

Krzysztof Najman

Uniwersytet Gdański

PROPOZYCJA ALGORYTMU SAMOUCZENIA SIĘ SIECI NEURONOWYCH TYPU GNG ZE ZMIENNYM KROKIEM UCZENIA

Streszczenie: Jednym z kluczowych parametrów procesu samouczenia się sieci neuronowych typu GNG jest szybkość zmiany pozycji w przestrzeni neuronu uczącego się i najbliższego połączonego z nim neuronu. Zależy ona od lokalnego błędu kwantyzacji i stałej nazywanej krokiem uczenia. Stała wartość kroku uczenia w szczególności niepotrzebnie zwalnia proces samouczenia się w początkowej jego fazie. W artykule proponuje się modyfikację algorytmu, wprowadzając zmienny krok uczenia oparty na liniowej funkcji iteracji między kolejnymi fazami wstawiania nowego neuronu do sieci. Przeprowadzone rozważania teoretyczne i eksperymenty symulacyjne potwierdzają zasadność proponowanej zmiany.

Słowa kluczowe: samoucząca się sieć neuronowa GNG, analiza skupień.

1. Wstęp

Samouczące się sieci neuronowe o zmiennej strukturze znajdują liczne zastosowania w grupowaniu i klasyfikacji danych. Do sieci tego typu zalicza się sieć SOM (Self Organizing Map) zaproponowaną przez Kohonena [Kohonen 1997; Deboeck, Kohonen 1998] i sieć GNG (Growing Neural Gas) zaproponowaną przez Fritzschego [1994]. Z rozważań teoretycznych, a także badań symulacyjnych [Migdał-Najman 2009; Najman 2009] wynika, że większy potencjał w grupowaniu danych wielowymiarowych prawdopodobnie ma sieć GNG. Mimo licznych zalet sieci tego typu, do których zaliczyć należy ich szybką zbieżność, oszczędną strukturę sieci, automatyczne rozpoznawanie liczby skupień separowalnych, sieć ta ma także wady. Jeżeli odległość między najbliższymi obiektami z różnych skupień jest tego samego rzędu co przeciętne odległości między obiektami w tych skupieniach, to sieć GNG może ich poprawnie nie rozdzielić. Z badań symulacyjnych wynika [Najman 2010], że jednym z najważniejszych parametrów wpływających na zdolność do rozdzielania skupień słabo separowalnych jest krok uczenia neuronu wygrywającego i jego najbliższego sąsiada. W decydujący sposób wpływa on także na czas sa-

mouczenia się sieci. W literaturze nie ma jednak wskazówek, jak krok ten powinien być ustalany.

Celem prezentowanych badań jest weryfikacja hipotezy o możliwości zwiększenia potencjału sieci GNG w analizie skupień przez uzmiennienie kroku samouczenia się sieci. Klasyczny algorytm samouczenia się zostanie zmodyfikowany przez wprowadzenie – w miejsce stałego kroku uczenia – kroku zmiennego. Zaprezentowane zostaną wyniki analizy porównawczej z algorytmem klasycznym.

2. Modyfikacja algorytmu samouczenia się sieci GNG

Istotą budowy sieci GNG jest to, że składa się ona jedynie z takiej liczby neuronów, która wystarczy do poprawnej identyfikacji skupień istniejących w zbiorze danych. Sieć powinna posiadać umiejętność odwzorowania skupień w przestrzeni o dowolnym wymiarze, o dowolnej konfiguracji w przestrzeni i autonomicznie wyznaczać liczbę skupień¹. Sieć GNG spełnia te wymagania, choć może mieć problemy z rozdzieleniem skupień słabo separowalnych [Najman 2009].

Szczególną niedoskonałością algorytmu samouczenia się sieci GNG jest stały krok uczenia neuronu wygrywającego ε_b i najbliższego połączonego z nim neuronu ε_n . W oryginalnym algorytmie szybkość zmiany położenia neuronu jest liniową funkcją błędu kwantyzacji i kroku uczenia:

$$\Delta w_{s_1} = \varepsilon_b (\xi - w_{s_1}) \quad (1)$$

$$\Delta w_i = \varepsilon_n (\xi - w_i) \quad (\forall i \in N_{s_1}), \quad (2)$$

