Модель Эластичной Сети с Неявным Использованием Информации о Степенях Узлов

Вадим Шергин, Лариса Чалая, Мария Погурская
Кафедра искусственного интеллекта
Харьковский национальный университет радиоэлектроники
Харьков, Украина
vadim.shergin@nure.ua, larysa.chala@nure.ua

Сергей Удовенко
Кафедра информатики и вычислительной техники
Харьковский национальный экономический университет имени Симона Кузнеца
Харьков, Украина
serhiy.udovenko@hneu.net

Model of Elastic Network with Implicit Use of Node Degree Data

Vadim Shergin, Larysa Chala, Mariya Pogurskaya
Artificial Intelligence dept.
Kharkiv National University of Radio Electronics
Kharkov, Ukraine
vadim.shergin@nure.ua, larysa.chala@nure.ua

Serhii Udovenko
Informatics and Computer Engineering dept
Simon Kuznets Kharkiv National University of Economics
Kharkov, Ukraine
serhiy.udovenko@hneu.net

Аннотация – Рассматривается проблема моделирования масштабно-инвариантных сетей. Предлагается модель эластичной сети с неявным использованием информации о степенях узлов. Основной особенностью модели является использование фактора копирования в качестве управляющего параметра модели вместо традиционно используемого количества связей входящего узла. Полученная модель сочетает ключевые преимущества модели с присоединением, основанным на посредничестве, и модели эластичной сети. Свойства разработанной модели исследованы аналитически и подтверждены результатами численного моделирования.

Abstract – The problem of modeling scale-free networks is considered. Model of elastic network with implicit use of node degree data is proposed. Using the copy-factor as the control parameter instead of traditionally used number of incoming links is the main feature of the proposed model. This model combines the key benefits of mediation-driven attachment model, and elasticity-based model. The properties of the proposed model have been studied analytically and fully confirmed by the simulation results.

Ключевые слова – масштабно-инвариантная сеть; эластичность; присоединение, основанное на посредничестве; фактор копирования

Keywords – scale-free network; elasticity; mediation-driven attachment; copy factor

I. ВВЕДЕНИЕ

Известно [1,2], что многие из сетей реального мира являются масштабно-инвариантными (МИС), например, социальные сети, сети сотрудничества, сети связи, нейронные и белковые сети, некоторые транспортные сети и т.д. Ключевым свойством МИС является степенной закон распределения числа связей у узлов таких сетей. В свою очередь, степенное распределение является характерной чертой фрактальных свойств системы и её способности к самоорганизации. Этот фактор является ещё одной причиной интереса к моделям МИС.

В основе большинства современных моделей МИС лежит модель, предложенная Барабаши и Альберт [3] (BA-модель). Хотя в настоящее время существует множество расширений и обобщений этой модели [4-9], многие из них обладают общими существенными недостатками, восходящими к BA-модели.

В этой связи наибольший интерес представляют модель присоединения, управляемого посредничеством [10], и модель эластичных сетей [11-12], позволяющие соответственно не использовать информацию о степенях узлов и строить плотные сети. Построение модели сети, сочетающей преимущества вышеназванных моделей, составляет предмет данного исследования.

II. Обзор моделей МИС и постановка задачи

Согласно концепции роста, начальная сеть состоит из n_0 узлов и L_0 связей между ними, а затем на каждом шаге к сети добавляется новый узел. Таким образом, размер сети n (количество узлов) можно использовать в качестве меры времени, номер узла (i) – как время его рождения, а величина $n-i$ служит возрастом узла.

Каждый узел сети характеризуется своей степенью (т.е. количеством связей, k_i) и другими свойствами P_i .

Свойства сети определяются правилом присоединения, то есть зависимостью вероятности π_i присоединения новой связи к узлу от его индивидуальных свойств (i, k_i, P_i) и от общих параметров сети (её размера n, начальных условий n_0, L_0 и управляющих параметров U):

$$\pi_i = f(i, k_i, n_0, L_0, P_i, U). \qquad (1)$$

Динамика роста сети определяется количеством ссылок $m(n)$, соединяющих входящий узел с существующими, или средней степенью узлов $\bar{k}(n)$.

