



МИБ - 2007

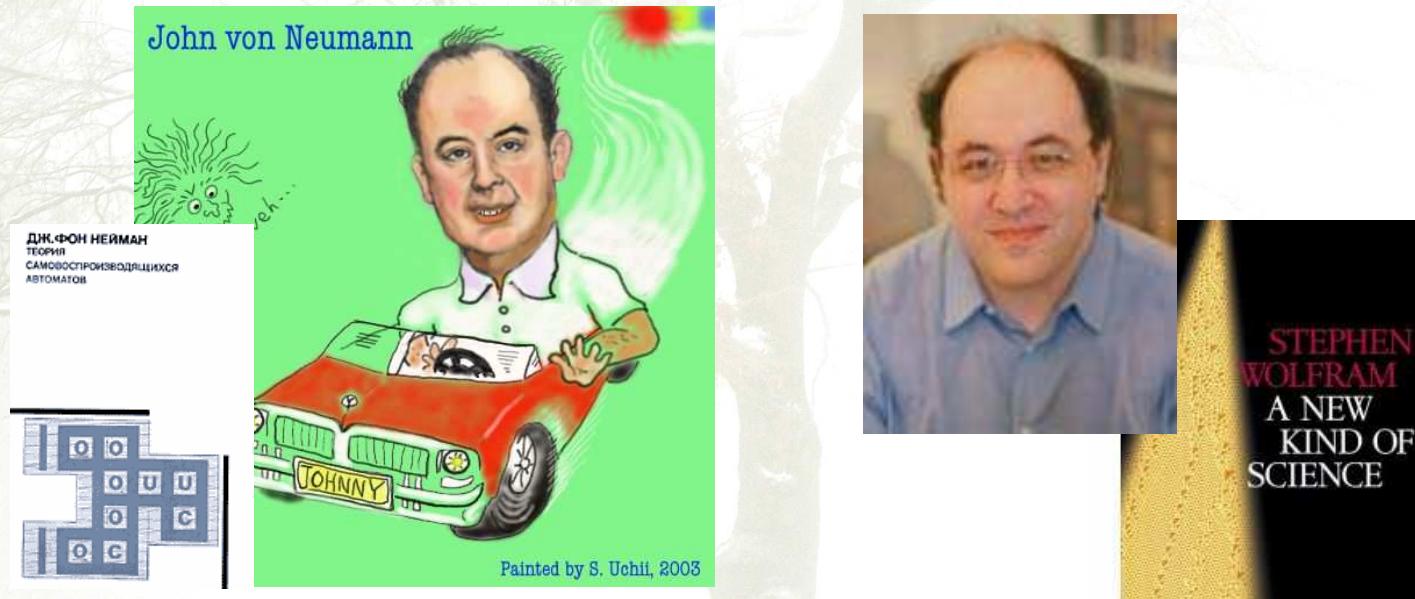
МОДЕЛЬ ДИФУЗИИ ИНФОРМАЦИИ

Д.В. Ландэ

ИЦ «ЭЛВИСТИ», НТУУ «КПИ»

Моделирование на основе концепции клеточных автоматов

Разнообразие поведения публикаций по различным тематикам и сложность взаимного влияния различных публикаций заставляют искать новые, ранее неизвестные в этой области методы. По-видимому, на данном этапе в области моделирования сложных информационных процессов успех может быть достигнут лишь путем синтеза достаточно простых алгоритмов и концепций.



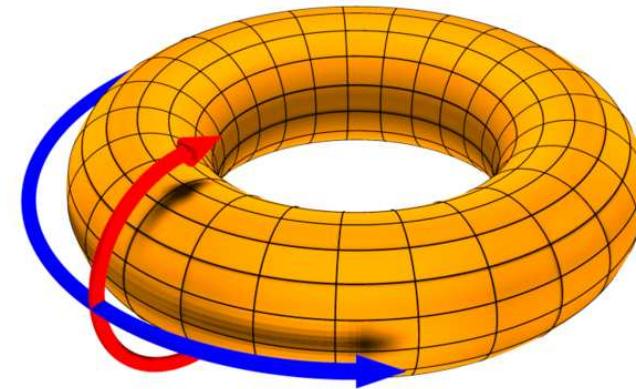
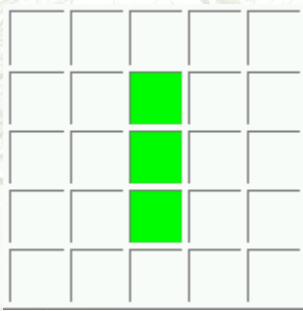
Перспективными в этой области можно признать теорию клеточных автоматов, впервые предложенную более тридцати лет тому назад Дж. фон Нейманом и развитую С. Уолфрамом в фундаментальной монографии.

