Spontaniczna struktura bezskalowa w grafach przepływu impulsów dla rekurencyjnych sieci neuronowych

Autoreferat rozprawy doktorskiej

Filip Piękniewski

8 października 2008

1 Wstęp

W pracy przedstawiamy wyniki, zarówno teoretyczne jak i numeryczne, opisujące samoorganizację funkcjonalnych grafów połączeń dla kilku modeli rekurencyjnych sieci neuronowych. Rezultatem badanej samoorganizacji jest przede wszystkim spontaniczne powstawanie w takich grafach struktur bezskalowych, czyli takich w których rozkład stopni wierzchołków spełnia prawo potęgowe. Dyskutowane szerzej w pracy sieci bezskalowe stanowią intensywnie eksplorowaną obecnie gałąź badań grafów losowych, szczególnie w kontekście doniesień, iż bardzo wiele grafów występujących w otaczającym nas świecie ma cechy bezskalowości. Do tej pory cechy takie stwierdzono w strukturze Internetu, sieciach współpracy naukowej/artystycznej, społecznych, lingwistycznych, metabolicznych i wielu innych. Tymczasem wyniki dotyczące struktury połączeń sieci neuronowych nie są jednoznaczne. Pewne doniesienia wskazują na wykładniczy zanik ogona rozkładów stopni wierzchołków w układach neuronowych niewielkich organizmów (na poziomie synaps), inne natomiast badania przeprowadzone na mózgu ludzkim za pomocą funkcjonalnego rezonansu magnetycznego (fMRI) wskazują, iż grafy indukowane przez centra aktywności są istotnie bezskalowe, z wykładnikiem zgodnym z naszymi przewidywaniami teoretycznymi. W dalszych rozdziałach tej pracy prezentujemy model matematyczny asynchronicznej impulsującej sieci neuronowej uogólniającej koncepcję maszyny Boltzmanna. Analizujemy minima energetyczne (stany bazowe) prezentowanego modelu, oraz graf indukowany przez przepływy potencjału pomiędzy jednostkami. Dowodzimy, że dla pewnych naturalnych założeń dotyczących funkcji energetycznej i w odpowiednio niskiej temperaturze wielkoskalowy opis układu sprowadza się do dynamiki typu “zwycięzca bierze wszystko” z czego można dalej wnioskować, iż indukowany graf przepływu impulsów jest bezskalowy z wykładnikiem $\gamma = 2$. Ponadto prezentujemy wyniki badań symulacyjnych, które wskazują, iż zaproponowany model dobrze odzwierciedla interakcje między izolowanymi grupami neuronów opartych o fenomenologiczny model Eugene M. Izhikevicha. Wskazujemy dalsze kierunki badań oparte o spontanicznie wykształcające się grupy neuronów oraz wielkoskalowe symulacje układu nerwowego.

2 Zarys problematyki: sieci bezskalowe a sieci neuronowe

Koncepcja sieci bezskalowych wywołała w ostatnich latach spore zainteresowanie, oferując zunifikowane narzędzie opisu dla szerokiej klasy grafów których mechanizmy organizacyjne są związane z pewnym stopniem losowości (Albert & Barabási, 2002). Badania te związane są z faktem, iż wiele sieci obecnych w rozmaitych aspektach rzeczywistości ma istotnie cechy bezskalowości. Do tej pory takie struktury zostały potwierdzone w strukturze Internetu (zarówno na poziomie stron WWW jak i warstwach niższych) (Albert *et al.* , 1999), sieciach współpracy naukowej (Barabási *et al.* , 2002), sieciach cytowań (Redner, 1998), sieciach ekologicznych (relacji w łańcuchach pokarmowych) (Montoya & V., 2002), sieciach lingwistycznych (i Cancho & Solé, 2001), metabolicznych (Jeong *et al.* , 2000) i wielu innych. Wiele z powyższych sieci było wcześniej modelowanych jako klasyczne grafy losowe Erdős'a-Rényi'ego, których struktura jednak znacznie odbiega od tej spotykanej w grafach bezskalowych. Obszerne wyniki empiryczne zdają się wskazywać, iż bezskalowość może być pożądaną cechą dla sieci przetwarzających informację. Grafy bezskalowe są stosunkowo rzadkie, jednak zachowują dobrą łączność (w terminach przeciętnej długości scieżek), są przy tym dość odporne na rozspójnienie. Takie cechy mogą być interesujące szczególnie dla sieci neuronowych. Dotychczasowe modele sieci neuronowych nie wymuszają jednak w żaden sposób architektur bezskalowych, ich struktura oscyluje pomiędzy kilkoma ustalonymi architekturami