Ключевым свойством МИС является распределение узлов по числу связей $p(k)$, которое должно (по крайней мере асимптотически), подчиняться закону Юла-Саймона, являющемуся в свою очередь дискретным аналогом степенного закона:

$$p(k) = C\frac{\Gamma(k)}{\Gamma(k+\gamma)} \propto k^{-\gamma}, \quad k \ge k_0 \qquad (2)$$

Параметр $\gamma \ge 2$ называется показателем скейлинга.

Самая простой и известной моделью МИС является модель ВА [3]. В её основе лежат две фундаментальные концепции: концепция роста и концепция предпочтительного присоединения, согласно которой вероятность того, что новый узел (с номером $n+1$) присоединится к существующему (номер i) пропорциональна его степени:

$$\pi_i = k_i / \sum k_i \qquad (3)$$

Легко видеть, что описанное правило (3) является простейшим частным случаем (1). Модель BA приводит к распределению степеней узлов по закону Юла-Саймона (2) с показателем скейлинга $\gamma = 3$. Управляющим параметром служит количество связей $m = const$, инцидентных входящему узлу.

В настоящее время существует множество различных расширений и обобщений модели ВА, которые тем не менее укладываются в общую форму (1) генерирующего механизма (1) (фитнесс-модель [4], модель с фактором старения [5], с фактором дополнительной привлекательности [2], на основе нелинейного предпочтительного присоединения [6], перенаправления связей [7], с учётом свойств узлов второго уровня соседства [8] и т.д.).

Не все из перечисленных моделей порождают именно МИС. Показано [6], что это происходит лишь в том случае, когда правило присоединения (1) является асимптотически линейным относительно k_i , т.е. $\pi_i \sim a_\infty k_i$. С другой стороны, несмотря на широкое распространение моделей МИС в теоретических исследованиях, этот класс сетей является далеко не единственным среди сетей реального мира [9].

Являясь ядром моделей МИС, ВА-модель несёт в себе ряд недостатков, унаследованных и перечисленными нисходящими моделями:

- В сетях реального мира средняя степень узлов $\bar{k}(n)$ имеет тенденцию к росту с размером сети, в то время как в ВА-модели она постоянна. Таким образом, эта модель генерирует очень разреженные сети;
- Модель ВА порождает сети, нейтральные по ассортативности [7, 13], в то время как социальные сети существенно ассортативны, а биологические и технические – дизассортативны;
- Предполагается, что информация о степени узлов k_i используется явно, т.е. входящий узел "знает" (или может оценить) степени всех существующих узлов сети;
- Управляющие параметры моделей (например, m), являются, как правило, масштабно-зависимыми, что противоречит самой концепции масштабной инвариантности.

Для преодоления первого из указанных недостатков была предложена модель эластичных сетей [11-12]. В этой модели предполагается, что относительные темпы роста числа связей и узлов различны. Их отношение называется показателем эластичности:

$$\lambda = \frac{\delta L(t)}{\delta n(t)} = \frac{\Delta L(t)}{L(t)} \Big/ \frac{\Delta n(t)}{n(t)} = n\frac{L(n+1)-L(n)}{L(n)}. \qquad (4)$$

При $\lambda = const$ $(1 \le \lambda < 2)$ общее количество связей в сети растёт асимптотически по степенному закону с показателем λ . Средняя степень узлов $\bar{k}(n)$ также является не постоянной (равной m), а растёт по степенному закону

$$\bar{k}(n) = L(n)/n \propto n^{\lambda-1}. \qquad (5)$$

Так, при $\lambda > 1$ каждый новый узел приносит больше связей, чем текущее среднее их значение (5). Например, чем больше научных статей было написано, тем больше ссылок должен пересмотреть автор новой статьи.

Таким образом, использование концепции эластичности обобщает исходную модель БА и позволяет создавать плотные МИС.

