

The next level of modeling social interaction

How to detect, quantify and utilize emotional influence

Frank Schweitzer

fschweitzer@ethz.ch

in collaboration with:

David Garcia, Antonios Garas

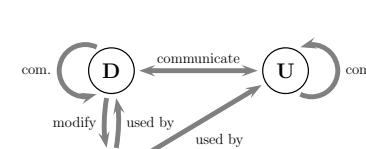


Chair of Systems Design at ETH Zurich

■ Main Research Areas

■ Economic Networks & Social Organizations

- e.g. ownership networks, R&D networks, financial networks, ...
- e.g. online communities, OSS projects, animal societies, ...



■ Methodological Approach: Data Driven Modeling

- **economic databases:** ORBIS, Bloomberg, patent databases
- **online data:** user interaction, communication records, blogs

Outline

- 1 Motivation: Why Collective Emotions
- 2 Quantifying Emotions
- 3 Modeling Cyber Emotions
- 4 Applications
- 5 Outlook

Motivation: What drives our behavior?

■ rational agent: calculates utility

- perfect knowledge?, how to quantify utility?

■ social “ingredients”: for the good and the bad

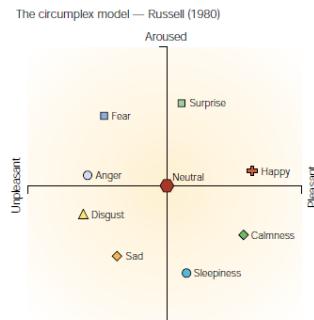
- *individual:* (dis)trust, empathy, aggression, emotions
- *collective:* herding, group feeling, collective emotions



What are emotions?



- short-lived psychological states that consume individual's energy and strongly bias behavior (for example expression)



(Credit: Calder et al. 2001)

Russell's dimensional model

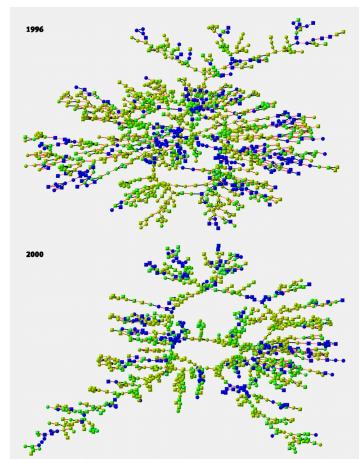
Valence

Pleasure associated with the emotion.

Arousal

Degree of activity induced by the emotion.

A big difference: Happiness network



^aFowler, Christakis, 2008

- time aggregated clusters of happy individuals based on two snapshots within 20 years
- correlations don't show collective emotional states, but global lifetime happiness
- hypothesis of happiness contagion is not verified

Outline

1 Motivation: Why Collective Emotions

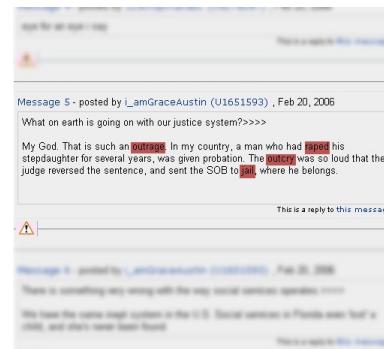
2 Quantifying Emotions

3 Modeling Cyber Emotions

4 Applications

5 Outlook

Emotional Posts in Fora



Example of a negative (left) and a positive (right) post

- Threads analysed in two different ways

- text-based emotion classification ⇒ *sentiment analysis*
- measuring physiological responses of users

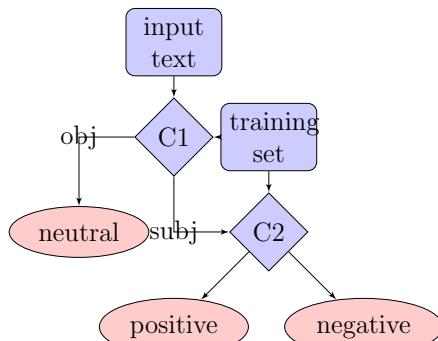
Text-based emotion classification

Annotated lexicon

- positive and negative score for predefined words

Supervised learning

- training set: annotated text, output: subjectivity, polarity



Survey based lexicon (ANEW)

- dataset of word emotionality \Rightarrow valence, arousal, dominance
- improvement: stemming of words \Rightarrow better accuracy, recall