- skierowanej warstwowej (obecna w klasycznych perceptronach wielowarstwowych) jest niejako wymuszona przez ideę uczenia takich sieci za pomocą algorytmu wstecznej propagacji
- całkowicie rekurencyjna (np. sieci Hopfielda) jest wygodna do analizy, nie jest jednak dobrze umotywowana biologicznie, a także nie jest optymalna ze względu na pamięć (ilość połączeń rośnie z kwadratem rozmiaru sieci)
- lokalna struktura jest używana w niektórych zastosowaniach miedzy innymi dotyczących przetwarzania obrazu, jednak jej zastosowanie jest zasadniczo

ograniczone do specyficznych zadań, także brakuje tu motywacji biologicznej

W ostatnich latach pojawiły się pewne prace dotyczące zaimportowania struktur bezskalowych (a także grafów *małego świata*) w klasycznych sieciach neuronowych, co przyniosło interesujące rezultaty (Perotti *et al.* , 2006; Stauffer *et al.* , 2003), jednak struktura jest tutaj jedynie narzucona na wcześniej istniejące modele. Badania empiryczne dotyczące niewielkiego robaka Caenorhabditis elegans (Amaral *et al.* , 2000; Koch & Laurent, 1999) wskazują, iż na poziomie pojedynczych neuronów prawdopodobnie bezskalowość nie występuje (przynajmniej w statycznej anatomii). Badania dotyczące funkcjonalnego działania ludzkiego mózgu przeprowadzone za pomocą fMRI (Eguíluz *et al.* , 2005) pokazują już jednak zupełnie inny obraz - sieci funkcjonalne łączące skorelowane ośrodki aktywności mózgu przy rozwiązywaniu pewnych zadań są bezskalowe z wykładnikiem $\gamma \approx 2$, a także mają cechy *małego świata*.

Obecnie istnieje wiele modeli kontrukcji sieci, produkujących jako wyniki losowe sieci bezskalowe. Duża część tych konstrukcji wywodzi się od znanego modelu Barabási'ego i Albert (Barabási & Albert, 1999) opartego o *wzrost* i *preferencyjne podczepanie*[1]. W książce (Chung & Lu, 2006) analizowane są także modele bezpośrednio konstruujące graf o zadanym, oczekiwanym rozkładzie stopni wierzchołków (uogólniające model Erdős'a-Rényi'ego), a także modele oparte o duplikowanie wierzchołków, które produkują grafy bezskalowe z wykładnikiem prawa potęgowego mniejszym niż 2 (takie sieci często pojawiają się w biochemii). Żadne z tych modeli jednak nie odzwierciedlają zjawisk występujących w układzie nerwowym odpowiedzialnych za generowanie wspomnianych grafów funkcjonalnych (w których stale przesyłane są pewne impulsy między jednostkami). Nasz model zaproponowany w (Piękniewski & Schreiber, 2007) oraz rozwinięty w (Piękniewski & Schreiber, 2008) ma na celu rozpoczęcie wypełniania tej luki.

3 Sieci przepływu impulsów

W pracy doktorskiej proponujemy nowy model (Piękniewski & Schreiber, 2007, 2008) impulsującej sieci neuronowej opartej na klasycznej maszynie Boltzmanna. Naszym celem było stworzenie modelu, który mógłby stanowić pomost pomiędzy sieciami neuronowymi a grafami bezskalowymi, i mógłby wskazać, jakich cech należy się spodziewać w dynamice układów impulsujących, aby indukowany graf miał cechy bezskalowości. Prezentowany model (zwany umownie dalej *modelem przepływowym*) potencjalnie może stanowić także opis innych zjawisk, takich jak przepływ kapitału w pewnym układzie ośrodków finansowych (który nota bene

[1]preferential attachment

też ma cechy bezskalowe, co zauważył już w swoich wykładach Vilfredo Pareto (1896-1897)), czy też inne sieci wymiany (np. wymiany danych w Internecie).

