

Point process modelling for directed interaction networks

Patrick O. Perry and Patrick J. Wolfe

New York University and University College London

Interaction data

emails

mobile phone calls

transit cards

credit cards

movement in public places

blog entries

online social networks

These transactions leave digital traces that can be compiled into comprehensive pictures of both individual and group behavior

-Lazer et al. (2009)

Raw data + Point process model = Insight

Insight: Which traits and behaviors are predictive of interaction

Raw data: Enron e-mail dataset

4

156 Employees, 21635 Messages, Nov 1998 – June 2002

Message-ID: <7303996.1075860726914.JavaMail.evans@thyme>

Date: Wed, 10 Oct 2001 08:51:16 -0700 (PDT)

From: kenneth.lay@enron.com

To: benjamin.r@enron.com

Subject: RE: Power Trading Group

Ben -

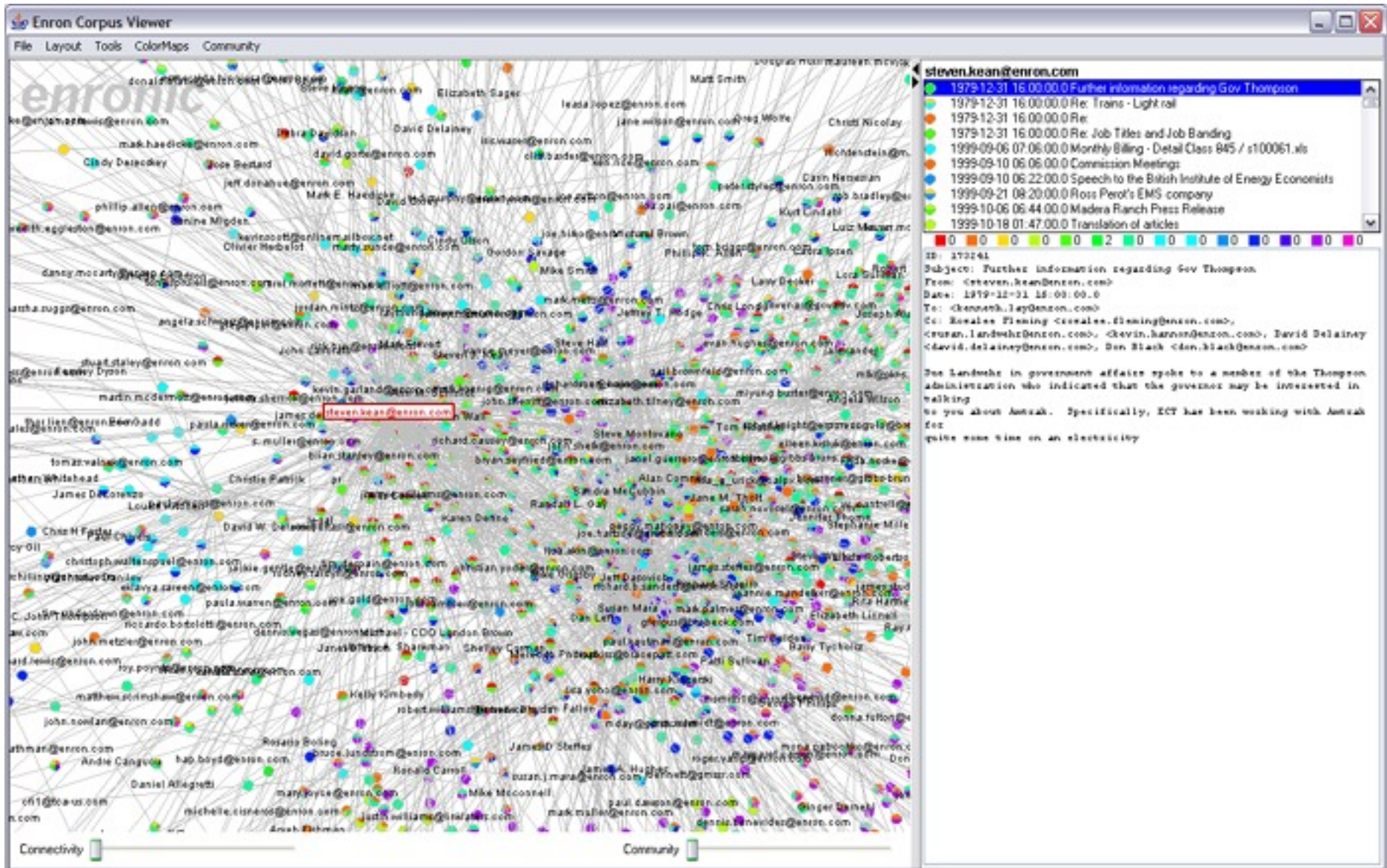
I likewise was glad to see you. Sorry we didn't have a chance to talk.