Lyrics for
Michael Jackson's Billie Jean

"She was more like a beauty queen
from a movie scene.
:
And mother always told me,
be careful who you love.
And be careful of what you do
'cause the lie becomes the truth.
Billie Jean is not my lover,
She's just a girl who claims
that I am the one.
:"

ANEW words	v_k	f_k	$v_{text} = \frac{\sum v_k f_k}{\sum f_k}$
$k=1$, love	8.72	1	
2, mother	8.39	1	
3, baby	8.22	3	
4, beauty	7.82	1	
5, truth	7.80	1	
6, people	7.33	2	
7, strong	7.11	1	
8, young	6.89	2	
9, girl	6.87	4	
10, movie	6.86	1	
11, perfume	6.76	1	
12, queen	6.44	1	
13, name	5.55	1	
14, lie	2.79	1	

*

*P. S. Dodds, C. M. Danforth, 2010

What is problematic here?

1 performance of algorithm

- context sensitivity*: emotions only in words?
- subtle meaning*: That's not bad ... – how good is it?

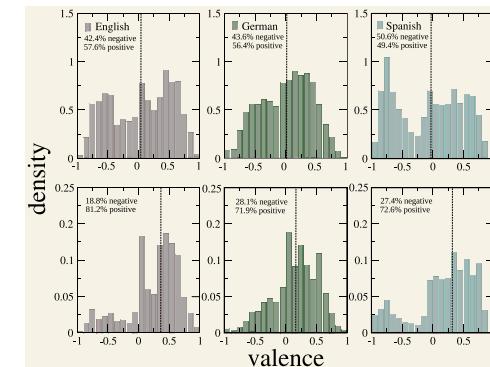
2 quality of lexicon

- human ratings*: English (1034 w), German (2902 w), Spanish (1034 w)
- validation against independent measurements (*physiology*)?

3 inherent properties of used language

- how emotional is "neutral" communication?
- what is the reference point for "normal" valence? (zero???)

How emotional is used language?

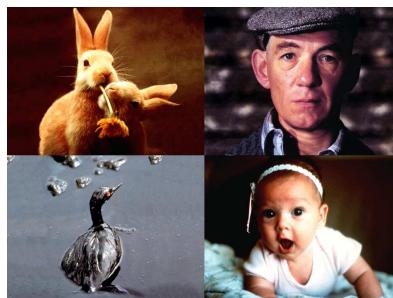


- positive words are more frequently used (Polyanna hypothesis)

- lexica: *no bias*, full range of valence \Rightarrow *neutral* (mean, median)
- frequency of word usage from Google N-gram dataset (10^{12} token)
- example*: $v = 0.715$: "party" ($144.7 \cdot 10^{-6}$) "sunrise" ($6.8 \cdot 10^{-6}$)

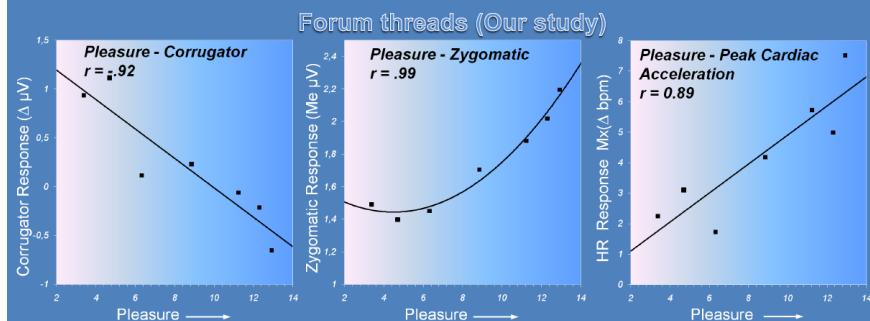
*D. Garcia, A. Garas, F.S., EPJ Data Science 1:3 (2012)

Measuring physiological response



- physiological response to classified pictures[†] and fora
 - monitoring heart rate, skin conductance, frowning and smiling
 - already known to correlate with valence and arousal

[†]IAPS - International affective picture system



Outline

- 1 Motivation: Why Collective Emotions
- 2 Quantifying Emotions
- 3 Modeling Cyber Emotions
- 4 Applications
- 5 Outlook