gdzie stałe ε_b i ε_n są zwykle małymi ułamkami [Fritzke 1994; Najman 2009]. To rozwiązanie powoduje, że krok uczenia jest większy na początku procesu uczenia się sieci (duży błąd kwantyzacji), a w miarę jego postępów maleje. Ponieważ stałe ε_b i ε_n przyjmują wartości bliskie zero (dla zapewnienia małych błędów kwantyzacji w końcowej fazie uczenia się sieci), to początkowa szybkość uczenia się sieci jest w praktyce mała. Szczególnie jest to widoczne w przypadku większych zbiorów. W kolejnych iteracjach sieć uczy się przypadków położonych w dowolnej części przestrzeni obiektów i choć błąd kwantyzacji maleje, to ogólny efekt uczenia jest niewielki. Sytuacja poprawia się w miarę kolejnych iteracji, ale głównie za przyczyną wzrostu liczby neuronów. Podobnie problematyczne jest końcowe uczenie się sieci. Gdy neurony są już w pobliżu odwzorowywanych obiektów i błąd kwantyzacji jest małym ułamkiem, szybkość uczenia się jest niemal zerowa.

¹ Szczegóły budowy algorytmu uczenia się sieci GNG znajdują się w: [Fritzke 1994; Najman 2009].

W konsekwencji początkowe i końcowe uczenie się sieci jest nadmiernie zwolnione. Dla grupowania dużych zbiorów danych ważniejsza jest początkowa szybkość uczenia się sieci. Jeśli neurony szybko wypełnią całą przestrzeń zmiennych, dalsze uczenie się będzie bardzo szybkie. Końcowy etap uczenia się sieci ma kluczowe znaczenie jedynie dla rozróżniania skupień słabo separowalnych. Z powyższych powodów w prezentowanych badaniach zostanie przedstawiona propozycja zwiększenia początkowej szybkości uczenia się sieci GNG.

Rozwiązaniem problemu zbyt wolnego uczenia się sieci w pierwszym etapie tego procesu może być zmienny krok uczenia się sieci². Proponuje się, aby stałe ε_b i ε_n zastąpić wartościami odpowiedniej funkcji. Optymalnie byłoby, gdyby wartości te malały wraz ze wzrostem liczby iteracji. Wybór odpowiedniej funkcji jest tu wtórny. Niemal każda funkcja, nawet liniowa, spełnia postawione wymagania (choć w różny sposób wpływa na dynamikę zmian wartości parametrów). Jedyny warunek, jaki musi być spełniony, to taki, aby wartości funkcji były z przedziału $[0, 1]$. Dla wartości większej niż 1 neuron, zamiast się zbliżać do obiektu, mógłby go „przeskakiwać”. Dla wartości mniejszych niż 0 neuron oddalałby się od obiektu.

Ważniejszym problemem niż wybór funkcji jest ustalenie jej argumentów. Najprostszym rozwiązaniem byłoby uzależnienie wielkość kroku uczenia od ogólnej liczby iteracji samouczenia się sieci. Krok uczenia wówczas malałby wraz ze wzrostem liczby iteracji. Parametr ten ma jednak bardzo małe znaczenie w algorytmie samouczenia się sieci GNG. Jest w zasadzie tylko ograniczeniem zadanym *a priori* na czas pracy algorytmu. Jeżeli nie uda się osiągnąć zakładanego małego średniego błędu kwantyzacji sieci, to algorytm zostanie przerwany po zadanej liczbie iteracji. Ponieważ nie można w żaden sposób przewidzieć, jaka liczba iteracji jest konieczna do osiągnięcia tego celu, ustala się ją arbitralnie. Badając nieznaną zbiór danych zwykle ustala się tę liczbę tak, aby czas pracy sieci był co najwyżej kilkuminutowy. Gdy w wyniku kolejnych przybliżeń zostaną ustalone pozostałe parametry pracy sieci, liczbę iteracji zwiększa się – ale jednocześnie traci ona znaczenie. Liczba ta bowiem jest osiągnięta tylko w przypadku porażki sieci, a więc nieosiągnięcia zakładanego średniego błędu kwantyzacji. Uzależnienie kroku uczenia od ogólnej liczby iteracji skutkowałoby znaczącym zwolnieniem pracy sieci, ponieważ zmiany pozycji neuronów byłyby bardzo małe.

Od ogólnej liczby iteracji ważniejsza wydaje się liczba iteracji, która mija między kolejnymi fazami wstawiania nowego neuronu do sieci. Nowy neuron pojawia się zawsze w tej części sieci, która charakteryzuje się największym lokalnym błędem kwantyzacji³. Jest to w danym momencie najgorszy neuron, a więc także ten,

² Jest tu pewna analogia do algorytmów uczenia sieci nadzorowanych, z których wiele zawiera w sobie zmienny krok uczenia. Nie można jednak tej idei zaczerpnąć wprost z tych algorytmów ze względu na duże różnice w konstrukcji samej sieci, a także całkowitą odmienną algorytmu samouczenia się sieci GNG.