Good to hear you're doing well. You're with a great group and, yes, the company will soon be doing a lot better.

Thanks,

Ken

156 nodes, 21635 messages



(Heer, 2004)

The big question

Employee Traits

Gender:

Female (43)
Male (113)

Seniority:

Junior (82)
Senior (74)

Department:

Legal (25)
Trading (60)
Other (71)

Question: Which traits and behaviors are predictive of interaction?

Raw data

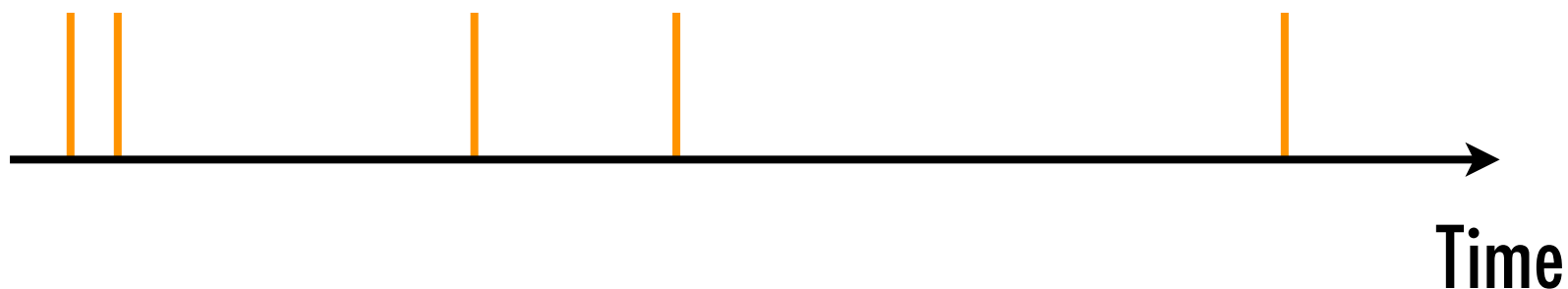
Messages

| Time | Sender | Receiver |
|----------|----------|----------|
| t_1 | i_1 | j_1 |
| t_2 | i_2 | j_2 |
| \vdots | \vdots | \vdots |
| t_n | i_n | j_n |

1. Continuous time
2. Events, not links

Point process model

Messages from i to j :



Model via intensity, $\lambda_t(i, j)$:

$$\lambda_t(i, j) dt = \text{Prob}\{i \text{ sends to } j \text{ in } [t, t + dt)\}$$

Employee traits

| Variate | Characteristic of actor i | Count |
|---------|----------------------------------|-------|
| $L(i)$ | member of the Legal department | 25 |
| $T(i)$ | member of the Trading department | 60 |
| $J(i)$ | seniority is Junior | 82 |
| $F(i)$ | gender is Female | 43 |

20 edge-specific traits: $L(j)$, $L(i)*L(j)$, $T(i)*L(j)$, $J(i) *L(j)$, ...

Notation: $x(i, j) \in \mathbb{R}^{20}$

First attempt: Cox model

Rate of i-j message exchange

$$\lambda_t(i, j) = \bar{\lambda}_t(i) \exp\{\beta^T x(i, j)\}$$

Baseline send rate

Edge-specific covariate vector

Coefficient vector

$$\lambda : \mathbb{R} \times 156 \times 156 \rightarrow \mathbb{R}_+$$

$$\bar{\lambda} : 156 \rightarrow \mathbb{R}_+$$

$$x : 156 \times 156 \rightarrow \mathbb{R}^{20}$$

Problem: Sparsity

Messages from Tania J.

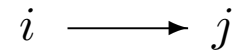
| | | | | | | | | | | | | | | | |
|-----|-----|---|-----|-----|---|---|-----|---|-----|---|-----|-----|---|---|---|
| 33 | 0 | 0 | 192 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 275 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 |
| 405 | 0 | 0 | 0 | 407 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| 67 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 126 | 0 | 0 | 1 | 0 |
| 0 | 3 | 0 | 30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 166 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 271 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 221 | 0 | 0 | 0 | 0 | 1 | 8 | 0 | 507 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 7 | 0 | | | | |

