

Министерство образования и науки Украины  
Российская ассоциация искусственного интеллекта  
Национальный технический университет Украины "КПИ"  
Факультет прикладной математики  
Институт прикладного системного анализа  
Издательство "Просвіта"

***Международная научная конференция  
имени Т. А. Таран***

**Интеллектуальный анализ информации  
ИАИ-2014**

**Киев, 14-16 мая 2014 г.**

**Сборник трудов**

Рекомендовано ученым советом  
факультета прикладной математики

**«Просвіта»  
Киев  
2014**

УДК 004.8/.9+001.102](06)

ББК 32.973я43+73я43

I-73

Редакционная коллегия:

д.т.н., проф. Валькман Ю.Р., д.т.н., проф. Гладун В.П., д.т.н., проф. Голенков В.В., к.т.н., доц. Григорьев А.В., д.т.н., проф. Дичка И.А., д.ф.-м.н., проф. Донской В.И., д.т.н., проф., академик НАНУ Згуровский М.З., д.т.н., проф. Кузнецов О.П., д.т.н., проф. Молчанов А. А., д.ф.-м.н., проф. Осипов Г.С., д.т.н., проф. Стефанюк В. Л., д.т.н., проф. Тарасенко В.П., д.т.н., проф. Финн В.К., д.т.н., проф. Хорошевский В.Ф., к.т.н., доц. Чертов О.Р.

Главный редактор к. т. н. Сирота С. В.

Ответственный редактор Темникова Е.Л.

Ответственный за выпуск Копычко С. Н.

Рекомендовано ученым советом факультета прикладной математики НТУУ

“КПИ” протокол №10 от 26 мая 2014 г.

«Интеллектуальный анализ информации», междунар. науч. конф. им. Т.А.Таран (14 ; 2014 ; Киев).

I-73 Международная научная конференция имени Т. А. Таран "Интеллектуальный анализ информации" ИАИ-2014, Киев, 14–16 мая 2014 г. : сб. тр.– К. : Просвіта, 2014. – 260 с. : ил.

**ISBN 978–617–7010–03–5**

В сборнике опубликованы доклады, представленные на конференции по следующим направлениям: интеллектуальный анализ данных и машинное обучение, правдоподобные рассуждения, интеллектуальные методы классификации, интеллектуальный поиск и анализ информации в локальных и глобальных сетях, прикладные системы интеллектуальных данных, интеллектуальный анализ данных в социальной сфере и гуманитарных исследованиях, инструментальные средства интеллектуального поиска и анализа информации.

УДК 004.8/.9+001.102](06)

ББК 32.973я43+73я43

Використання матеріалів збірки можливе за умови обов'язкового посилання.

Использование материалов сборника возможно при условии обязательной ссылки.

© ПТФ Просвіта, 2014

**ISBN 978–617–7010–03–5**

УДК 681.3:004.7

## Эффект «взрывной перколяции» в контентной сети

*Снарский А.А.<sup>1,2</sup>, д.ф.-м.н, проф., Ландэ Д.В.<sup>2,1</sup>, д.т.н., с.н.с.*

*<sup>1</sup>НТУУ «Киевский политехнический институт», Киев, Украина*

*<sup>2</sup>Институт проблем регистрации информации НАН Украины,  
Киев, Украина*

*dwlande@gmail.com, asnarskii@gmail.com*

Предлагается метод построения контентной сети (CNC, Content Network Connections), являющийся развитием алгоритма «взрывной перколяции». Изучается поведение данной растущей сети, а также фазовый переход первого рода, возникающий при построении такой сети.

### Введение

Исследование динамики сложных сетей (Complex Networks) [1-2], изменения количества связей, уровней кластерности, других параметров, является актуальной научной и практической задачей, так как большинство реальных сетей (Интернет, WWW, социальные, биологические и др. сети) являются динамичными, растущими. Для таких сетей, в которых течением времени увеличивается число узлов и/или связей, очень важным является эффект возникновения так называемого гигантского кластера (GC, Giant Cluster), при котором количество узлов самого большого кластера в ней становится равным по порядку общему количеству узлов. Условия возникновения GC и его свойства хорошо изучены для перколяционных [3] и некоторых других сложных сетей, например, для сетей Эрдоша-Реньи. В частности, для таких сетей условие возникновения GC определяется степенями узлов и задается критерием Моллоя-Рида [4]:  $\langle k^2 \rangle - 2\langle k \rangle \geq 0$ , где  $k$  – степень узла, а  $\langle \dots \rangle$  означает усреднение по всем узлам.

