* информација за динамиката на системот. Единствената информација е векторот на состојба во момент t:

x_t ...... хоризонталната позиција на количката, [m]

$\dot{x}_t$ ...... хоризонталната брзина на количката, [m/s]

φ_t ...... аголот меѓу нишалото и вертикалата, [rad]

$\dot{\varphi}_t$ ...... аглова брзина на нишалото, [rad/s]

и "сигналот на неуспех" r кој добива вредност 1 ако аголот φ_t ја надмине вредноста $\pm 12^o$ или количката излезе надвор од лимитот на шината.

Употребени се равенки на движење и вредности на параметрите како во [1] поради полесна споредба на резултатите:

$$\ddot{\varphi}_t = \frac{g\sin\varphi_t + \cos\varphi_t\left[\,-F_t - m_p l\,\dot{\varphi}_t^2 \sin\varphi_t\right] / (m_c+m_p)}{l\left(4/3 - m_p\cos^2\varphi_t / (m_c+m_p)\right)} \tag{8}$$

$$\ddot{x}_t = \left[F_t + m_p l\,(\dot{\varphi}_t^2\sin\varphi_t - \ddot{\varphi}_t\cos\varphi_t)\right] / (m_c+m_p) \tag{9}$$

Системот е симулиран со интегрирање на равенките на движење (4) и (5) со Euler-овата метода со чекор $\Delta\tau = 0.02$c, при што управувачката сила во секој чекор е определена со правилото добиено со индуктивно учење како во [8].

Задачата во експериментот е да се научи од примери модел на пресликувањето

$$\ddot{\varphi}_t = f_1(\varphi_t, \dot{\varphi}_t, F) \tag{10}$$

$$\ddot{x}_t = f_2(\varphi_t, \dot{\varphi}_t, F) \tag{11}$$

Независните променливи φ_t, $\dot{\varphi}_t$, F се избрани поради физичката природа на проблемот, а и поради погодност. Јасно, F, изведена според правило од информацијата за векторот на состојба, всушност е зависна од ($x_t, \dot{x}_t, \varphi_t, \dot{\varphi}_t$). Меѓутоа, информацијата за сигналот F е директно достапна на системот и затоа е поприродно да се изрази $\ddot{\varphi}_t$ во зависност од F; (имплицитната) зависност на $\ddot{\varphi}_t$ и $\ddot{x}_t$ од целокупниот вектор на состојба притоа е осигурена преку зависноста од F. Со тоа се редуцира димензионалноста на проблемот на учење: векторот на примери има само 3 компоненти ($\varphi_t, \dot{\varphi}_t, F$) наместо 4 (доколку беше употребен векторот на состојба).

5. ЕКСПЕРИМЕНТАЛНИ РЕЗУЛТАТИ И ДИСКУСИЈА

Првиот експеримент се однесува на брзината и точноста на учењето. Пет множества на примери беа употребени, $L_1,\ldots,L_5$ со 200, 160, 120, 80, и 40 примери респективно. (Секој пример е чифт на влезно-излезни вектори $\{(\varphi_t, \dot{\varphi}_t, F); (\ddot{\varphi}_t, \ddot{x}_t)\}$. Овие примери беа добиени со регуларно земање на одбироци од излезната датотека (за првите 500 чекори) при симулацијата на системот со употребување на равенките на Anderson и правилото на управување на Sammut ([8]), и униформна редукција на бројот на примерите на соодветните вредности. Множествата $L_1,\ldots,L_5$ потоа се употребени како влезови на две верзии на алгоритмот за учење, наречени RL и RQ. Времето на учење T_L (CPU sec.) беше забележено за секој обид. По извршеното учење системот беше тестиран за точност на репродукција на првите 1000 чекори (еволуцијата на системот се одвива под истите услови како погоре), земајќи ја нормализираната грешка е дефинирана со (3) и пресметана за точно пресметаната секвенца $\{(\ddot{\varphi}_t^e, \ddot{x}_t^e)\}$ и за апроксимираната секвенца $\{(\ddot{\varphi}_t^a, \ddot{x}_t^a)\}$ респективно (за целиот интервал од 1000 чекори). Резултатите се прикажани на Слика 2 и тие го потврдуваат заклучокот за линеарната временска комплексност на овој алгоритам за учење во зависност од бројот на примерите (*vide supra*). Исто така се гледа дека варијантата RL е типично двојно поспора отколку варијантата RQ, но се одликува со поголема точност.