Messages predicted by model

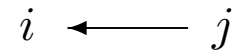
| | | | | | | | | | | | | | | | |
|-------|------|------|-------|-------|-----|-----|-------|------|-------|------|-------|------|------|-------|-----|
| 2.3 | 2.3 | 1.4 | 166.7 | 1.4 | 1.4 | 7.1 | 2.3 | 7.1 | 1.6 | 2.3 | 0.4 | 0.4 | 1.4 | 166.7 | 3.2 |
| 7.1 | 2.3 | 1.4 | 78.5 | 3.2 | 1.4 | 0.4 | 4.4 | 78.5 | 166.7 | 3.2 | 1.4 | 1.4 | 2.3 | 2.3 | 3.2 |
| 2.3 | 78.5 | 3.2 | 0.4 | 0.4 | 3.2 | 1.4 | 2.3 | 1.4 | 1.4 | 2.3 | 0.4 | 78.5 | 78.2 | 4.4 | 4.4 |
| 166.7 | 2.3 | 1.4 | 3.2 | 78.5 | 2.3 | 2.3 | 4.4 | 4.4 | 33.7 | 0.0 | 4.4 | 4.4 | 1.4 | 4.6 | 1.4 |
| 7.1 | 4.6 | 4.4 | 4.4 | 0.4 | 2.3 | 0.4 | 7.1 | 0.4 | 0.4 | 2.3 | 2.3 | 78.2 | 2.3 | 2.3 | 4.4 |
| 4.4 | 4.4 | 2.3 | 2.3 | 0.4 | 0.4 | 0.4 | 1.6 | 2.3 | 2.3 | 33.7 | 166.7 | 1.4 | 4.6 | 166.7 | 0.4 |
| 0.4 | 2.3 | 1.4 | 0.4 | 33.7 | 0.4 | 0.4 | 1.6 | 0.4 | 3.2 | 0.4 | 1.4 | 78.2 | 0.4 | 0.4 | 3.2 |
| 78.5 | 1.6 | 0.4 | 1.4 | 166.7 | 3.2 | 3.2 | 166.7 | 4.4 | 78.5 | 2.3 | 4.4 | 0.4 | 0.4 | 0.4 | 2.3 |
| 3.2 | 78.2 | 78.5 | 0.4 | 0.4 | 2.3 | 2.3 | 2.3 | 3.2 | 78.5 | 1.4 | 1.6 | 0.4 | 4.6 | 2.3 | 4.6 |
| 7.1 | 0.4 | 4.4 | 7.1 | 2.3 | 0.4 | 0.4 | 0.4 | 4.4 | 4.4 | 4.4 | 2.3 | | | | |

Solution: Network effects

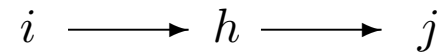
send



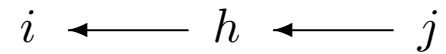
receive



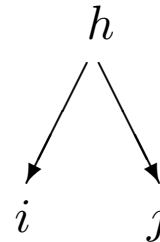
2-send



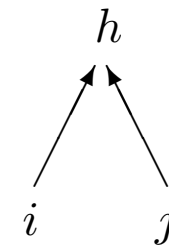
2-receive



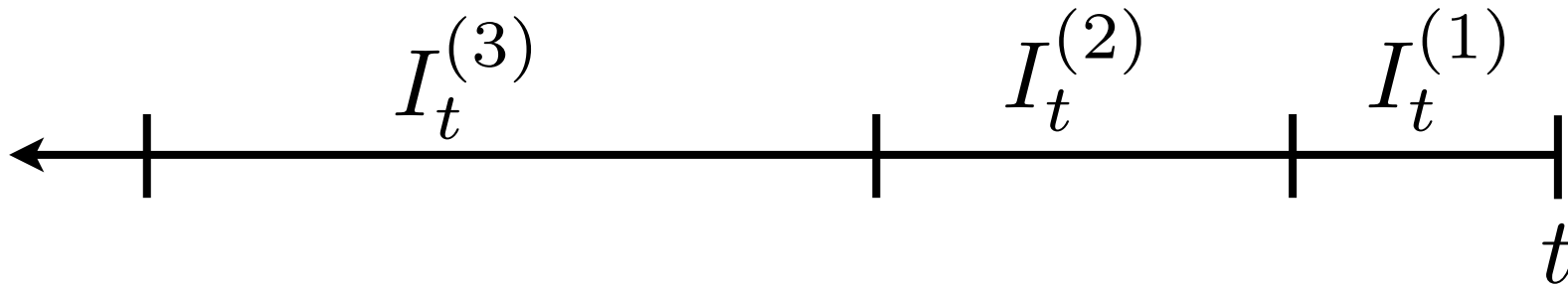
sibling



cosibling



Interval-dependent network effects ¹³



$$\mathbf{send}_t^{(k)}(i, j) = \#\{i \rightarrow j \text{ in } I_t^{(k)}\},$$

$$\mathbf{receive}_t^{(k)}(i, j) = \#\{j \rightarrow i \text{ in } I_t^{(k)}\};$$