На фазе возникновения и роста GC, как в перколяционных, так и в ER и в других сложных сетях, например, в безмасштабных (SF, Scale Free Networks) [1], некоторые параметры этих сетей (например, размеры максимальных групп связанных узлов – кластеров), ведут себя аналогично параметрам порядков в теории фазовых переходов второго рода [5].

## «Взрывная перколяция»

В работе [6] был предложен алгоритм построения растущей сети, который приводит к так называемой взрывной перколяции (EP, Explosive Percolation). В отличие от алгоритма Эрдоша-Реньи, т.е. стандартного случайного порождения (путем «набрасывания») новых связей, этот алгоритм предполагает альтернативу. В соответствии с ним на каждом шагу случайным образом порождается не одна, а две связи, из которых только одна остается (реализуется), а вторая отбрасывается. Критерий отбора реализованной связи определяется сравнением произведений размеров кластеров (количества узлов), соединяемых связями. Из двух выбирается та связь, для которой это произведение наименьшее. Как оказалось, при реализации такого алгоритма, формирование и рост GC принципиально отличается, его размер изменяется скачкообразно, аналогом чего является фазовый переход первого рода, а не второго.

## Модель контентной сети

Авторами предлагается следующая модель контентной сети. Предполагается, что сеть состоит из  $N$  узлов. Существует  $M$  различных документов, отдельные экземпляры которых (копии) распределены в узлах контентной сети. Пусть  $i$ -й узел сети содержит  $n_i$  различных документов, причем только по одному экземпляру каждого из них. Узлы можно интерпретировать как библиотеки, веб-сайты, файловые хранилища. Соответственно, документы можно интерпретировать как книги, веб-страницы, файлы, фрагменты файлов и т.п. Предполагается, что распределение количества документов по узлам соответствует закономерности Парето (аналогично распределению богатства в социальных системах), т.е. степенное.

В этом случае для построения растущей сети предложен новый алгоритм порождения связей в контентной сети (CNC, Content Network Connections), являющийся развитием алгоритма EP. В соответствии с ним на каждом шагу случайным образом из всего многообразия документов задаются два различных документа, а затем случайным образом выбираются две пары узлов, содержащих данные документы. В данном случае если хотя бы один выбранный узел не содержит необходимого документа, то происходит случайный выбор следующего узла. Эта процедура выполняется циклически до достижения цели, естественно, должны учитываться условия выхода из «зацикливания», например, путем перехода к выбору другой пары документов.

После этого между выбранными узлами, содержащими разные документы, также как и в [6] устанавливаются две связи, из которых также остается (реализуется) только одна, а вторая отбрасывается. Критерий отбора реализованной связи в этом случае определялся сравнением общего количества документов, содержащихся в кластерах, узлы которых соединяются связями с выбранными узлами. Из двух связей выбирается та, для которой сумма документов в соответствующих кластерах оказывается наименьшей. Таким образом, документы являются основой построения связей между узлами-библиотеками. На рис.1 представлена абстрактная модель сети из 12 узлов-библиотек, содержащих документы, обозначенные различными фигурами (закрашенными и не закрашенными треугольниками, квадратами, ромбами и т.д.).