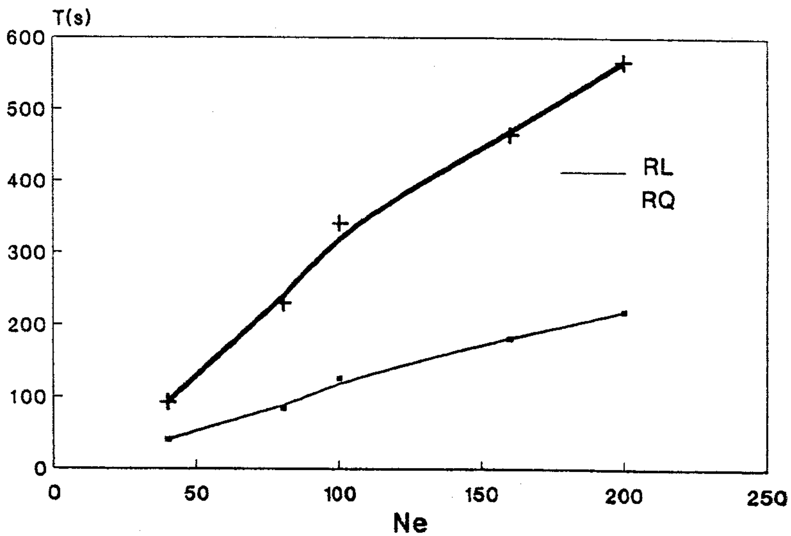


Fig. 2 Time complexity vs. sample size

Графикот на нормализираната грешка на генерализација при процесот на репродукција на првите 1000 чекори од еволуцијата на системот е прикажан на Сл. 3. Се гледа дека грешката на генерализација е останува практично константна (типично доста под 10 %) при редукција на бројот на примерите, се додека не се дојде до критична минимална вредност (40 примери во овој случај). Во тој момент грешката рапидно расте до 62 %. Сепак, резултатите укажуваат дека дури и толку "слаб" мемориски модел може да се употреби за симулација во проблемот на инвертираното нишало.

Интересно е да се повлечат некои споредби меѓу карактеристиките на оваа метода и резултати од други истражувања. Учењето на нелинеарни пресликувања е мошне истражуван проблем и во областа на вештачките неурални мрежи ([10], [5]). И покрај тоа што се развиени успешни методи, вклучувајќи и наши ([11], [13]), тие типично бараат големи мрежи и долго време на учење ([5]). Наспроти тоа, овој алгоритам во некои експерименти со учење на модел на динамика на робот

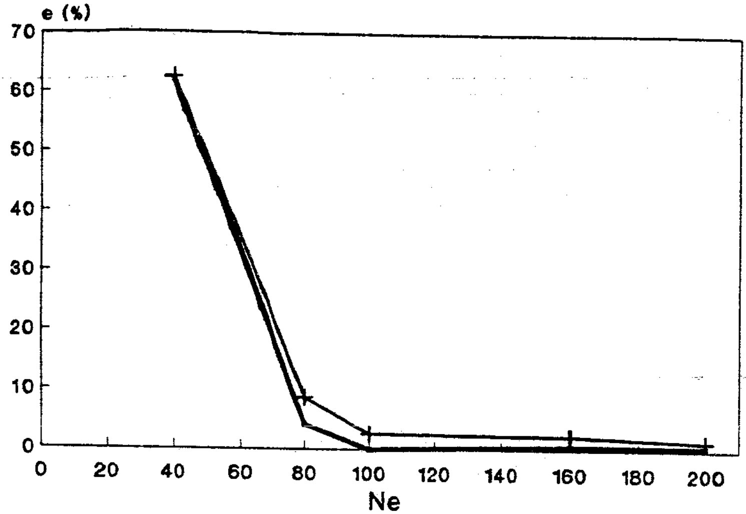


Fig. 3 Generalisation accuracy vs sample size

PUMA 560 се покажа подобар од најдобрите неурални методи во однос на брзината и точноста на учењето, како и способноста за генерализација ([11] и [12]).