Triadic network effects

$$\mathbf{2-send}_t^{(k,l)}(i, j) = \sum_{h \neq i, j} \#\{i \rightarrow h \text{ in } I_t^{(k)}\} \cdot \#\{h \rightarrow j \text{ in } I_t^{(l)}\},$$

$$\mathbf{2-receive}_t^{(k,l)}(i, j) = \sum_{h \neq i, j} \#\{h \rightarrow i \text{ in } I_t^{(k)}\} \cdot \#\{j \rightarrow h \text{ in } I_t^{(l)}\},$$

$$\mathbf{sibling}_t^{(k,l)}(i, j) = \sum_{h \neq i, j} \#\{h \rightarrow i \text{ in } I_t^{(k)}\} \cdot \#\{h \rightarrow j \text{ in } I_t^{(l)}\},$$

$$\mathbf{cosibling}_t^{(k,l)}(i, j) = \sum_{h \neq i, j} \#\{i \rightarrow h \text{ in } I_t^{(k)}\} \cdot \#\{j \rightarrow h \text{ in } I_t^{(l)}\}.$$

Final model

$$\lambda_t(i, j) = \bar{\lambda}_t(i) \exp\{\beta^T x_t(i, j)\}$$

| | |
|----------------------|--|
| $\lambda_t(i, j) dt$ | Prob{i sends j a message in time $[t, t+dt)$ } |
| $\bar{\lambda}_t(i)$ | Baseline intensity for sender i |
| β | Vector of coefficients |
| $x_t(i, j)$ | Vector of time-varying covariates |

(cf. Butts 2008 , Vu et al. 2011)

MPLE asymptotics

Theorem (POP & PJW): Under regularity conditions:

1. $\hat{\beta}_n \xrightarrow{P} \beta$
2. $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} \text{Normal}(0, \Sigma(\beta))$

Cox (1975): heuristic argument ("under mild conditions implying some degree of independence... and that the information values are not too disparate")