³ Błąd kwantyzacji badanego neuronu i drugiego, najbliższego i połączonego z nim neuronu.

który będzie się uczył w pierwszej kolejności. To także ten neuron wymaga najszybszego uczenia. Zastosowanie zmiennego kroku uczenia opartego na liczbie iteracji między kolejnymi fazami wstawiania nowego neuronu spowodowałoby, że nowo wstawiony neuron będzie uczył się bardzo szybko, prowadząc do szybkiego spadku lokalnego błędu kwantyzacji. Oczywiście całkowity błąd kwantyzacji sieci także będzie malał szybciej niż w klasycznym algorytmie. Jednocześnie liczba ta jest tylko małym ułamkiem ogólnej liczby iteracji, więc zmiany kroku w kolejnych iteracjach będą znacznie większe. Zmienny krok uczenia oparty na powyższej zasadzie można zapisać następująco:

$$\varepsilon_{bi} = \varepsilon_{b(i-1)} - \frac{\varepsilon_{b(i-1)}}{\max(i)}, \quad (3)$$

$$\varepsilon_{ni} = \varepsilon_{n(i-1)} - \frac{\varepsilon_{n(i-1)}}{\max(i)}. \quad (4)$$

Zmiana położenia neuronu, uwzględniająca lokalny błąd kwantyzacji i zmienny krok uczenia oparty na funkcji liniowej i liczbie iteracji między fazami wstawiania neuronu, może być ostatecznie zapisana następująco:

$$\Delta w_{s1} = \varepsilon_{bi} (\xi - w_{s1}), \quad (5)$$

$$\Delta w_i = \varepsilon_{ni} (\xi - w_i). \quad (6)$$

Konieczne jest jeszcze ustalenie początkowych wartości ε_{b0} i ε_{n0} . Proponuje się, aby wartości te były ułamkami bliskimi 0,5. Większe wartości spowodują, że w jednym kroku neuron nauczy się idealnie najbliższego obiektu i nie będzie się już więcej uczył. Pokonanie tylko połowy odległości gwarantuje większą zdolność neuronu do generalizacji, a więc odwzorowywania większej liczby podobnych obiektów.

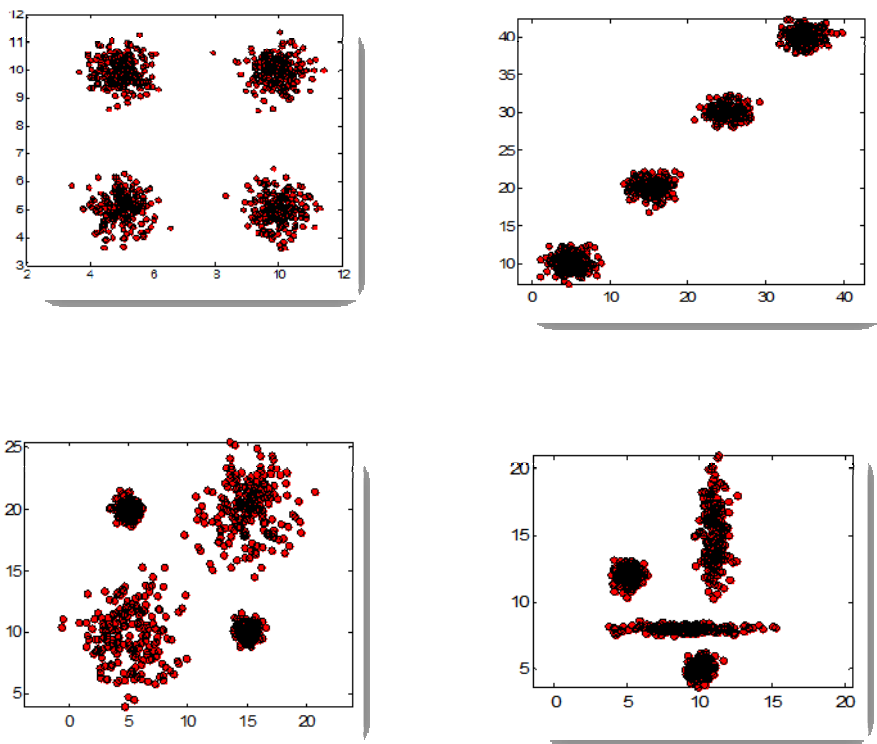
Należy się spodziewać, że w momencie wstawiania nowego neuronu będzie się on uczył bardzo szybko. W kolejnych iteracjach pozycje neuronów będą się zmieniać coraz wolniej aż do momentu wstawienia kolejnego neuronu, kiedy nastąpi znaczący skok w położeniu neuronów. Strategia ta może potencjalnie przyczynić się także do ograniczenia liczby neuronów uczących się „na pamięć”⁴.