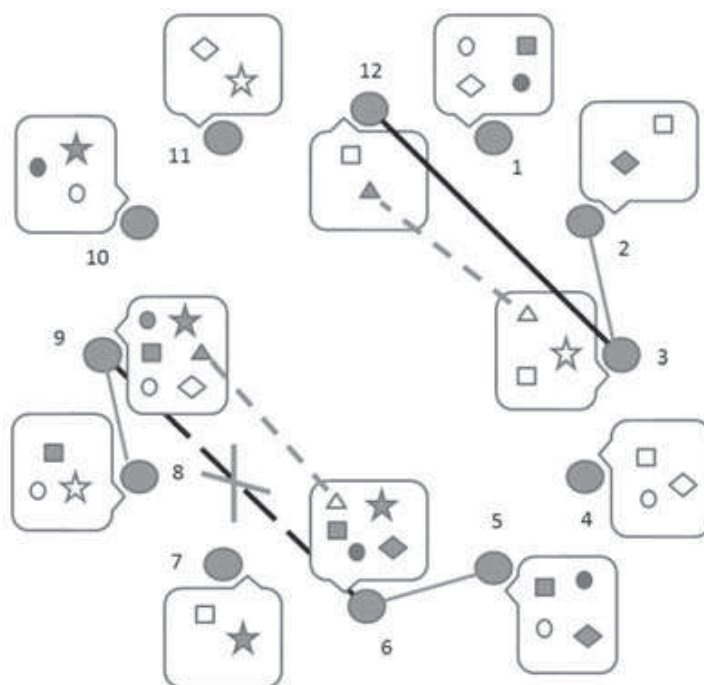


Рис. 1 – Установление связи в модели контентной сети

Здесь первое условие, задаваемое при построении новой связи, заключается в установлении связей между узлами, содержащими документы-треугольники – закрашенный и не закрашенный. Случайным образом выбраны две пары узлов, содержащих требуемые документы: первая – 12 и 3, вторая – 6 и 9. Очевидно, кластер первой пары, содержащий документы из 12-го, 3-го и 2-го узлов (всего 7 документов), значительно меньше по количеству документов кластеру второй пары, содержащей документы из 6-го,

9-го, 5-го и 8-го узлов (всего 18 документов). Таким образом остается только первая связь (между 12-м и 3-м узлами), а вторая отбрасывается.

Выбор двух документов в алгоритме CNC может интерпретироваться, например, как установление связей между библиотеками при подборе многотомного издания, или как сборка целого файла при объединении различных фрагментов, расположенных в различных узлах пиринговой сети.

## Изменение размера гигантского кластера

Авторами были проведены эксперименты для различных  $N$  и  $M$  с целью определения зависимости числа узлов (библиотек) наибольшего кластера  $p$  от числа связей для предложенного алгоритма CNC. На рис. 2 показана типичная зависимость для эталонного примера, соответствующего алгоритму Эрдоша-Реньи, когда устанавливается связь между первыми выбранными узлами, содержащими требуемую пару документов (рис 2a), и для CNC (рис 2b) при  $M = N = 250$ .

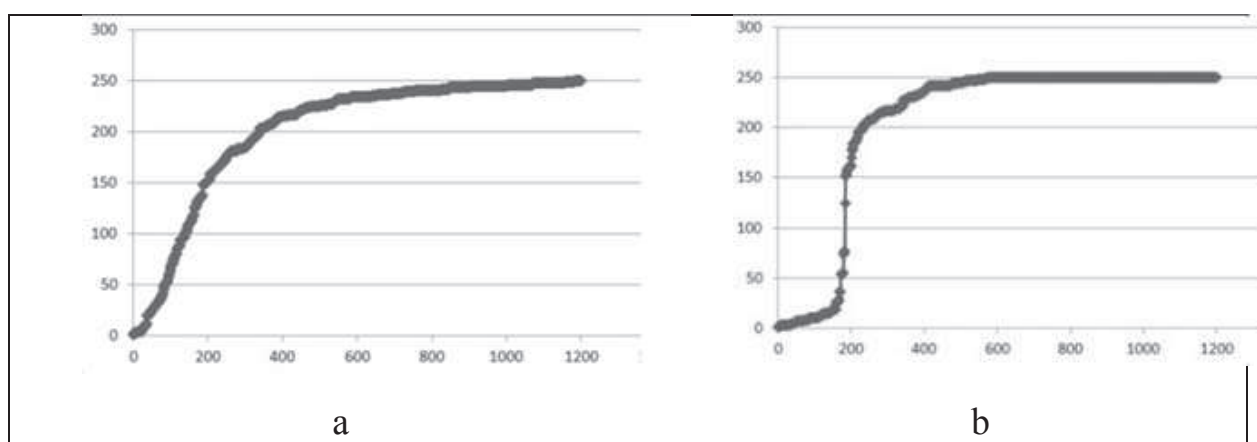


Рис. 2 – Зависимость размера наибольшего кластера от числа связей при  $M = N = 250$