Modeling framework: Brownian agents

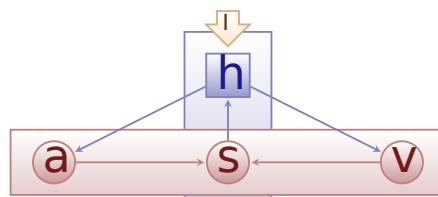
- emotional state of agent i : $E_i(t) = \{v_i(t), a_i(t)\}$
- without external/internal excitation: $v_i(t) \rightarrow 0, a_i(t) \rightarrow 0$
 - relaxation into a 'silent' mode
- dynamics of the Brownian agent:

$$\dot{v}_i = -\gamma_{vi} v_i(t) + \mathcal{F}_v + A_{vi} \xi_v(t)$$

$$\dot{a}_i = -\gamma_{ai} a_i(t) + \mathcal{F}_a + A_{ai} \xi_a(t)$$
- γ_{vi}, γ_{ai} : decay on valence and arousal
- $\mathcal{F}_v, \mathcal{F}_a$: reflect specific influences

[†]Frank Schweitzer: *Brownian Agents and Active Particles. On the Emergence of Complex Behavior in the Natural and Social Sciences*, Berlin: Springer (2003)

Modeling framework: Schema



- agents described by *arousal* a , *valence* v , *expression* s
- arousal causes expression wrt on valence
- emotional information stored in field h

$$\dot{h}_{\pm} = -\gamma_{h\pm} h_{\pm}(t) + s n_{\pm}(t) + I_{\pm}(t)$$

- valence and arousal are affected by the field

† F.S., D. Garcia: D. *An agent-based model of collective emotions in online communities*, European Physical Journal B, vol. 77, no. 4 (2010) pp. 533–545, <http://arxiv.org/abs/1006.5305>

Valence and Arousal

- valence:** nonlinear influence of information

$$\mathcal{F}_v[h_{\pm}(t), v_i(t)] = h_{\pm}(t) \sum_{k=0}^n b_k v^k(t)$$

- arousal:** subthreshold dynamics: nonlinear response

$$\mathcal{F}_a \propto (h_+(t) + h_-(t)) \sum_{k=0}^n d_k a^k(t)$$

- $a_i(t) > \mathcal{T}_i$: agent takes action

- expresses emotions in blogs, fora, reviews, ...

$$s_i(t + \Delta t) = f[v_i(t)] \Theta[a_i(t) - \mathcal{T}_i]$$

- after expressing emotion, arousal is set back to zero

$$\dot{a}_i = \dot{a}_i(t) \Theta[\mathcal{T}_i - a_i(t)] - a_i(t) \Theta[a_i(t) - \mathcal{T}_i]$$

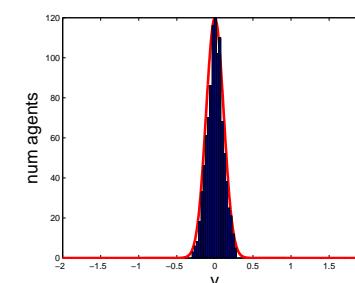
Valence dynamics

- Cubic dependence on the valence

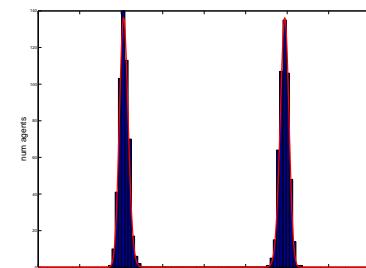
$$\dot{v} = -\gamma_v v(t) + h_{\pm}(t) \{ b_0 + b_1 v(t) + b_2 v^2(t) + b_3 v^3(t) \}$$

- allow for 'silent' mode: $v(t) \rightarrow 0$: $b_0 = 0$
- positive and negative valences 'equal': $b_2 = 0$
- collective emotions** emerge if $b_1 \cdot h_{\pm} > \gamma_v$
 \Rightarrow regime with high emotional information (!)