Model przepływowy składa się z N jednostek σ_i, które mogą przyjmować wartości (stany) w zbiorze liczb naturalnych. Całą konfigurację układu oznaczamy $\bar{\sigma}$. Każda para jednostek połączona jest z pewną wagą $w_{i,j}$ wylosowaną niezależnie z rozkładu Gaussa $\mathcal{N}(0,1)$. Układ wyposażony jest funkcję energetyczną postaci:

$$\mathcal{H}(\bar{\sigma}) := \frac{1}{2}\sum_{i \neq j} w_{ij}|\sigma_i - \sigma_j| \tag{1}$$

zatem dodatnia waga $w_{i,j}$ wspiera (w sensie minimalizacji energii) sytuację w której stany σ_i oraz σ_j są zgodne, natomiast ujemna waga $w_{i,j}$ odpowiednio wspiera sytuację w której stany σ_i oraz σ_j są rozbieżne. Dynamika układu przebiega następująco:

- Losowo, jednostajnie wybrana zostaje para uporządkowana (σ_i, σ_j). Jednostka σ_i będzie źródłem przepływu potencjału, zaś σ_j jego celem.
- Jeśli $\sigma_i = 0$ (nie ma żadnego potencjału), wracamy do punktu powyżej i losujemy nową parę.
- Jeśli $\sigma_i > 0$ (zatem istnieje pewien potencjał do zaoferownia), tworzymy konfigurację $\bar{\sigma}'$ powstałą z poprzedniej konfiguracji $\bar{\sigma}$ przez zmniejszenie wartości σ_i o 1 i zwiększenie wartości σ_j o tyle samo (przepływ jednostki potencjału z σ_i do σ_j).
- Jeśli $\mathcal{H}(\bar{\sigma}') < \mathcal{H}(\bar{\sigma})$, akceptujemy nową konfigurację. W przeciwnym wypadku akceptujemy ją z prawdopodobieństwem $\exp(-\beta[\mathcal{H}(\bar{\sigma}^*) - \mathcal{H}(\bar{\sigma})])$, $\beta > 0$, i odrzucamy z prawdopodobieństwem dopełniającym (gdzie β określa temperaturę odwrotną, zgodnie ze standardowymi oznaczeniami spotykanymi w mechanice statystycznej). Jeśli nowa konfiguracja nie została zaakceptowana, przywracamy poprzednią.
- Dodatkowo zakładamy, że układ zawiera skończoną ilość jednostek potencjału (z tego względu przestrzeń możliwych konfiguracji też jest skończona).

Łatwo zauważyć, że sumaryczna ilość dostępnego potencjału $\sum_i \sigma_i$ jest zachowana w trakcie dynamiki. Ponadto dla każdej konfiguracji początkowej $\bar{\sigma}^0$ każda inna konfiguracja $\bar{\sigma}$ spełniająca $\sum_i \sigma_i^0 = \sum_i \sigma_i$ jest osiągalna z pewnym (być może niewielkim) prawdopodobieństwem po pewnej ilości kroków, łatwo też sprawdzić aperiodyczność dynamiki. Standardowa weryfikacja równań dokładnego bilansu potwierdza, że stanami stacjonarnymi uzyskanego odwracalne łańcucha Markowa są rozkłady postaci:

$$\mathbb{P}_n(\bar{\sigma}) = \begin{cases} \frac{\exp(-\beta\mathcal{H}(\bar{\sigma}))}{\left(\sum_{\bar{\sigma}', \sum_i \sigma_i'=n} \exp(-\beta\mathcal{H}(\bar{\sigma}'))\right)}, & \text{jeśli } \sum_i \sigma_i = n, \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{2}$$

oraz ich kombinacje wypukłe. Zaprezentowany kształt Hamiltonianu może sugerować, iż stan bazowy rozważanego modelu może być bardzo skomplikowany i trudny do wysymulowania. Faktycznie jednak, stan bazowy ma dość prosty charakter (co dowodzimy w pracy):

Twierdzenie 1 *W modelu przepływowym dla wag $w_{i,j}$ wylosowanych z rozkładu Gaussa $\mathcal{N}(0,1)$ z dużym prawdopodobieństwem stan bazowy (minimum energetyczne) rozważanego układu ma postać:*

- *Jednostka σ_i dla której*
$$S_i := -\sum_{j\neq i} w_{ij}.$$
jest maksymalne przechowuje cały dostępny potencjał.
- *Wszystkie pozostałe jednostki mają zerowy stan posiadania.*

Oznacza to, że w układzie niezależnie od stanu początkowego przepływ zgodny z dynamiką ostatecznie kończy się w stanie, w którym "najlepsza" z jednostek przechwytuje cały potencjał - stąd wielkoskalowy opis sprowadza się do dynamiki "zwycięzca bierze wszystko".