## Оценка величины скачка

Как видно на рис. 2, при определенном значении  $p$ , в случае предложенного алгоритма CNC происходит скачкообразный рост наибольшего кластера. Усредненная по реализациям зависимости величины скачка  $\eta$  от

размера контентной сети  $N$  в случае  $M = N$  приведена на рис. 3. Как показывают измерения, тренд этой зависимости близок к линейному.

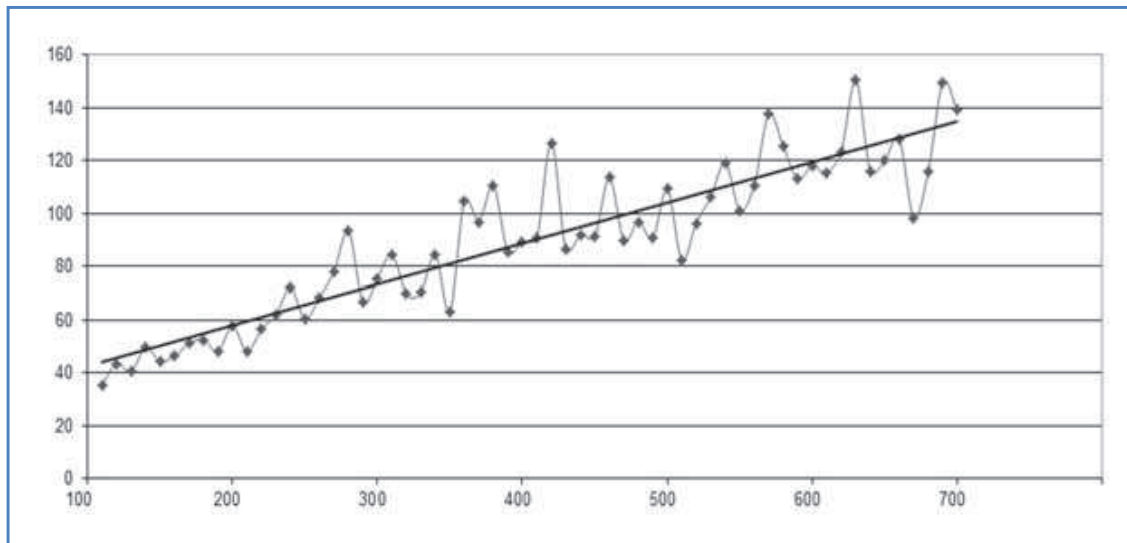


Рис. 3 – Зависимость величины скачка  $\eta$  от размера контентной сети  $N$

Пусть  $C$  – размер самой большой компоненты связности,  $t_0$  – наибольшее количество связей в растущей контентной сети, при которых  $C < \sqrt{N}$ , а  $t_1$  – наименьшее количество связей, при которых  $C > N/2$ . В случае непрерывных переходов, разность  $D = t_0 - t_1$  линейно зависит от  $N$ , в частности,  $D \approx 0.19N$  для сетей Эрдоша-Реньи. В отличие от данного поведения, в случае CNC зависимость не является линейной, а может аппроксимироваться зависимостью  $D \sim AN^\alpha$ , ( $A \approx 0,84$ ,  $\alpha \approx 0,79$ ), что соответствует поведению, описанному в [6].

## Выводы

При реализации алгоритма CNC, образование и рост GC соответствует поведению, наблюдаемому при реализации алгоритма EP, его величина также изменяется скачкообразно, что также соответствует фазовому переходу первого рода, т. е. так называемой «взрывной перколяции». Как и в [6] небольшие изменения в правилах формирования связей приводит к принципиально новому характеру в протекании процесса формирования сети, что необходимо учитывать в процессе формирования и управления контентными сетями.