Најпосле, интересно е да се одбележи и дека овој тип на асоцијативна меморија може да се структурира и во вид на *k-d* стебло, со мала модификација на алгоритмот за учење. Репрезентациите на сензоримоторното знаење во вид на стебло, конструирано со учење, го предизвикуваат вниманието и на други истражувачи [8]. Со таква реформулација, нашата метода може да се поврзе и спореди со методите за индуктивно симболичко учење [7].

6. ЗАКЛУЧОК

Во овој труд е презентирана нова метода за учење на нелинеарни пресликувања од примери. При учењето се формира репрезентација на примерите во вид на посебен тип на асоцијативна меморија (наречена AICAM), адресибилна според содржината. Методата беше применета за учење на модел на нелинеарен динамички систем и за употреба на научениот модел во симулација на еволуцијата на системот. Теориски изведените заклучоци за комплексноста на алгоритмот беа и експериментално потврдени. Посебно, беше покажано дека: а) временската комплексност на алгоритамот за учење е линеарна во зависност од бројот *N* на примерите за учење, б) мали множества на примери се доволни за формирање на многу точни модели кои можат ефективно да се употребат наместо експлицитните, математички модели во симулацијата.

Овие резултати водат кон заклучокот дека оваа метода, како и воопшто и други геометриски репрезентации и алгоритми слични на неа (или изведени од неа) можат да бидат ефективен алат за аквизиција на знаење во комплексни интелигентни системи. Наша цел во понатамошните истражувања ќе биде разработка на овие резултати во неколку насоки, а посебно во воспоставувањето на врски помеѓу неуралните, алгоритмичко-геометриските и симболичките методи на автоматно учење во контекстот на динамичките системи.

7. ЛИТЕРАТУРА

[1] C. Atkeson and D. Renkensmeyer: "Using content-addressable memories to control robots", In *Proceedings of the IEEE International Conference on Robotics and Automation, Scottsdale, AZ, 1989.*

[2] C. W. Anderson: "Strategy learning with multilevel connectionist representations", *Proceedings of the 4th International Workshop on Machine Learning, 1987.*

[3] A Barto et al.: "Neuronlike elements that can solve difficult control problems", *IEEE Trans. SMC, Vol. 13, pp. 835-846, July 1983.*

[4] J. D. Farmer and J. J. Sidorowich: "Predicting chaotic time series", *Physical Review Letters, Vol. 59, pp. 845-848, August 1987.*

[5] M. Kawato: "Computational schemes and neural network models for formation and control of multijoint arm trajectory", in *Proceedings of the NSF Workshop on Applications of Neural Networks to Robotics and Control, New Hampshire, 1989.*

[6] D. Michie D. and R. Chambers: "Boxes: An experiment in adaptive control", in *Machine Intelligence 2 (E. Dale and D. Michie, Eds.). Edinburgh: Oliver and Boyd, 1968.*

[7] J. R. Quinlan: "Induction of decision trees", *Machine Learning, Vol. 1, pp. 81-106, 1986.*

[8] C. Sammut: "Experimental results from an evaluation of algorithms that learn to control dynamical systems", *in Proc. 5th Int'l Machine Learning Workshop, San Francisco, CA, 1988.*

[9] C. Stanfill and D. Waltz: "Toward memory-based reasoning", *Comm. ACM, Vol. 29, pp. 1213-1228, December 1986.*

[10] B. Widrow and M. A Lehr: "Thirty years of adaptive neural networks", *Proceedings of the IEEE, Vol. 78, pp. 1415-1443, September 1990.*

[11] Z. Zografski: *Neuromorphic, Algorithmic, and Logical Models for the Automatic Synthesis of Robot Action, DSc Thesis, University of Ljubljana, Yugoslavia, 1989.*

[12] Z. Zografski: "A learning method for the construction of asociative interpolating memory models of robot inverse dynamics", in *Proc. Int.l Workshop on Sensorial Integration for Industrial Robots, Zaragoza, Spain, 1989, pp. 306-310.*

[13] Z. Zografski: "Efficient learning of robot dynamics in a neuromorphic system", *in Proc. Int'l Neural Networks Conf., Paris. Dodrecht: Kluwer Academic Publishers, 1990, pp. 217-220.*