Przepływy impulsów w układzie indukują nowy graf w którym waga jest liczbą naturalną określającą ile razy dana krawędź była wykorzystana do przesłania impulsu. Powstały graf jest nietrywialny (niektóre krawędzie są rzadko uczęszczane, inne za to są oblegane). Głównym rezultatem opisanym w pracy jest następujące

Twierdzenie 2 *Dla modelu przepływowego zdefiniowanego jak wyżej, indukowany graf przepływu impulsów jest bezskalowy z wykładnikiem $\gamma = 2$.*

W pracy rozważane są także sytuacje w których pojemność jednostek jest ograniczona. Dynamika w takich wariantach modelu nie różni się znacznie od wersji "nieograniczonej" gdy ograniczenie na pojemność jednostki jest w miarę duże (a tylko takie przypadki są analizowane w pracy).

4 Potwierdzenia numeryczne

Praca zawiera szereg symulacji numerycznych potwierdzających otrzymane rezultaty. Najbardziej oczywistą symalacją było zaimplementowanie bezpośrednio modelu przepływowego (Piękniewski & Schreiber, 2007). Obliczenia te w pełni potwierdziły nasze oczekiwania. Ciekawsza jest symulacja (Piękniewski, 2007) wykonana w oparciu o fenomenologiczny model impulsujący zaproponowany w (Izhikevich, 2003) i pokazująca jak zjawisko bezskalowości może pojawić się w bardziej naturalnym z biologicznego punktu widzenia modelu. W tym przypadku neurony są zebrane w izolowane grupy, graf połączeń otrzymywany jest na podstawie miary korelacji aktywności (metodologia jest podobna do tej stosowanej w empirycznych badaniach aktywności mózgu Eguíluz *et al.* (2005)). W tym przypadku także widoczna jest struktura bezskalowa z wykładnikiem $\gamma \approx 2$, która znika po zastąpieniu grup pojedynczymi neuronami. Daje to wskazówki, dlaczego biologiczne sieci neuronowe na poziomie pojedynczych synaps i neuronów nie są bezskalowe - być może nie jest to odpowiednia skala do występowania tego zjawiska. Grupy neuronów mają znacznie bardziej skomplikowaną dynamikę, stanowią pewnego rodzaju rezerwuary w którym nadchodząca aktywność może być przez pewien czas przechowywana i wysyłana dalej, co upodobnia je do modelu przepływowego.

5 Kierunki dalszych badań

Badania przedstawione w pracy mają naturalne kierunki kontynuacji. Jednym z takich kierunków jest poszukiwanie podobnych zjawisk w bardziej realistycznych modelach układów nerwowych. Preliminarz takich badań dyskutowany jest ostatnim rozdziale pracy, opisującym implementację modelu samoistnie wykształcających się grup neuronowych (Izhikevich *et al.* , 2004). Wyniki tej symulacji nie potwierdzają występowania bezskalowości, jest to prawdopodobnie spowodowane dosyć arbitralnym kryterium określającym pojęcie *grupy* w tym modelu, oraz trudnym do estymowania szumem. Model ten należy raczej potraktować jako wstęp do bardziej skomplikowanych modeli opartych o dane anatomiczne, takich jak proponowany w (Izhikevich & Edelman, 2008).

Kolejnym kierunkiem jest czysto matematyczna analiza grafów powstających w wyniku działania modelu przepływowego, w szczególności analiza własności spektralnych i relacji modelu przepływowego z innymi sposobami konstrukcji sieci bezskalowych.

W dalszej perspektywie ciekawym wydaje się także badanie innych wariantów modelu przepływowego, w szczególności zadając na nim strukturę topologiczną (np. w oparciu o dane dotyczące struktury układu nerwowego).

Literatura

Albert, Réka, & Barabási, Albert-László. 2002. Statistical mechanics of complex networks. *Reviews of modern physics*, January, 47–97. 2

Albert, Réka, Jeong, Hawoong, & Barabási, Albert-László. 1999. Diameter of the world-wide web. *Science*, **401**(Septmeber), 130–131. 2

Amaral, L. A., Scala, A., Barthelemy, M., & Stanley, H. E. 2000. Classes of small-world networks. *Proc natl acad sci u s a*, **97**(21), 11149–11152. Dostępne na: `http://dx.doi.org/10.1073/pnas.200327197`, `doi:10.1073/pnas.200327197`. 3

Barabási, Albert-László, & Albert, Réka. 1999. Emergence of scaling in random networks. *Science*, October, 509–512. 3