Andersen & Gill (1982): survival analysis, fixed time interval

Duplication

From: Alice
To: Bob, Carol, Dan

=

From: Alice
To: Bob

From: Alice
To: Carol

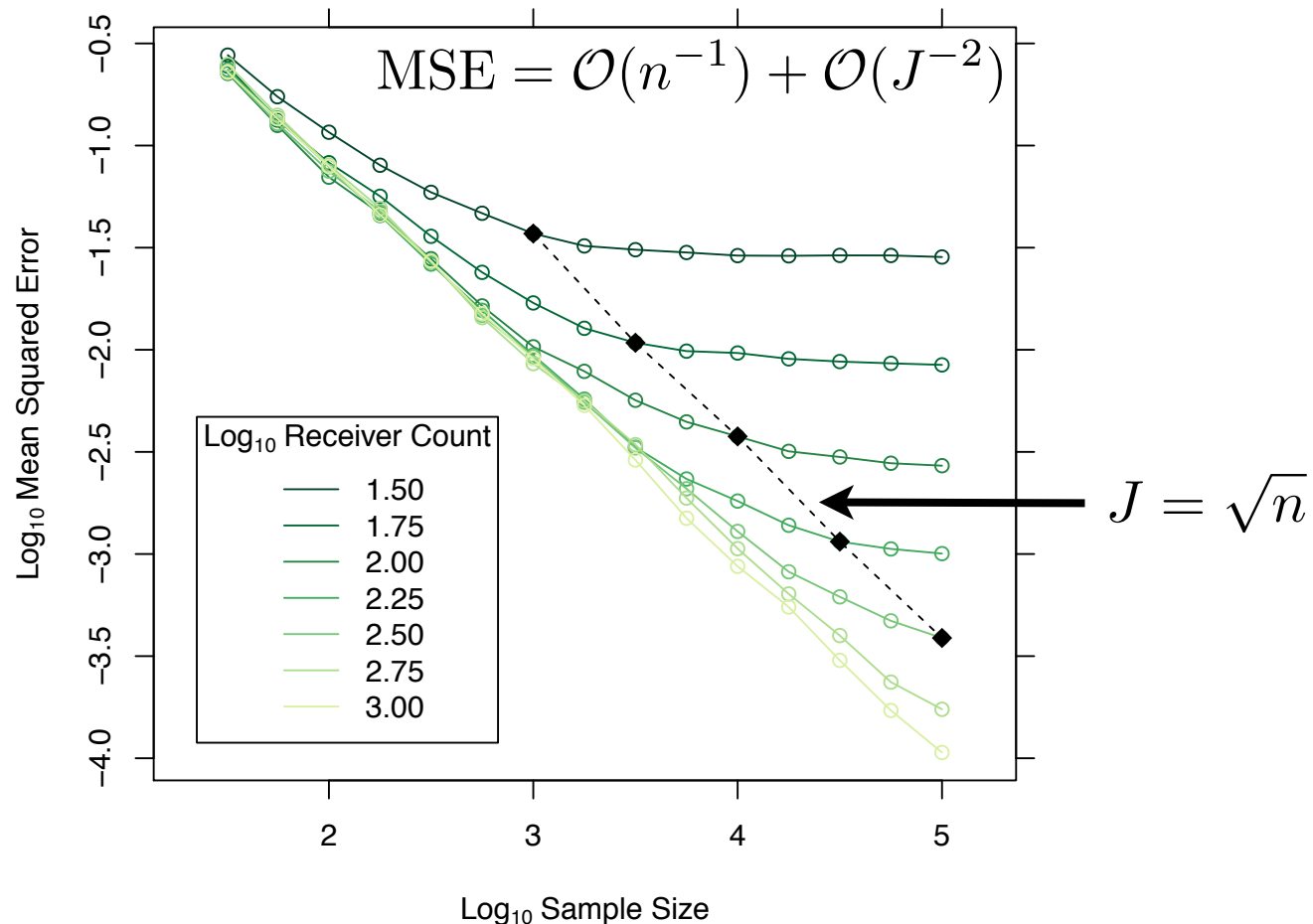
From: Alice
To: Dan

?

(21635 to 35567)

Approximation error

Theorem (POP & PJW): Under regularity conditions, using message duplication introduces bias of order $(\text{nodes})^{-1}$.



Summary so far

1. Interaction data: (t,i,j) tuples
2. Proportional intensity model; capture group effects and reciprocation through covariates
3. Consistent estimates via MPLE

Next: implementation

Enron results

Data

156 employees

21635 messages

Covariates

20 group-level covariates (static)

216 network effects (dynamic)

Time to fit: 15 minutes

Goodness of fit

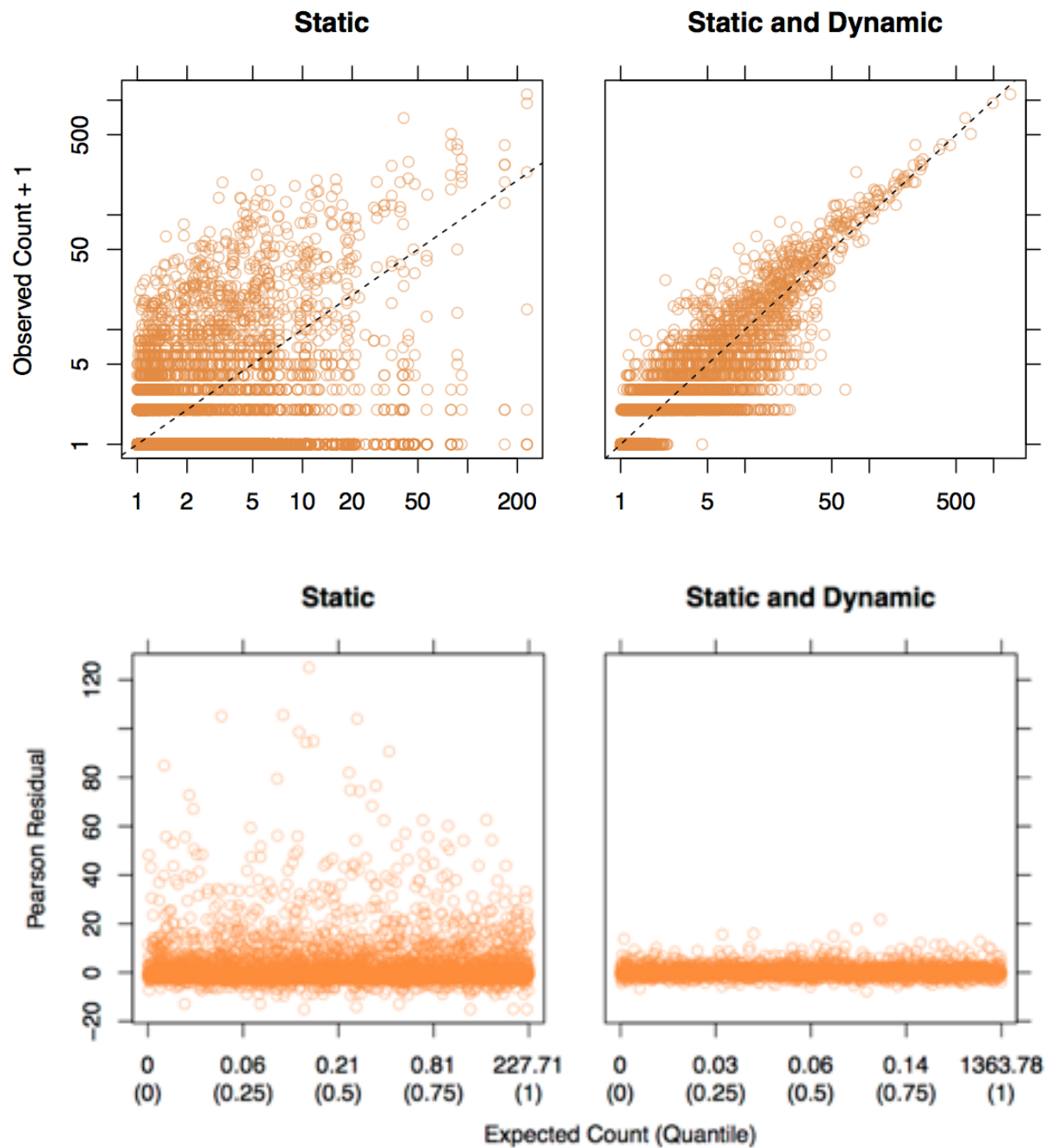
Messages from Tania J.