Клеточные автоматы

Клеточный автомат представляет собой дискретную динамическую систему, совокупность одинаковых клеток, образом соединенных между собой. Все клетки образуют сеть клеточных автоматов. Состояние каждой клетки определяются состоянием клеток, входящих в ее локальную окрестность. Окрестностью конечного автомата с номером j называется множество его «ближайших соседей». Состояние j -го клеточного автомата в момент времени $t + 1$, таким образом, определяется следующим образом:

$$y_j(t+1) = F(y_j(t), O(j), t),$$

где F - некоторое правило, которое можно выразить, например, языком булевой алгебры, $O(j)$ - окрестность, t - такт.



Клеточные автоматы в традиционном понимании удовлетворяют таким правилам:

- изменение значений всех клеток происходит одновременно (единица измерения времени - такт);
- сеть клеточных автоматов является однородной, т.е. правила изменения состояний для всех клеток одинаковы;
- на клетку могут повлиять лишь клетки из ее локальной окрестности;
- множество состояний клетки конечно.

Правила «диффузии информации»

Динамике распространения информации присущи некоторые свойства, которые были учтены в модели. Предполагается, что клетка может быть в одном из трех состояний: 1 - «свежая новость» (клетка окрашивается в черный цвет); 2 - новость, устаревшая, но сохраненная в виде сведений (серая клетка); 3 - клетка не имеет информации, переданной новостным сообщением (клетка белая, информация не дошла или уже забыта).

Правила распространения новостей следующие:

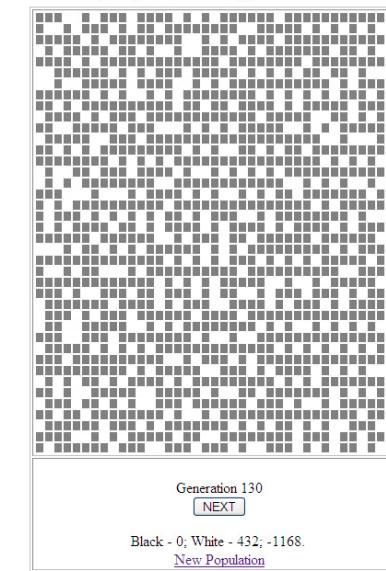
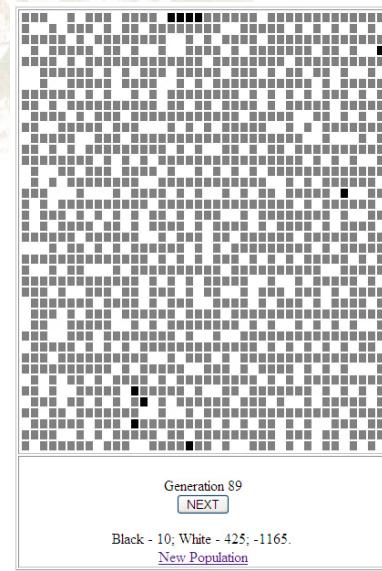
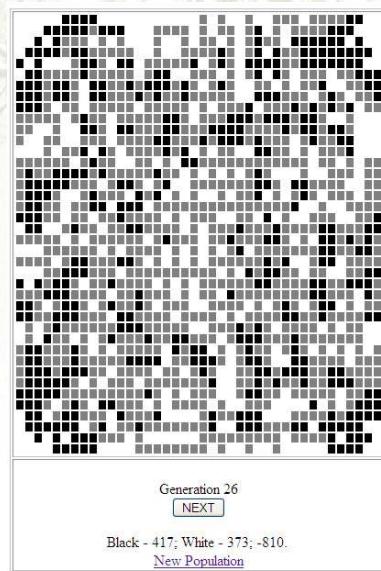
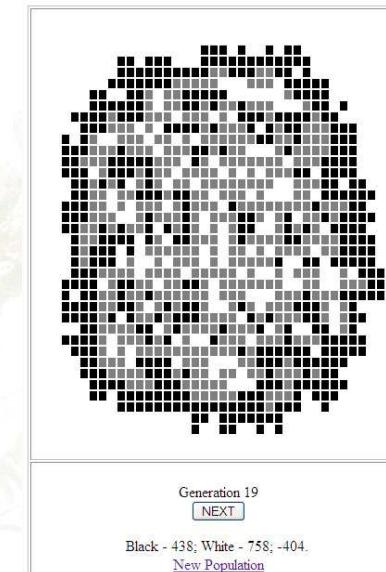
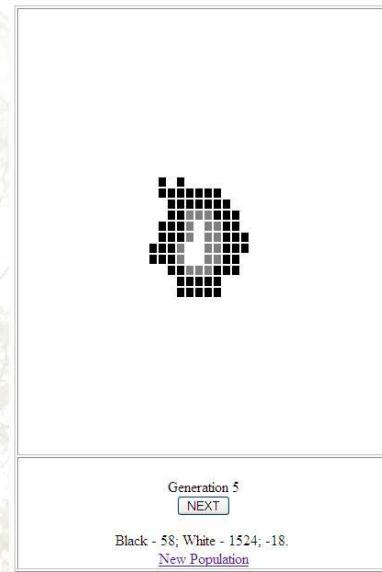
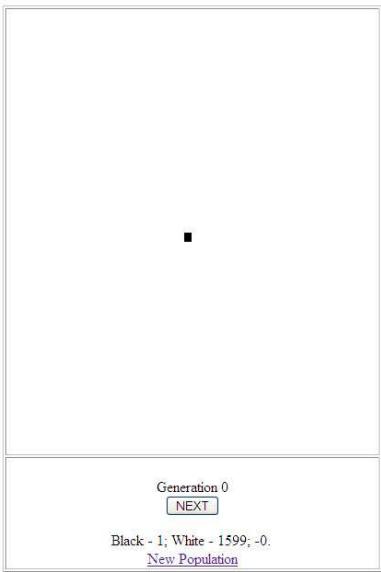
- изначально все поле состоит из белых клеток за исключением одной - черной, которая первой «приняла» новость;
- белая клетка может перекрашиваться только в черный цвет или оставаться белой (она может получать новость или оставаться «в неведении»);
- белая клетка перекрашивается, если выполняется условие:

$$pt > 1,$$

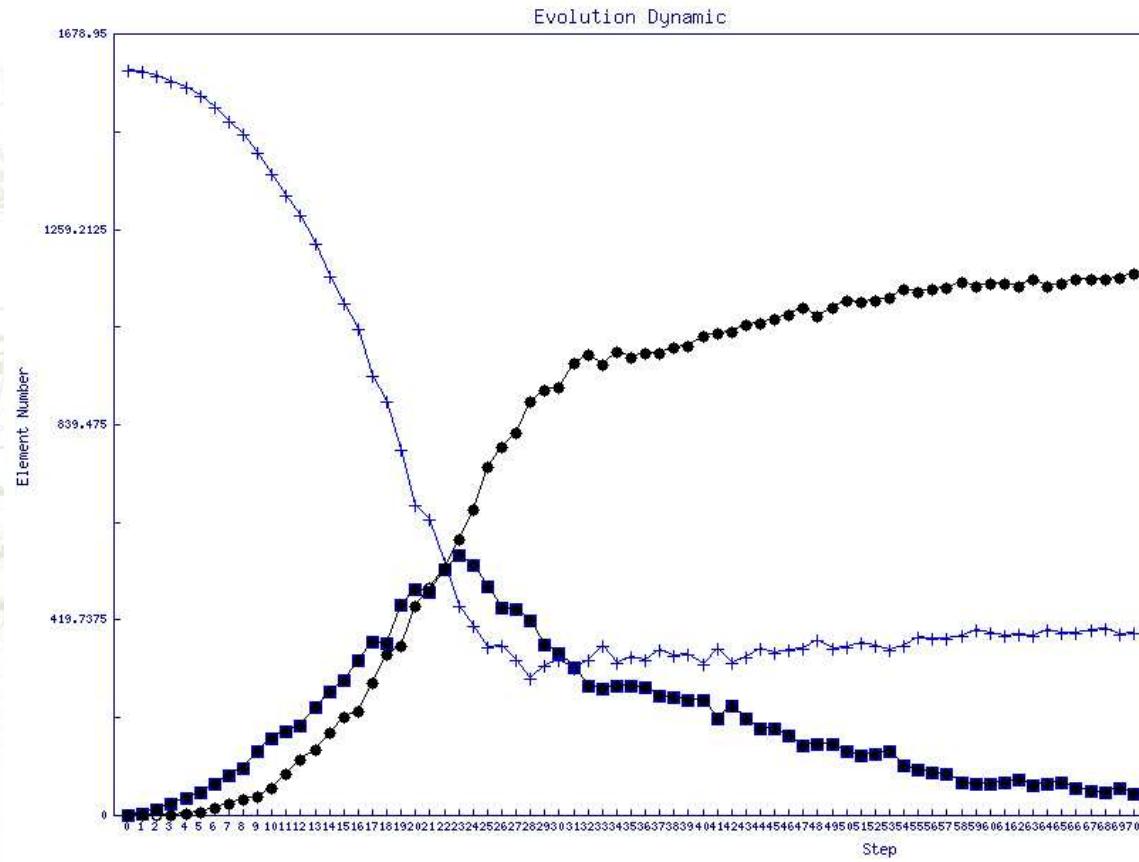
p - вероятность «принятия информации», t - количество «проинформированных» соседей;