Модель присоединения, управляемого посредничеством (MDA), предложенная Хассаном и др. в [10], позволяет преодолеть второй из перечисленных недостатков семейства BA-моделей: явное использование информации о степенях узлов.

Согласно модели MDA, новый узел выбирает одного из существующих в качестве посредника (равновероятно) и затем соединяется (также равновероятно) с m соседями этого посредника. Показано [10], что при $m > 14$ вероятность присоединения нового узла к существующему зависит только от его степени, однако в отличие от семейства BA-моделей эта зависимость неявная.

Для большинства реальных сетей правило MDA выглядит намного естественнее, чем правило BA, но также имеет некоторые существенные недостатки. В первую очередь, в силу того, что параметр m считается постоянным, что противоречит свойствам реальных сетей.

На основе проведённого анализа, очевидно, что применение концепции эластичности к MDA-модели позволит соединить достоинства обоих подходов. Построение модели эластичной сети с присоединением, управляемым посредничеством (EMDA) и исследование её свойств является целью настоящей работы.

III. Модель эластичной сети с присоединением, управляемым посредничеством

Согласно проведённому анализу предметной области, мы предлагаем [14] изменить правило MDA, применив в качестве управляющего параметра фактор копирования (q) – вероятность, с которой узел-сосед посредника связывается с новым узлом. Также установим, что узел-посредник сам всегда связывается с новым узлом (рис. 1).

Таким образом, количество связей нового узла является не константой (как в правиле MDA), а равно $m = 1 + q \cdot k_{med}$ (где k_{med} – степень узла-посредника). Тогда математическое ожидание этой величины составляет

$$E\{m(n)\} = 1 + q \cdot E\{k_{med}\} = 1 + q \cdot \bar{k}(n) = 1 + q\frac{L(n)}{n}. \quad (6)$$

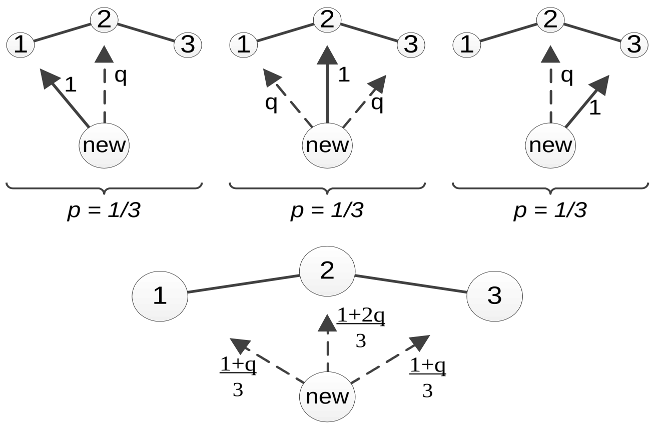

Рис. 1. Вероятности присоединения нового узла согласно предложенному эластичному правилу MDA.

Уравнение динамики роста количества связей сети:

$$L(n+1) - L(n) = 2(1 + q \cdot L(n)/n). \quad (7)$$

С учётом начальных условий $L(1) = 0$, из (7) получим зависимость общего количества связей сети $L(n)$ и средней степени её узлов $\bar{k}(n) = L(n)/n$ от размера сети:

$$\bar{k}(n) = \frac{2}{2q-1}\left(\frac{\Gamma(n+2q)}{\Gamma(n+1)\Gamma(1+2q)} - 1\right). \quad (8)$$

Сравнив (5) и (8), можно убедиться, что предложенная модель ЕМДА является эластичной с показателем эластичности $\lambda = 2q$. При $0 < q < 0.5$ средняя степень узлов возрастает до конечного предела $\bar{k}_{\lim} = 2/(1-2q)$, а при $0.5 < q < 1$ – растёт неограниченно по асимптотически степенному закону $\bar{k}(n) \propto n^{2q-1}$.