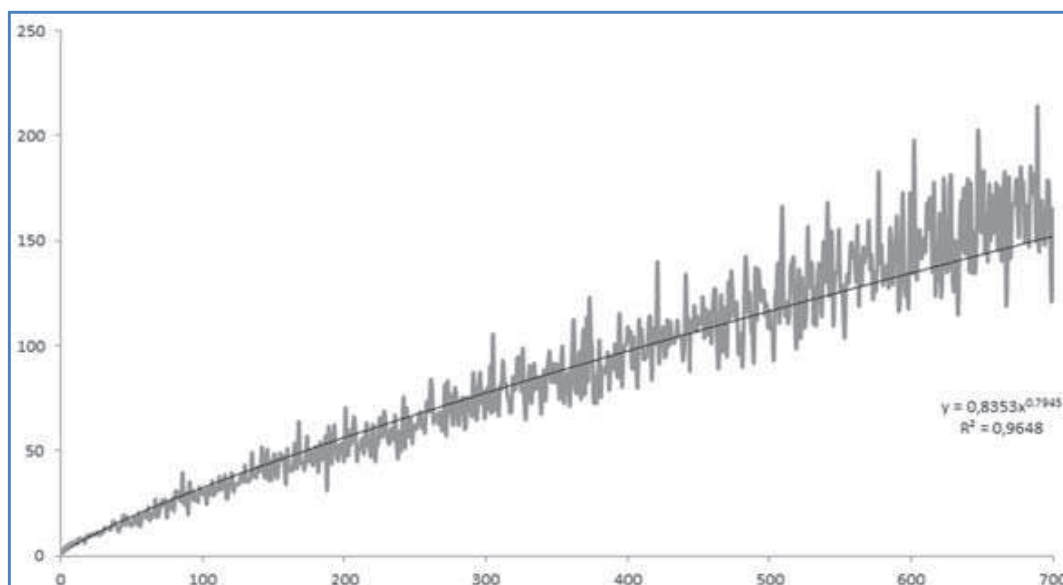


Рис. 4 – Зависимость величины  $D$  от размера контентной сети  $N$

## Литература

1. *Albert R., Barabási A.-L.* Statistical mechanics of complex networks // *Rev. Mod. Phys.*, 2002 – 74. – P. 47-97.
2. *Dorogovtsev S.N., Mendes J.F.F.* Evolution of Networks. From Biological Nets to the Internet and WWW. – Oxford: Oxford University Press, 2003. – 280 pp.
3. *Stauffer D., Aharony A.* Introduction to Percolation Theory. London: Taylor&Francis, 1992. – 180 pp.
4. Dorogovtsev S.N., Goltsev A.V., Mendes J.F.F. Critical phenomena in complex networks // *Rev. Mod. Phys.*, 2008. – 80. – P. 1275-1335.
5. *Ландау Л.Д., Лифшиц Е.М.* Статистическая физика. Ч.1. – М.: Наука, 1976. – 584 с.
6. *Achlioptas D., D'Souza R.H., Spenser J.* Explosive Percolation in Random Networks // *Science*, 2009. – 323. – P. 1453-1455.



# Содержание

<b>Святогор Л. А., Величко В. Ю.</b> Памяти В. П. Гладуна .....	4
<b>Валькман Ю. Р.</b> О гештальтах в мышлении: определения, классификация и свойства .....	11
<b>Голенков В. В., Гулякина Н. А.</b> Смысловые модели представления и обработки знаний .....	23
<b>Кузнецов О. П.</b> Еще одна концепция когнитивной семантики .....	28
<b>Бабенко А. Е.</b> Интеллектуализация удаленного процесса обучения иностранному языку .....	37
<b>Байбуз О. Г., Луценко О.П.</b> Модель оцінювання ймовірності виникнення розладання у нестационарних процесах з множинними розладаннями .....	44
<b>Боргест Н. М., Бурдюгова В. В.</b> Разработка онтологии документооборота выполнения заказа на предприятии .....	49
<b>Боргест Н. М., Сокова Н. А.</b> Разработка онтологии нештатных ситуаций на космической станции .....	54
<b>Валькман Ю. Р., Жолнарская А. А.</b> Анализ средств организации хранилищ данных в информационно- аналитических системах статистических данных .....	58
<b>Вовк О.Л.</b> Моделирование зрительного анализатора человека для выделения контуров объектов изображений .....	71
<b>Гаврилов С.В., Дробышев Ю.П.</b> Класифікація руху людини з використанням методу швидкого відключення для рекурентних нейронних мереж .....	77
<b>Дробязко І.П., Клятченко Я.М., Тарасенко-Клятченко О.В., Тарасенко В.П., Тесленко О.К.</b> Методика аналізу та оцінки достовірності функціонування логічних мереж на основі суперпозицій булевих функцій .....	83