| | | | | | | | | | | | | | | | |
|-----|-----|---|-----|-----|---|---|-----|---|-----|---|-----|-----|---|---|---|
| 33 | 0 | 0 | 192 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 275 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 |
| 405 | 0 | 0 | 0 | 407 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 1 | 0 | 0 | 0 |
| 67 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 126 | 0 | 0 | 1 | 0 |
| 0 | 3 | 0 | 30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 166 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 271 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 221 | 0 | 0 | 0 | 0 | 1 | 8 | 0 | 507 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 7 | 0 | 0 | 0 | 0 | 0 |

Messages predicted by model

| | | | | | | | | | | | | | | | |
|-------|-------|-----|-------|-------|-----|-----|-------|-----|-------|-----|------|-------|-----|------|-----|
| 8.9 | 0.4 | 0.3 | 223.6 | 0.3 | 0.3 | 6.0 | 0.3 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.3 | 19.8 | 0.3 |
| 0.4 | 0.3 | 0.3 | 0.5 | 5.3 | 0.3 | 0.2 | 0.3 | 0.5 | 267.2 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |
| 0.3 | 0.9 | 0.3 | 0.4 | 0.2 | 0.3 | 0.5 | 0.3 | 0.3 | 0.4 | 0.3 | 0.2 | 29.5 | 0.5 | 0.2 | 3.8 |
| 447.3 | 0.3 | 0.3 | 0.3 | 233.9 | 0.3 | 0.3 | 0.3 | 0.2 | 39.9 | 0.0 | 0.4 | 6.6 | 0.4 | 0.3 | 0.3 |
| 65.6 | 0.5 | 0.3 | 0.2 | 0.2 | 0.3 | 0.2 | 0.2 | 0.2 | 0.2 | 0.3 | 2.7 | 11.5 | 0.3 | 0.4 | 0.3 |
| 0.2 | 0.3 | 0.3 | 0.3 | 0.3 | 0.2 | 0.2 | 0.3 | 0.3 | 0.5 | 1.2 | 90.4 | 0.3 | 0.3 | 1.5 | 0.2 |
| 0.2 | 3.7 | 0.3 | 4.8 | 0.5 | 0.2 | 0.2 | 0.4 | 0.2 | 0.3 | 0.2 | 0.3 | 108.0 | 0.4 | 0.2 | 0.3 |
| 16.2 | 0.3 | 0.2 | 0.3 | 0.5 | 0.3 | 0.3 | 226.1 | 2.5 | 0.9 | 0.4 | 0.3 | 0.2 | 0.2 | 0.2 | 0.3 |
| 0.3 | 206.6 | 0.5 | 0.2 | 0.2 | 0.3 | 7.7 | 3.9 | 0.3 | 655.8 | 0.3 | 0.3 | 0.2 | 0.3 | 0.4 | 0.5 |
| 0.2 | 0.3 | 0.4 | 0.3 | 0.3 | 0.3 | 0.2 | 0.2 | 0.2 | 21.6 | 3.8 | 0.4 | 0.4 | 0.4 | 0.4 | 0.5 |