Valence distribution



$p_s(v)$ under low h



$p_s(v)$ under high h

- polarization of emotions emerges under high information exchange
- agreement of analytical results with simulations

Arousal dynamics

- quadratic dependence on the arousal

$$\dot{a} = -\gamma_a a(t) + h(t) \{ d_0 + d_1 a(t) + d_2 a^2(t) \}$$

- response to total information $h(t) = h_+(t) + h_-(t)$

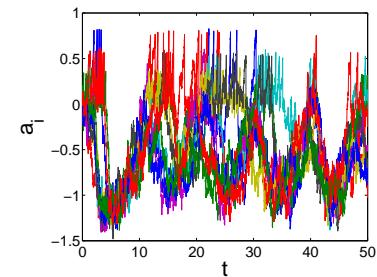
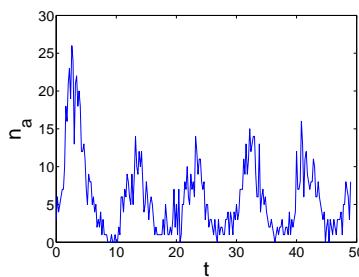
- initial bias to positive arousal $d_0 > 0$

- if $d_2 \neq 0$, two possible solutions

- two cases:

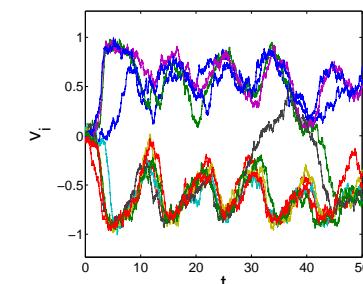
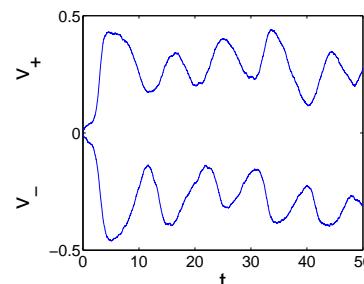
- 1 $d_2 < 0$ lower solution unstable, higher stable \Rightarrow one CE
- 2 $d_2 > 0$ lower solution stable, higher unstable \Rightarrow fluctuating CEs

Collective arousal $\mathfrak{T}_i \sim U(\mathfrak{T}_{min}, \mathfrak{T}_{max})$



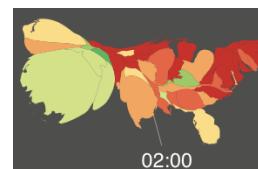
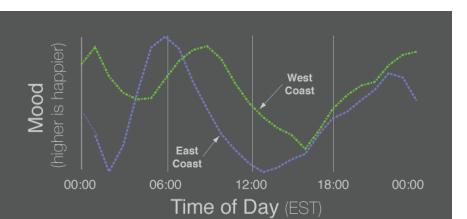
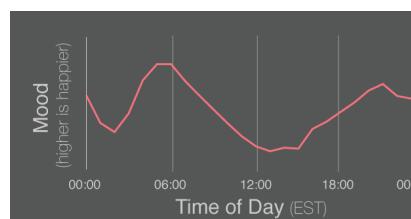
- amount of agents expressing emotions fluctuates
- appearance and fading of collective emotions can be observed

Collective emotions oscillate



- valence polarizes with activity fluctuations
- agent trajectories show change in emotions

A big difference: Expression patterns



- U.S. daily mood changes inferred from Twitter
- no self-organized collective emotion, but daily/weekly effects
- possible origin: tweets 'good morning', 'good night' dominate pattern

[†]<http://www.ccs.neu.edu/home/amislove/twittermood/>

Outline

- 1 Motivation: Why Collective Emotions
- 2 Quantifying Emotions
- 3 Modeling Cyber Emotions
- 4 Applications
- 5 Outlook

Let's start an emotional discussion ...



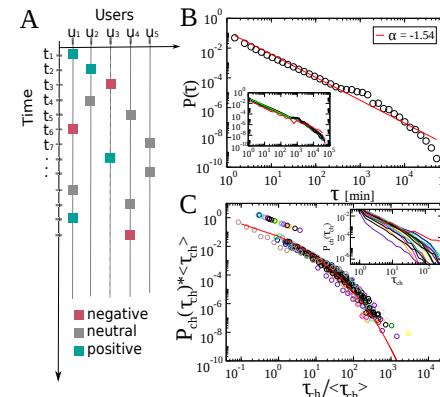
Zürich als Hauptstadt?