Для нахождения распределения узлов по количеству связей, запишем уравнение динамики математического ожидания числа связей у произвольного узла i:

$$k_i(n+1) - k_i(n) = \frac{1 + q \cdot k_i(n)}{n}. \quad (9)$$

Общее решение (9) имеет вид

$$k_i(n) = C_i \frac{\Gamma(n+q)}{\Gamma(n)} - \frac{1}{q}. \quad (10)$$

Коэффициенты C_i могут быть найдены из условия, что ожидаемое значение степени узла i в момент времени i равно ожидаемому количеству рёбер, добавленных на шаге $i-1$ (6), т.е. $k_i(i) = E\{m(i-1)\}$. Следовательно,

$$C_i = \frac{1}{2q-1}\left(\frac{\Gamma(i-1+2q)}{\Gamma(i+q)\Gamma(2q)} - \frac{\Gamma(i)}{\Gamma(i+q)}\frac{1-q}{q}\right). \quad (11)$$

Согласно (11) при $i \gg 1$ этот коэффициент асимптотически растёт согласно степенному закону от i:

$$C_i \propto \begin{cases} \frac{1}{(2q-1)\Gamma(2q)} i^{q-1}, & 0.5 < q < 1 \\ \frac{q-1}{(2q-1)q} \cdot i^{-q}, & 0 < q < 0.5 \end{cases} \quad (12)$$

В силу монотонности зависимостей (11)-(12) относительно номера узла, его можно рассматривать как ранг распределения узлов по числу связей. Таким образом, ранговое распределение узлов асимптотически следует степенному закону с показателем

$$\beta = \min\{q, 1-q\}. \quad (13)$$

Степенному закону рангового распределения степеней узлов с показателем (13) соответствует степенной закон частотного распределения (2) с показателем скейлинга

$$\gamma = 1 + 1/\beta = 1 + 1/\min\{q, 1-q\}. \tag{14}$$

Таким образом, в результате теоретического исследования показано, что предложенная EMDA-модель порождает сети, являющиеся эластичными с $\lambda = 2q$ и масштабно-инвариантными с показателем скейлинга (14).

IV. Численное моделирование EMDA-сети

На первом этапе численного эксперимента исследовалась динамика роста числа связей у узлов EMDA-сети. Изначально сеть состоит из двух узлов, соединённых ребром. На каждом шаге добавлялся один узел и связывался с существующими согласно предложенному правилу. Численные результаты и соответствующие теоретические зависимости средней степени узлов от размера сети и фактора копирования показаны на рис. 2.

Затем было исследовано ранговое распределение степеней узлов в зависимости от q при фиксированном размере сети $n = 4096$. Полученные численные зависимости вместе с теоретическими (15)-(16) показаны на рис.3.