Goodness of fit



Analysis of deviance

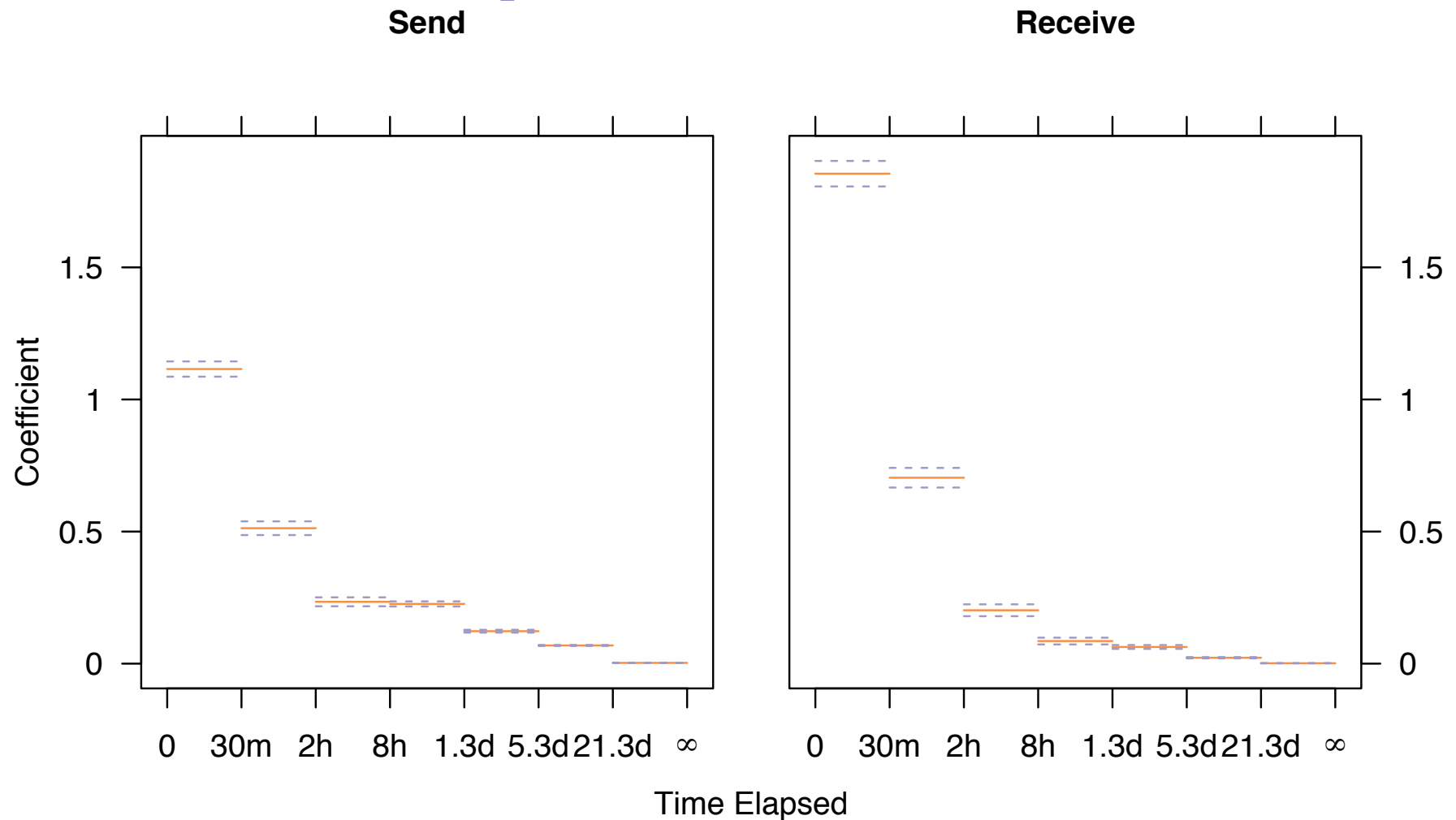
| Term | Df | Deviance | Resid. Df | Resid. Dev |
|-----------|----|----------|-----------|------------|
| Null | | | 32261 | 325412 |
| Static | 20 | 50365 | 32241 | 275047 |
| Send | 8 | 107942 | 32233 | 167105 |
| Receive | 8 | 5919 | 32225 | 161186 |
| Sibling | 50 | 3601 | 32175 | 157585 |
| 2-Send | 50 | 516 | 32125 | 157069 |
| Cosibling | 50 | 1641 | 32075 | 155428 |
| 2-Receive | 50 | 158 | 32025 | 155270 |

Group effects

| Sender | Receiver | | | |
|--------|-----------------|-----------------|-----------------|-----------------|
| | L | T | J | F |
| 1 | -0.91 (0.04) | -0.36 (0.04) | -0.34 (0.04) | 0.04 (0.03) |
| L | 0.63 (0.05) | 0.28 (0.05) | 0.22 (0.04) | 0.15 (0.04) |
| T | 0.32 (0.07) | 0.43 (0.05) | 0.27 (0.05) | -0.07 (0.05) |
| J | 0.06 (0.05) | 0.28 (0.04) | 0.37 (0.03) | -0.13 (0.03) |
| F | 0.59 (0.05) | -0.21 (0.05) | -0.09 (0.04) | 0.15 (0.03) |