ZÜRICH. «Was zum Teufel ist Bern? Ich bin mir sicher, dass man das Bundeshaus auch in Zürich aufstellen könnte. Yes, we can!» Mit solchen Parolen fordern bereits über 5000 Facebook-Mitglieder die Verlegung der Schweizer Hauptstadt nach Zürich. Mehrere Gruppen haben

zum Ziel, über 100000 Einträge für eine Petition zu sammeln. Schliesslich sei Zürich als Verkehrsknotenpunkt und internationaler Finanzplatz die einzige «Global City» der Schweiz. Zudem sei die Limmatstadt zum siebten Mal in Folge ausgezeichnet worden. Die Berner wehren sich mit Händen und Füßen gegen die Bewegung: 1500 Mitglieder zählt die Gruppe «Anti-Petition, Bern bleibt Bundesstadt». Berns Stadtpolitiker Alexander Tschäppat bleibt gelassen: «Schön, dass uns die Zürcher derart beneiden.»

† January 2009: Emotional discussion of more than 5.000 facebook users to make Zurich the Swiss capital, instead of the much smaller Berne ('What the hell is Berne?'). This raised an anti-campaign of another 1.500 facebook users to keep Berne as the capital.

User temporal activity in IRC channels



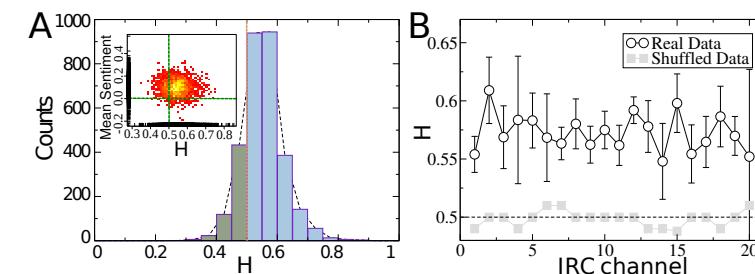
- B) Power-law distribution of user delay: **individual time dynamics are indepent of conversation**
- C) Distribution of inter-event time in each channel: **channels have a natural time delay**

Dialog systems can sample believable individual behavior from **B** and keep conversations natural if delays follow **C**

† A. Garas, D. Garcia, M. Skowron, F.S., *Scientific Reports* 2:402 (2012)

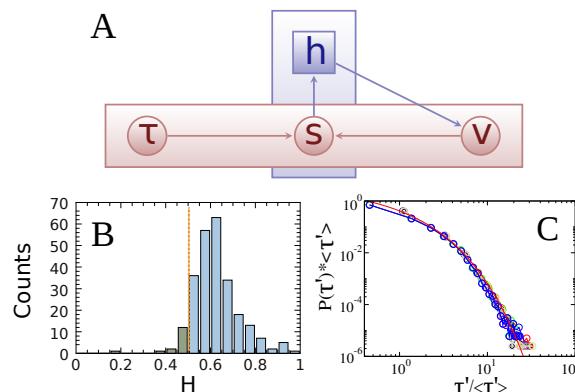
Emotional persistence of user behavior

- A) Most individual users are persistent wrt emotions
 - Hurst exponent H measures deviation from random behavior ($H = 0.5$)
- B) Conversations are persistent (social norms)



† A. Garas, D. Garcia, M. Skowron, F.S., *Scientific Reports* 2:402 (2012)

An agent-based model for chatroom users

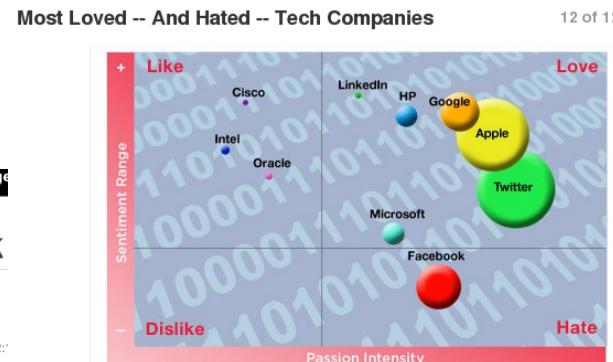


$$N = 10^4, V_- = -0.15, V_+ = 0.05, \gamma_v = 0.2, A_v = 0.2, b = 0.01, c = 0.05, \gamma_h = 0.9$$

Distribution of H (B) and rescaled distribution of inter-message time (C) similar to real data.

†A. Garas, D. Garcia, M. Skowron, F.S., *Scientific Reports* 2:402 (2012)

How do customers really feel?