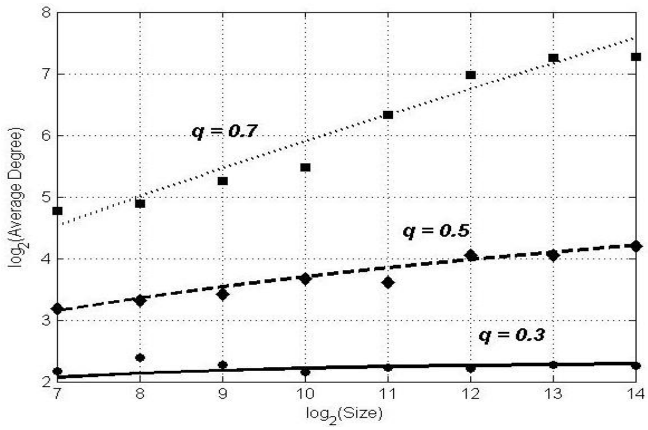

Рис. 2. Зависимость средней степени узлов от размера сети и фактора копирования

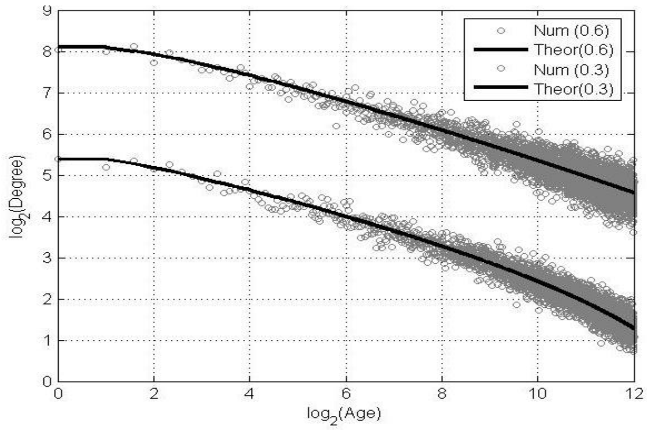

Рис. 3. Ранговое распределение степени узлов при q = 0.6 и q = 0.3

V. Выводы

Рассмотрена проблема моделирования МИС. Проведён анализ существующих моделей МИС и соответствующих правил присоединения узлов. Выявлено, что основными недостатками традиционных моделей являются линейная зависимость количества связей в сети от числа узлов и явное использование информации о свойствах существующих узлов сети. Модель эластичной сети и модель MDA позволяют исправить указанные недостатки, однако, лишь по отдельности.

Предложена модель эластичной сети с присоединением, управляемым посредничеством (EMDA-модель). Полученная модель сочетает в себе достоинства обоих родительских моделей: является эластичной и не использует статистику узлов.

Аналитически получена зависимость средней степени узлов от фактора копирования и размера сети, а также зависимость показателя рангового распределения степени узлов от фактора копирования. Результаты проведённого численного моделирования полностью согласуются с теоретическими.

Анализ свойств ассортативности предложенной EMDA-модели МИС рассматривается как перспективное направление дальнейших исследований.

Литература References

[1] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law", *Contemporary Physics*, 2005, 46(5). p.323-351.

[2] S. N. Dorogovtsev, J. F. F. Mendes, "Evolution of Networks: From Biological Networks to the Internet and WWW", Oxford, USA: Oxford University Press, 2003. – 280 p.

[3] R. Albert, A.-L. Barabási, "Statistical mechanics of complex networks", *Rev. Mod. Phys.*, 2002, - V. 74. - p. 42-97.

[4] G. Bianconi, A.-L. Barabási, "Competition and multiscaling in evolving networks", *Europhysics Letters*, 2001, 54 (4):436–442.

[5] G. Caldarelli, Guido; Catanzaro, Michele (2012). Networks: A Very Short Introduction. Oxford University Press. p. 78

[6] P. L. Krapivsky, S. Redner; F. Leyvraz, "Connectivity of Growing Random Networks", *Phys. Rev. Lett.*, 2000, 85: 4629–4632.

[7] R. Noldus, P. Van Mieghem, "Assortativity in complex networks", *J. Complex Networks*, 2015, vol. 3, pp. 507-542.

[8] Ch. Dangalchev, "Generation models for scale-free networks", *Physica A*, 2004, 338, 659

[9] A. D. Broido, A. Clauset, "Scale-free networks are rare". *Nat Commun* 10, 1017 (2019). https://doi.org/10.1038/s41467-019-08746-5

[10] M. K. Hassan, L. Islam, S. A. Haque, "Degree distribution, rank-size distribution, and leadership persistence in mediation-driven attachment networks". *Physica A: Stat. Mech. and its Appl.* 2017. 469 (2017): 23–30

[11] V. L. Shergin, L. E. Chala and S. G. Udovenko, "Fractal dimension of infinitely growing discrete sets," 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Slavske, 2018, pp. 259-263.

[12] V. Shergin, L. Chala, "The concept of elasticity of scale-free networks," 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 257-260.

[13] V. Shergin, S. Udovenko, L. Chala, "Assortativity Properties of Barabási-Albert Networks," In *Data-Centric Business and Application*. Springer, 2021.

[14] V. Shergin, L. Chala, S. Udovenko, M.Pogurskaya, "Elastic Scale-Free Networks Model Based on the Mediaton-Driven Attachment Rule", 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2020, in press.