- если клетка черная, а вокруг нее исключительно черные и серые, то она перекрашивается в серый цвет (новость устаревает, но сохраняется как сведения);
- если клетка серая, а вокруг нее исключительно серые и черные, то она перекрашивается в белый цвет (происходит «забывание» сведений при их общеизвестности).

Процесс эволюции системы клеточных автоматов



Количество клеток каждого типа в зависимости от такта



Белые клетки - (+); серые клетки - (●); черные клетки - (■)

Аналитическое моделирование динамики «диффузии» информации

$$x_g = f(t, \tau_g, \gamma_g), \quad x_w = 1 - f(t, \tau_w, \gamma_w).$$

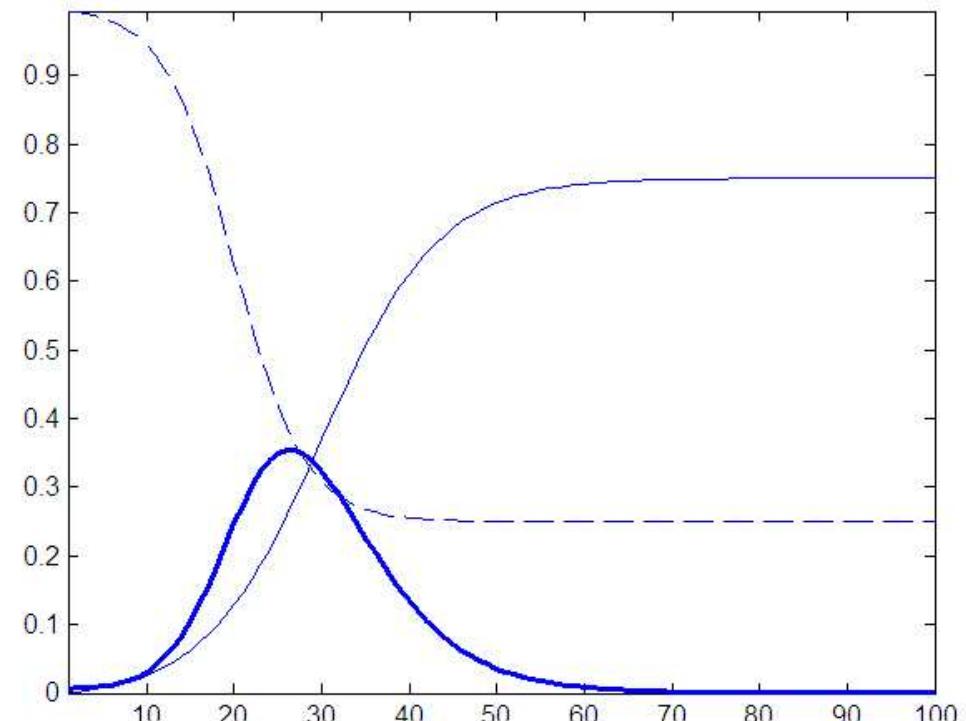
$$x_b = 1 - x_g - x_w = f(t, \tau_w, \gamma_w) - f(t, \tau_g, \gamma_g).$$