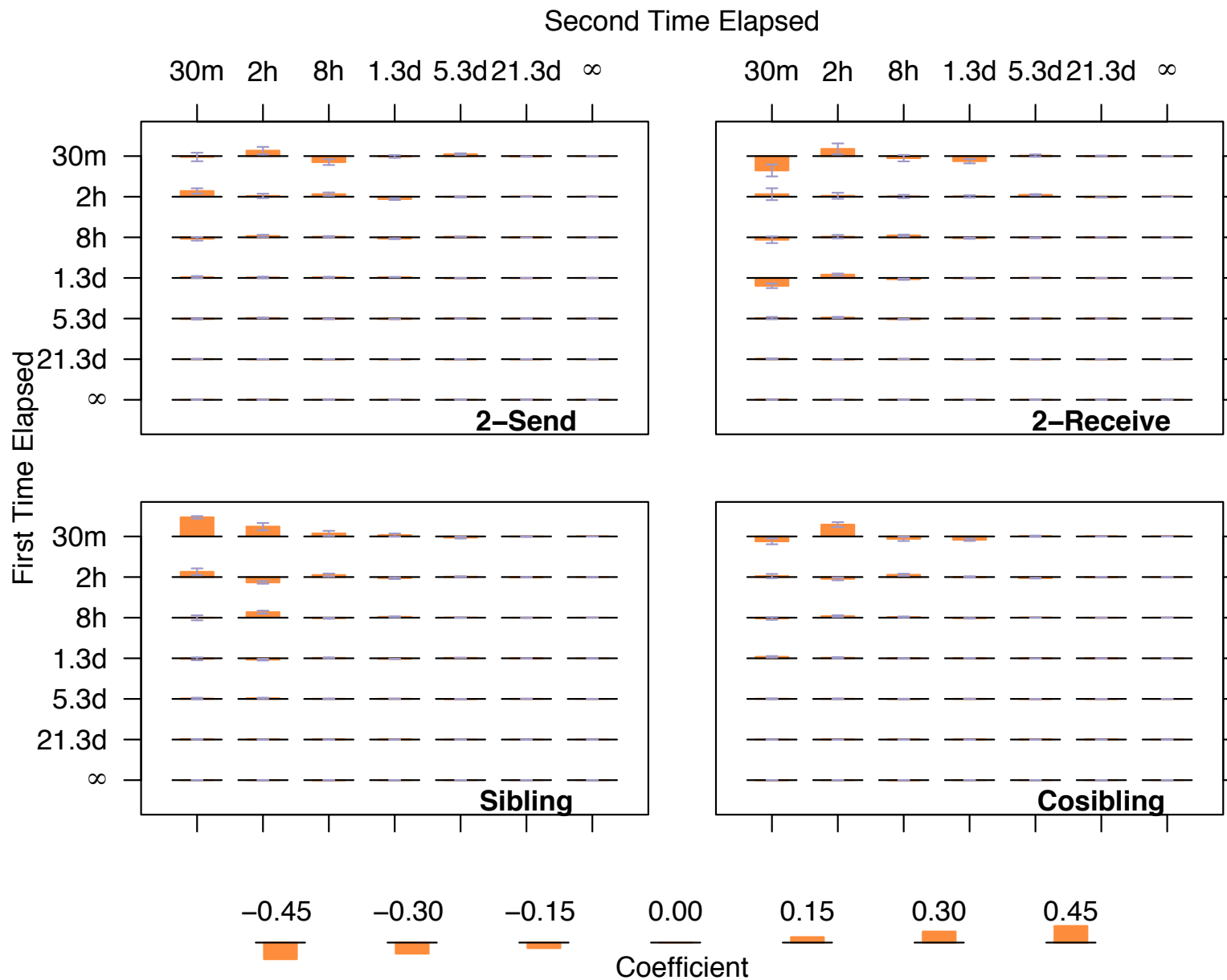
Example: All other factors being equal, **Junior** sends to **Junior** $e^{-0.34 + 0.37} - 1 = 4\%$ more than **Junior** sends to **Senior**; also, **Senior** sends to **Senior** $e^{-(-0.34)} - 1 = 40\%$ more than **Senior** sends to **Junior**.

Dyadic effects



Example: All other factors being equal, every message j has sent i in the last 30 minutes increases the relative i -to- j sending rate by $e^{1.8} = 6$; every message sent between 30 minutes and 2 hours increases the relative rate by $e^{0.7} = 2$.

Triadic effects



What have we learned?

- 1. Employees exhibit trait-based homophily in their message sending behavior.**
- 2. History-dependent network effects are far more predictive than trait-based effects.**
- 3. The predictive strength of the network effects decays rapidly in time.**

References

Andersen, P. K. and Gill, R. D. (1982) Cox's regression model for counting processes: a large sample study. *Ann. Statist.*, **10**, 1100-1120.

Butts, C. T. (2008) A relational event framework for social action. *Sociol. Methodol.*, **38**, 155-200.

Cox, D. R. (1975) Partial likelihood. *Biometrika*, **62**, 269-276.

Heer, J. (2004) Exploring Enron: Visualizing ANLP Results. <http://hci.stanford.edu/~jheer/projects/enron/v1/> .

Perry, P. O. and Wolfe, P. J. (2013) Point process modelling for directed interaction networks. *J. R. Statist. Soc. B*, **75**, 821-849.

Vu, D. Q., Asuncion, A., Hunder, D. and Smyth, P. (2011) Continuous-time regression models for longitudinal networks. *Adv. Neurl. Inform. Process. Syst.*, **24**, 2492-2500.