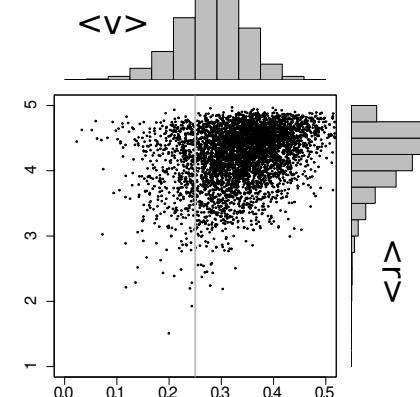


Sentiment Analysis Gives Companies Insight Into Consumer Opinion

Kia, Best Buy, and Viacom are using new tools to mine comments on the Web to see what consumers really think of their brands

Emotional influence on the internet: Reviews

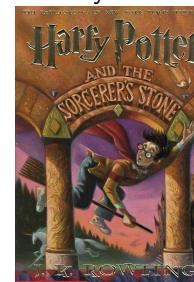
- rating vs. expressed emotion: any difference?
 - high ratings do not correlate with positive emotions
 - emotional charge as additional information



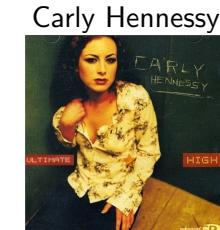
How do consumers really decide?

- often wrong predictions in market research
- social effects (herding, emotions) commonly neglected

Harry Potter



Rejected by 12 publishers.
More than 400 million sales.



2.2 Million dollars invested.
Sold 378 copies in the first three months.

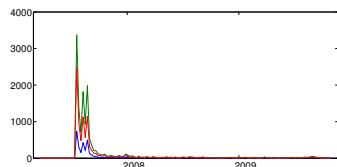
Manos: the hands of fate (1966)



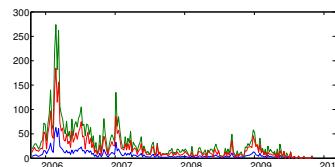
Worst movie ever (IMDB.com). DVD sales boost, sequel in production.

Hot selling vs slow moving products: Emotions

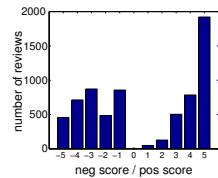
Weekly statistics (ratings: blue, positive: green, negative: red)



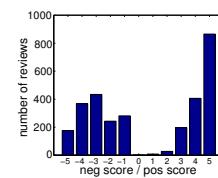
(left) "Harry Potter and the Deathly Hallows", (right) "Twilight: New Moon"



Distribution of emotional scores

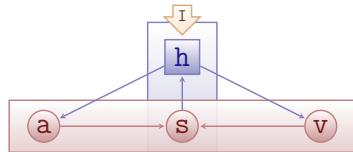


(left) "Harry Potter and the Deathly Hallows", (right) "Marley and Me"

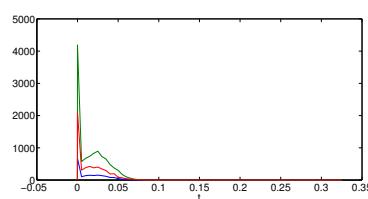
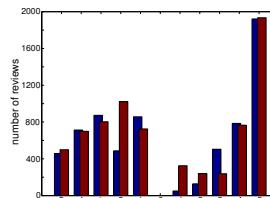


†D. Garcia, FS, IEEE SocialCom (2011), pp. 483-488

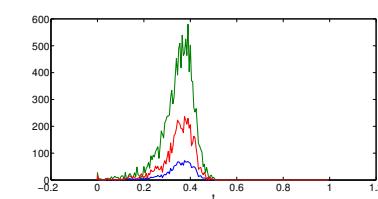
An agent-based model for emotions in reviews



(right) Distribution of emotional scores "Harry Potter" (blue), simulations (red)



Weekly statistics (left) hot selling, (right) slow moving product



†D. Garcia, FS, IEEE SocialCom (2011), pp. 483-488

Outline

- 1 Motivation: Why Collective Emotions
- 2 Quantifying Emotions
- 3 Modeling Cyber Emotions
- 4 Applications
- 5 Outlook

Project: Productivity in OSS

SNF project: *Impact of social interactions on software evolution*

■ impact on emergence of collaboration

- Why do people contribute to OSS at all?
- emotions solve cooperation paradox \Rightarrow explain responsibility