$$f(t, \tau, \gamma) = \frac{C}{1 + e^{\gamma(t-\tau)}},$$

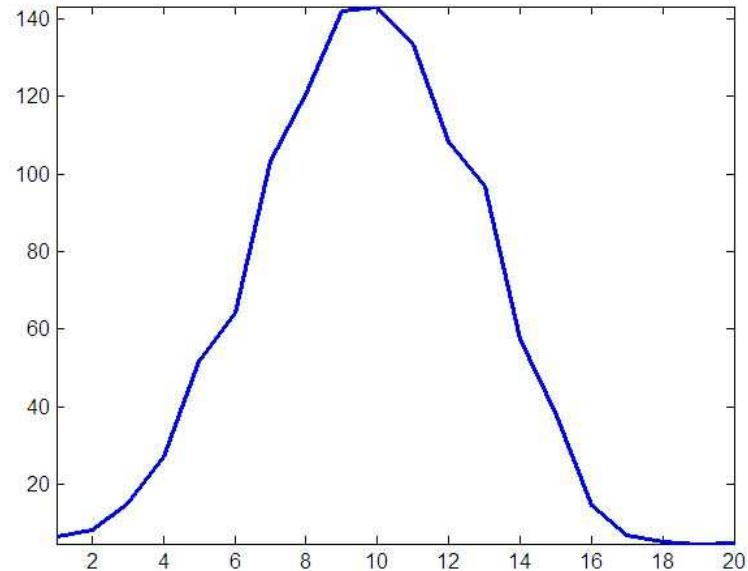
$$x_g = \frac{0.75}{1 + e^{-0.15(t-30)}};$$

$$x_w = 1 - \frac{0.75}{1 + e^{-0.25(t-20)}};$$

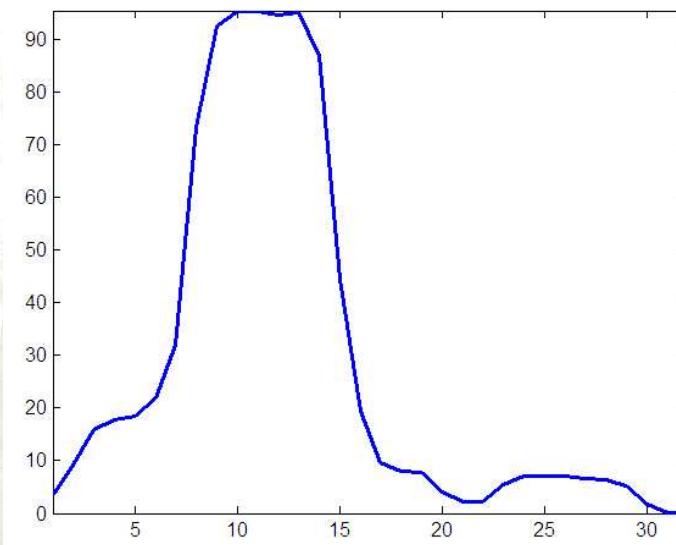
$$x_b = 0.75 \left(\frac{1}{1 + e^{-0.25(t-20)}} - \frac{1}{1 + e^{-0.15(t-30)}} \right).$$



Динамика реальных информационных потоков



Динамика понятия «ЦВК, БЮТ,
регистрация, но не референдум»



Динамика понятия «четвертый
указ», в системе интеграции
новостей *InfoStream*

Заключение

Многочисленные эксперименты с рассмотренной системой клеточных автоматов, доступной в настоящее время в сети Интернет по адресу <http://edu.infostream.ua/newsk.pl> показывают, что период его сходимости составляет от 80 до 150 тактов.

Полученная колоколообразная зависимость диффузии новостей на интернет-источниках (веб-сайтах) хорошо согласуется с реальным поведением тематических информационных потоков, а на локальных временных промежутках с известными моделями, например, экспоненциальной и логистической.

Литература

- 1) Ландэ Д.В., Фурашев В.Н., Брайчевский С.М., Григорьев А.Н. Основы моделирования и оценки электронных информационных потоков. - К.: Инжиниринг, 2006. - 176 с.
- 2) Нейман Дж. Теория самовоспроизводящихся автоматов. - М.: Мир, 1971. - 382 с.
- 3) S. Wolfram. *A New Kind of Science*. Champaign, IL: Wolfram Media Inc., 2002. - 1197 pp.
- 4) Плотинский Ю.М. Модели социальных процессов. - Изд. 2-е. - М.: Логос, 2001. - 296 с.
- 5) S.C. Bhargava, A. Kumar, A. Mukherjee. A stochastic cellular automata model of innovation diffusion // *Technological forecasting and social change*. - 1993. - Vol. 44. № 1. -P. 87-97.
- 6) Григорьев А.Н., Ландэ Д.В., и др. InfoStream. Мониторинг новостей из Интернет: технология, система, сервис. Научно-методическое пособие. - Киев: ООО "Старт-98", 2007. - 40 с.



**Спасибо за
внимание!**