<b>Yu.Zorin</b>	
An Artificial Bee Colony algorithm for data clustering.....	92
<b>Жилякова Л.Ю.</b>	
Сетевая парадигма в когнитивных науках .....	99
<b>Карасюк В.В.</b>	
Представлення правових знань на основі онтологій. ....	107
<b>Комаров-Ермолов А. И.</b>	
Новые подходы к извлечению знания: архитектоническая парадигма .....	114
<b>Копичко О.С., Рибачок Г.І., Копичко С.М., Мороченець О.І.</b>	
Інтелектуальна система автоматизованої ідентифікації клієнтів мобільної мережі.....	121
<b>Кравець П.І., Шимкович В.М., Зубенко Г.А.</b>	
Моделі штучних нейронних мереж при їх апаратно-програмній реалізації на FPGA.....	127
<b>Кулаков Ю. А., Воротников В. В., Данелюк Д.О.</b>	
Межкластерная многопутевая маршрутизация в сложной сети.....	132
<b>Литвинов В.А., Майстренко С. Я., Хурцилава К.В.</b>	
К оценке качества базового словаря в системе автоматического обнаружения ошибок пользователя и возможностей его улучшения ...	139
<b>Олефір О.С., Гусак Б.І.</b>	
Прогнозування волатильності акцій на основі гетероскедастичних моделей. ....	144
<b>Путренко В.В.</b>	
Інтелектуальний аналіз небезпеки виникнення природних пожеж на основі геоінформаційних технологій.....	150
<b>Рогущина Ю.В.</b>	
Разработка методов самообучения поисково-рекомендующей системы с целью усовершенствования ее поведения .....	155
<b>Савеленко О.К.</b>	
Кластерний аналіз в задачах обробки інформації.....	162
<b>Савчук О.В., Кривенко К.С.</b>	
Інтелектуальний аналіз діагностичної інформації складних технічних комплексів .....	172

<b>Северин Н. В., Леоненко Л. Л.</b> О некоторых методах “борьбы с неустранимой неоднозначностью” в сетях телекоммуникаций .....	178
<b>Сирота С.В., Величко В.Ю., Ліскін В.О.</b> Застосування паралельних обчислень для побудови зростаючих пірамідальних мереж. ....	185
<b>Снарский А.А., Ландэ Д.В.</b> Эффект «взрывной перколяции» в контентной сети. ....	190
<b>Танченко С. С., Титенко С. В., Гагарін О. О.</b> Усунення мовної неузгодженості в тестових завданнях, згенерованих на основі понятійно-тезисної моделі. ....	196
<b>Темников В.А., Темникова Е.Л., Конфорович И.А.</b> Построение системы аутентификации операторов и распознавания слов по голосу .....	201
<b>Терлецький Д.О.</b> Представлення знань за допомогою об’єктно-орієнтованих динамічних мереж .....	207
<b>Титенко С.В.</b> Інформаційно-логічні та архітектурні засади універсальних систем керування web-контентом .....	214
<b>Трифонов Е.О.</b> Сжатие данных на основе бинарного дерева.....	221
<b>Хала К. А.</b> Функциональность и методы прикладной правовой онтологии. ....	227
<b>Хіміченко І.В.</b> Щодо деяких методів підвищення часової ефективності при фрактальному стисненні зображень .....	234
<b>Чертов О.Р., Тавров Д.Ю.</b> Метод штрафных функций для учета ограничений на решения задачи обеспечения групповой анонимности .....	241
<b>Казаков А.И., Обухов Л.И., Ткаченко Н.А., Хоружий В.П., Хоружая И.А.</b> К 15-летию Киевской лаборатории искусственного интеллекта.....	250