⁴ W algorytmach uczenia klasycznych sieci warstwowych stosuje się strategię polegającą na okresowym zaburzaniu wag wszystkich neuronów o małe losowe wartości. Dzięki temu zabiegowi neuron, który dobrze się nauczył odwzorowywania pewnej klasy obiektów, po kilku iteracjach powróci na swoje miejsce. Jednocześnie neurony, które „utknęły” w lokalnym minimum funkcji błędu uczenia, a więc nauczyły się bezbłędnie identyfikować, ale tylko jeden obiekt, tracąc w ten sposób zdolność do generalizacji, mogą zostać „wybite” z tego minimum i w kolejnych iteracjach mogą zna-

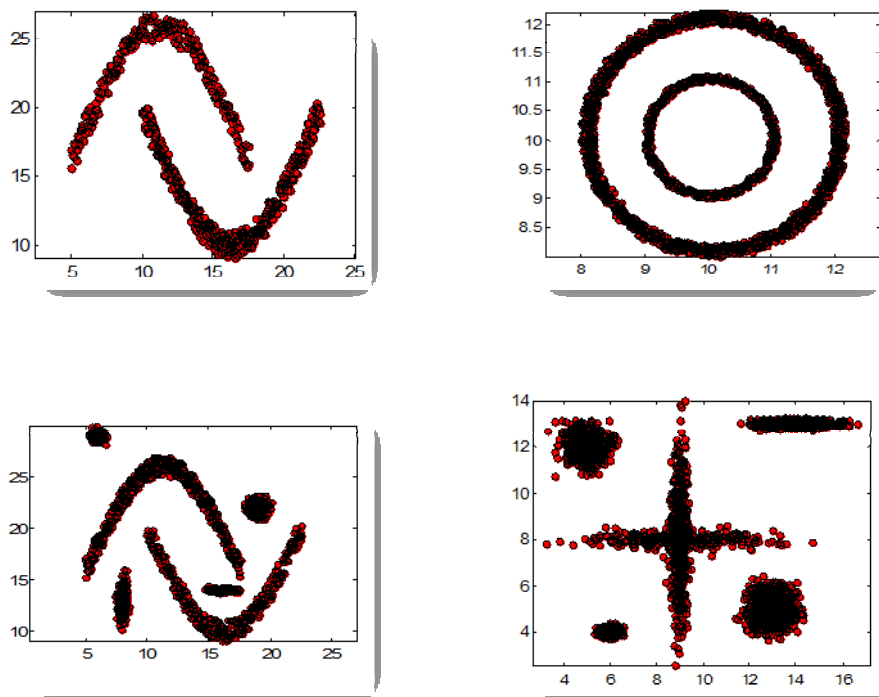
3. Eksperyment badawczy

W celu weryfikacji powyższych hipotez przeprowadzono szereg eksperymentów. Przygotowano osiem zbiorów danych umownych o znanej strukturze grupowej i przeprowadzono analizę porównawczą efektywności klasycznego algorytmu samouczenia się sieci GNG z algorytmem ze zmiennym krokiem uczenia. Zbiory te przedstawione są na rysunku 1. Różnią się liczbą obiektów, liczbą skupień i ich konfiguracją w przestrzeni. Niektóre z nich mają prostą strukturę grupową inne bardziej złożoną. Dla każdego ze zbiorów przeprowadzono grupowanie klasycznym algorytmem GNG i zmodyfikowanym.

Algorytm zmodyfikowany różni się od klasycznego wyłącznie elementem uzmiennienia kroku uczenia. Porównano trzy parametry uzyskanych rozwiązań: uzyskaną liczbę neuronów (przy danym średnim błędzie kwantyzacji im mniej, tym lepiej), czas



leżąc minimum globalne. Dobrze nauczona sieć po takim wstrząsie w niewielkiej liczbie iteracji odtworzy swoją strukturę. Słabo nauczona sieć może znacząco poprawić swoją efektywność.



Rys. 1. Zbiory testowe

Źródło: opracowanie własne.

uczenia się sieci (przy danym średnim błędzie kwantyzacji im krótszy, tym lepiej) i średni błąd kwantyzacji (im mniejszy, tym lepiej). Wyniki symulacji przedstawiono w tab. 1. Podkreślono najlepsze wyniki w każdej kategorii.

