# Information spirals in social network dynamics for large Internet databases

YouTube, financial markets, Cyber-risks, Social unrests, Conflicts



Didier Sornette Chair of Entrepreneurial Risks ETH Zurich (Swiss Federal Institute of Technology, Zurich) Department of Management, Technology and Economics http://www.er.ethz.ch/

Collaborators:

Y. Ageon, J. Andersen, R. Crane, F. Deschatres, T. Gilbert, A. Helmstetter, A. Johansen, N. Johnson, Y. Malevergne, T. Maillart, J.F. Muzy, S. Pillai, B. Roehner



## **Collective human activity**

## Not so simple trends

Products







Movies









Response function reveals information

# The big questions

How do we understand non-trivial herding?

How does the timing of individual activity get modified (renormalized) by interactions?

Model has 2 ingredients:

Timing of individual activity

**Epidemic Branching Process** 

## **Timing of human activity**



7



# A Shocking Look At...



NAMES OF A DESCRIPTION OF

2007.10.05

Riley Crane – DMTEC – Entrepreneurial Risks – rcrane@ethz.ch





#### Most Recent Page



MARIANUSPIPESTOBACCO

Views: 0





## **YouTube viewer dynamics**



#### Exogenous





#### Endogenous





## **Typical Relaxation Following Peak**



# How to explain these dynamics?

pdf of waiting times between cause and action and rate of activity

$$P(\tau) \sim \frac{1}{\tau^{1+\theta}}$$
 with  $\theta \ge 0$ 

- **Priori-queuing** (Abate & Whitt, 1997)
- Time-varying activity rate with feedback (Vasquez, 2007)
- Random walk crossing condition

## Epidemic cascades in social networks

# pdf of waiting times between cause and action and rate of activity





with  $\theta > 0$ 

$$P(\tau) \sim \frac{1}{\tau^{1+\theta}}$$

## Epidemic processes by word-of-mouth









### **Distribution of relaxation exponents**





Occurence (Critical)

- 1. Class 1 is defined by  $80\% \le F \le 100\%$ . 2. Class 2 is defined by 20% < F < 80%.





#### What about precursory information?







- Book, CD music sales...
- Internet searches
- Open source software projects
- Ethical Hacking security
- NATO (National Association of Theatre Owners)
- Recommendation Blogs

#### AMAZON BOOK SALES

- Amazon.com posts a "live" ranking of all its products
- Book ranks in the top 10,000 are updated every hour according to a secret weighting of recent sales and entire history

#### Top 100 Bestsellers Updated Hourly



 <u>The South Beach Diet</u> by Arthur Agatston (Author)
 Price: \$14.97 You Save: \$9.98 (40%) <u>Used & new</u> from \$13.24



2. <u>The Da Vinci Code</u> by Dan Brown (Author) Price: \$14.97 You Save: \$9.98 (40%) <u>Used & new</u> from \$10.80



3. <u>The Last Juror</u> by John Grisham (Author) Price: **\$19.57 You Save:** \$8.38 (30%)

4. <u>South Beach Diet Good</u> <u>Fats/Good Carbs Guide</u> by Arthur Agatston (Author)
Price: \$7.99 Used & new from \$7.80

## The Original "Crisis"

- On Friday January 17, 2003, Sornette's recent book jumped to rank # 5 on Amazon.com's sales ranking (with Harry Potter as #1!!!)
- Two days before: release of an interview on MSNBC's MoneyCentral website



## **Book sales dynamics**





## **Empirical Implications**

- If buys were mainly initiated via news and advertisements, the model predicts an exponent of  $1+\theta$
- So the power-law exponents being smaller than 1 indicates:
  - Sales dynamics is dominated by cascades involving high-order generations
  - This implies that n ~ 1, i.e. the social network is close to critical
- Identification of critical niches for optimal marketing strategy



# Economics of Cyber Crime

- vulnerabilities / patches dynamics
  - core problem leading to economic arbitrage...
  - …deeply in favor of crime !
- there is no strong business incentive to make secure software (people are happy already when software works properly!
- complexity of IT systems prevent fast



with Google Sygitzerland FIG. 2: This figure shows the decay in proportion of installed browsers of an old version (N-1) after a new minor release

#### Cyber-risks: worms never disappear...



Power-law fit to the decay of the rate of active infections after removing seasonality, following the outbreak of the blaster virus on the Swiss SWITCH network, as a function of time from 2003 to 2008.

# Application to conflict early warning

with P. Meier (Tufts Univ., Boston) and R. Woodard (ETH Zurich)

- Data extracted using Virtual Research Associates, Inc. (VRA) Reader http://www.vranet.com
- Software parses Reuters Business Briefing newswire
- Database of source and target actors, type of event
- 157 event types, 'all' countries, aggregated into conflict and cooperation time series, 1990-2005

#### 24 conflict event types

Abduction	Armed actions
Armed battle	Arrest and detention
Artillery attack	Assassination
Beatings	Bodily punishment
Coups and mutinies	Criminal arrests
Crowd control	Force use
Hijacking	Hostage and kidnapping
Mine explosion	Missile attack
Physical assault	Political arrests
Riot	Sexual assault
Small arms attack	Suicide bombing
Torture	Vehicle bombing
	-

#### 27 cooperation event types.

Acknow. respons.	Agree to mediation
Agree to negotiate	Agree to peacekeeping
Agree to settlement	Apologize
Collaborate	Demobilize armed forces
Ease sanctions	Empathize
Engage in negotiation	Evacuate victims
Forgive	Grant asylum
Host a meeting	Improve relations
Mediate talks	Observe truce
Offer peace proposal	Offer to Negotiate
Offer to mediate	Promise to mediate
Provide shelter	Relax curfew
Request mediation	Request withdrawal or ceasefire
Travel to meet	-

![](_page_35_Figure_0.jpeg)

Predicting the rise and fall of social and economic interactions by monitoring and modeling internet activities and commercial sales

- Books, Music, DVD,
- Electronics (audio and video, cameras and photography, software, computers and video games, cell phones...)
- Office
- Children and Babies
- Home and Garden (which includes pets)
- Gifts, Registries, Jewellery and Watches
- Apparel and Accessories
- Food
- Health, Personal Care, Beauty
- Sports and Outdoors
- Services (movies, restaurants, travel, cars, ...)
- Arts and Hobbies
- Friends and Favourites

## **Research and Application questions**

I- ACTION METRICS: what is the efficiency of ads compaigns? How to design them better? How to manage in real time (or near-real time) and steer complex social systems, commercial sales, reputation branding, etc?

II- PREDICTING SUCCESS: forecast which product will be successful, where and how to allocate resources.

III-DYNAMICAL ENDO-EXO SEARCH: better search engine using the time dynamics of the endo-exo approach.