Barabási, Albert-László, Jeong, Hawoong, Néda, Zoltan, Ravasz, Erzsebet, Schubert, A., & Vicsek, Tamas. 2002. Evolution of the social network of scientific collaborations. *Physica a*, **311**(4), 590–614. 2

Chung, Fan, & Lu, Linyuan. 2006. *Complex graphs and networks (cbms regional conference series in mathematics)*. Boston, MA, USA: American Mathematical Society. 3

Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., & Apkarian, A. V. 2005. Scale-free brain functional networks. *Phys rev lett*, **94**(1). Dostępne na: `http://view.ncbi.nlm.nih.gov/pubmed/15698136`. 3, 6

i Cancho, Ramon Ferrer, & Solé, Ricard V. 2001. The small-world of human language. *Proceedings of the royal society of london b*, **268**(1482), 2261–2265. 2

Izhikevich, Eugene M. 2003. Simple model of spiking neurons. *Ieee transactions on neural networks*, 1569– 1572. Dostępne na: `http://www.nsi.edu/users/izhikevich/publications/spikes.pdf`. 6

Izhikevich, Eugene M., & Edelman, Gerald M. 2008. Large-scale model of mammalian thalamocortical systems. *Proceedings of the national academy of sciences*, February, 3593–3598. Dostępne na: `http://dx.doi.org/10.1073/pnas.0712231105`, `doi:10.1073/pnas.0712231105`. 6

Izhikevich, Eugene M., Gally, Joe A., & Edelman, Gerald M. 2004. Spike-timing dynamics of neuronal groups. *Cerebral cortex*, 933– 944. Dostępne na: `http://vesicle.nsi.edu/users/izhikevich/publications/reentry.pdf`. 6

Jeong, Hawoong, Tombor, B., Albert, Réka, Oltvai, Zoltan N., & Barabási, Albert-László. 2000. The large-scale organization of metabolic networks. *Nature*, **407**(6804), 651–653. 2

Koch, Christof, & Laurent, Gilles. 1999. Complexity and the Nervous System. *Science*, **284**(5411), 96–98. Dostępne na: http://www.sciencemag.org/cgi/content/abstract/284/5411/96, arXiv:http://www.sciencemag.org/cgi/reprint/284/5411/96.pdf, doi:10.1126/science.284.5411.96. 3

Montoya, Jose M., & V., Ricard V. Solé. 2002. Small world patterns in food webs. *Journal of theoretical biology*, **214**(3), 405–412. 2

Pareto, Vilfredo. 1896-1897. *Cours d'économie politique*. Rouge, Lausanne. 4

Perotti, J. I., Tamarit, F. A., & Cannas, S. A. 2006. A scale-free neural network for modelling neurogenesis. *Physica a statistical mechanics and its applications*, **371**(Nov.), 71–75. doi:10.1016/j.physa.2006.04.079. 3

Piękniewski, F. 2007. Emergence of scale-free graphs in dynamical spiking neural networks. *Pages 755–759 of: Proc. IEEE international joint conference on neural networks*. Orlando, Florida, USA: IEEE Press. Dostępne na: http://www.ieeexplore.ieee.org/xpl/freeabs˙all.jsp?isnumber=4370891&arnumber=4371052&count=569&index=160. 6

Piękniewski, F., & Schreiber, T. 2007 (April). Emergence of scale-free spike flow graphs in recurrent neural networks. *Pages 357–362 of: Proc. IEEE symposium series in computational intelligence - foundations of computational intelligence*. Dostępne na: http://doi.ieeecomputersociety.org/10.1109/FOCI.2007.371496. 3, 6

Piękniewski, Filip, & Schreiber, Tomasz. 2008. Spontaneous scale-free structure of spike flow graphs in recurrent neural networks;. *Neural networks*, **In Press, Corrected Proof**, –. Dostępne na: http://www.sciencedirect.com/science/article/B6T08-4SVKSNF-4/2/0363e0991d6ebe0fadcdf859b906018a, doi:doi:10.1016/j.neunet.2008.06.011. 3

Redner, S. 1998. How popular is your paper? an empirical study of the citation distribution. *European physical journal b*, **4**(2), 131–134. 2

Stauffer, Dietrich, Aharony, Amnon, da Fontoura Costa, Luciano, & Adler, Joan. 2003. Efficient hopfield pattern recognition on a scale-free neural network. *The european physical journal b*, 395–399. doi:10.1140/epjb/e2003-00114-7. 3