Ogólna liczba neuronów uzyskana przez sieć uczoną algorytmem klasycznym jest zwykle mniejsza niż w przypadku algorytmu zmodyfikowanego, jednak różnice te nie są istotne. Większa liczba neuronów zostaje osiągnięta w krótszym czasie w każdym z badanych przypadków. Te różnice są znaczące i sięgają nawet 20%. W problemach większej skali taka zmiana będzie bardzo istotna z praktycznego punktu widzenia. Najważniejsze jednak jest to, że w większości przypadków sieć ze zmiennym krokiem uczenia się osiągnęła mniejszy średni błąd kwantyzacji. Różnice są w tym przypadku także znaczące, sięgają bowiem nawet 100% (zbiór ostatni). Co ważne, tak dużą poprawę jakości sieci uzyskano przy identycznej liczbie neuronów i znacząco krótszym czasie uczenia się sieci.

Tabela 1. Wyniki analizy porównawczej

Zbiór testowy	Liczba neuronów		Czas uczenia się (w sek.)		Średni błąd kwantyzacji	
	GNG	GNG	GNG	GNG	GNG	GNG
	krok uczenia					
	stały	zmienny	stały	zmienny	stały	zmienny
1	72	74	0,5763	0,4552	0,00082	0,00072
2	75	78	0,5706	0,4563	0,00440	0,00580
3	61	64	0,5651	0,4495	0,00360	0,00270
4	71	68	0,5693	0,4534	0,00087	0,00110
5	97	95	0,7513	0,6065	0,01420	0,00950
6	99	100	0,7682	0,6284	0,00110	0,00087
7	32	34	0,7328	0,5708	0,00910	0,00940
8	32	32	0,7549	0,6096	0,00210	0,00100

Źródło: opracowanie własne.

4. Wnioski

Prezentowane rozważania teoretyczne pozwalają sądzić, że możliwy jest wzrost użyteczności sieci GNG w analizie skupień. Wstępne badania symulacyjne zdają się potwierdzać tę tezę. Zaobserwowano istotne obniżenie czasu trwania procesu samouczenia się przy braku istotnego wzrostu rozmiarów sieci i możliwym zwiększeniu zdolności sieci do rozpoznawania struktury grupowej. Osiągnięte wyniki zachęcają do dalszych badań – szczególnie na dużych empirycznych bazach danych.

Literatura

- Deboeck G., Kohonen T., *Visual explorations in Finance with Self-Organizing Maps*, Springer-Verlag, London 1998.
- Jirayusakul A., Auwatanamongkol S., *A Supervised Growing Neural Gas Algorithm for Cluster Analysis*, „International Journal of Hybrid Intelligent Systems” April 2007, vol. 4, Issue 2, 2007, s. 129-141.
- Kohonen T., *Self-Organizing Maps*, Springer Series in Information Sciences, Springer-Verlag, Berlin – Heidelberg 1997.
- Fritzke B., *Growing cell structures – a self-organizing network for unsupervised and supervised learning*, „Neural Networks 1994, vol. 7, no. 9, s. 1441-1460.
- Migdał-Najman K., *Analiza porównawcza własności nienadzorowanych sieci neuronowych typu Self Organizing Map i Growing Neural Gas w analizie skupień*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 47, Taksonomia 16, Wrocław 2009, s. 205-213.
- Najman K., *Ocena wpływu parametrów sterujących procesem samouczenia się sieci GNG na ich zdolność do separowania skupień*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 107, Taksonomia 17, UE, Wrocław, 2010, s. 296-304.
- Najman K., *Zastosowanie nienadzorowanych sieci neuronowych typu Growing Neural Gas w analizie skupień*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 47, Taksonomia 16, Wrocław 2009, s. 196-204.

Qin A. K., Suganthan P. N., *Robust growing neural gas algorithm with application in cluster analysis*, „Neural Networks” 2004, vol. 17, no. 8-9, s. 1135-1148.

Prudent Y., Ennaji A., *An Incremental Growing Neural Gas Learns Topologies*, Proceedings of International Joint Conference on Neural Networks, 2005, s. 1211-1216.

PROPOSITION OF SELF-LEARNING ALGORITHM OF GNG NEURAL NETWORK WITH CHANGING LEARNING STEP

Summary: One of the key parameters of self-learning GNG neural network process is the speed of the change of position in the space of a learning unit and the nearest unit connected to it. It depends on a local quantization error and a learning step. In the initial phase the constant value of the learning step unnecessarily slows down the self-learning process. The article presents the proposal of algorithm modification with the changing learning step. Theoretical discussions and simulation experiments confirm the legitimacy of a proposed change.