■ impact on project success/failure

- role of emotional feedback of users and empathy of developers
- emotions as building blocks of self-organized, non-profit community



Example: OSS Forum fights

Open Source Software fora show bursts of negative conversations between users and developers:

The screenshot shows a Trac ticket page for Pidgin, specifically ticket #4986. The ticket title is "automatic chat input field resizing should be optional, regression from 2.3". The description includes a note that it's impossible to unsubscribe from a ticket in Trac. A comment from user 'G-sus' discusses the issue, mentioning that it's actually impossible to unsubscribe from a ticket in Trac at least not without manually mucking about in the trac databases. Another comment from 'Apewall' provides more context, stating that it's actually impossible to unsubscribe from these e-mails due to constant monitoring by bots. The ticket has been open for 3 years and last modified 15 months ago.

- OSS project hampered by CE
- Effect: fork of Pidgin into FunPidgin

EU Project on Cyber Emotions (CE)

<http://www.cyberemotions.eu/>

To understand the role of *collective emotions* in creating, forming and breaking-up *ICT mediated communities* as a spontaneous emergent behaviour occurring in complex techno-social networks

- Funded by 7th Framework Programme (start 02/2009)
- Collaboration of 8 European universities for 4 years



JSI Ljubljana



CE Project: Monitoring, Mitigating

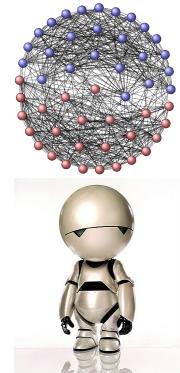
■ virtual humans

- emotional interaction beyond textual expression
- users realize “impact” of their text



■ visualization of collective emotions

- monitoring/ prediction of emotional status of communities
- when (and where) are issues heating up?



■ emotional chatbots

- mitigate emotional problems, online conflicts, encourage cooperation, interaction ⇒ *Artificial emotional intelligence*



■ The ultimate Touring Test

Conclusions

■ emotions

- differ from opinions(!), quantified by *valence*, *arousal*
- collective emotions important in decision processes; overcome dilemma

■ empirics on emotions/cyber emotions

- sentiment mining in text, plus *physiological responses*
- vast datasets to analyse: Myspace, IRC, Amazon reviews, Twitter ...

■ agent based model of collective emotions

- considers psychological variables (*arousal*, *valence*)
- provides *testable hypotheses* on agent's response
- predicts distribution of valence ⇒ data comparison
- framework applicable to IRC chats, product reviews, ...

■ applications

- understand viral marketing based on emotions
- impact on productivity: OSS
- developing bots to enhance user interaction

New Journal: EPJ Data Science

 **EPJ Data Science**
a SpringerOpen Journal

Search EPJ Data Science for Go
Advanced search

Home Articles Authors Reviewers About this journal My EPJ Data Science

Editors-in-Chief

Frank Schweitzer, ETH Zürich
Alessandro Vespignani, Northeastern University

[Editorial Board](#) | [Instructions for authors](#) | [FAQ](#)

Articles

Latest  Most viewed

Regular article Open Access

Effects of time window size and placement on the structure of an aggregated communication network

Krings G, Karsai M, Bernhardsson S, Blondel VD and Saramäki J
EPJ Data Science 2012, 1:4 (18 May 2012)

Regular article Open Access

Positive words carry less information than negative words

Garcia D, Garas A and Schweitzer F
EPJ Data Science 2012, 1:3 (18 May 2012)

Regular article Open Access

Long trend dynamics in social media

Wang C and Huberman BA
EPJ Data Science 2012, 1:2 (18 May 2012)

Aims & scope


The 21st century is currently witnessing the establishment of data-driven science as a complementary approach to the traditional hypothesis-driven method. This (revolution accompanying the paradigm shift from reductionism to complex systems sciences has already largely transformed the natural sciences and is about to bring the same changes to the techno-socio-economic sciences, viewed broadly.

EPJ Data Science offers a publication platform to address this evolution by bringing together all academic disciplines concerned with the same challenges:

- how to extract meaningful data from systems with ever increasing complexity
- how to analyse them in a way that allows new insights
- how to generate data that is needed



EPJ.org

Submit a manuscript

Register

Sign up for article alerts

Contact us

Follow @SpringerOpen