REM beyond dyads

relational hyperevent modeling with eventnet
(directed hyperevents)

Jürgen Lerner Alessandro Lomi

University of Konstanz
RWTH Aachen

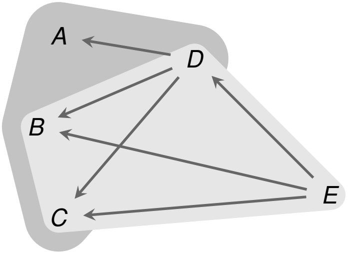
University of Lugano

EUSN Greenwich, 12–16 September 2022

https://github.com/juergenlerner/eventnet

Polyadic interaction: events involving several nodes.

$$e_1 = (t_1, \{D\}, \{A, B, C\})$$
$$e_2 = (t_2, \{E\}, \{B, C, D\})$$



Directed polyadic interaction:

- multicast (one-to-many) communication, email, texting
- citation networks: papers citing lists of references
- virus spreading from persons to several contacts

RHEM for directed hyperevents.

Here: only for events with a single source and arbitrary number of targets.

Hyperedge: can connect any number of nodes.

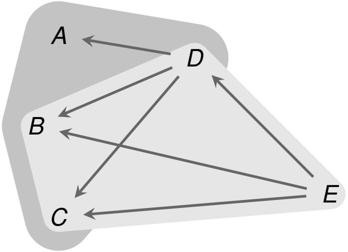
Hyperevent: hyperedge (event participants) with time stamp (event time).

Observed data.

Directed hyperevents $(t_1, i_1, J_1), \ldots, (t_n, i_n, J_n)$,
where for $e = (t_e, i_e, J_e)$

- t_e is the **time** of event e;
- $i_e \in \mathcal{I}_{t_e}$ is the **sender** of event e, taken from a set of possible senders $\mathcal{I}_{t_e}$;
- $J_e \subseteq \mathcal{J}_{t_e}(i)$ is the **set of receivers** of event e, taken from a set of possible receivers $\mathcal{J}_{t_e}(i)$.

$$e_1 = (t_1, \{D\}, \{A, B, C\})$$
$$e_2 = (t_2, \{E\}, \{B, C, D\})$$



Dyadic REM for directed hyperevents.

Perry and Wolfe (2013)

Intensity $\lambda_t(i, J)$; baseline $\overline{\lambda}_t(i, |J|)$; dyadic covariates $x_t(i, j)$.

$$\lambda_t(i, J) = \overline{\lambda}_t(i, |J|) \exp\left\{ \beta_0^{\mathrm{T}} \sum_{j \in J} x_t(i, j) \right\} \prod_{j \in J} \mathbf{1}\{j \in \mathcal{J}_t(i)\} \ .$$

Log partial likelihood; summation over $J \in \binom{\mathcal{J}_{t_e}(i_e)}{|J_e|}$

$$\log L_t(\beta) = \sum_{t_e \leq t} \left(\beta^{\mathrm{T}} \sum_{j \in J_e} x_{t_e}(i_e, j) - \log \left[\sum_J \exp\left\{ \beta^{\mathrm{T}} \sum_{j \in J} x_{t_e}(i_e, j) \right\} \right] \right) .$$

Perry & Wolfe (2013). **Point process modelling for directed interaction networks**. *J RSSB.*

From dyadic REM to RHEM.

It is

$$\begin{aligned}\lambda_t(i,J) &= \overline{\lambda}_t(i,|J|)\exp\left\{\beta_0^{\mathrm{T}}\sum_{j\in J} x_t(i,j)\right\}\prod_{j\in J}\mathbf{1}\{j\in\mathcal{J}_t(i)\} \\ &= \overline{\lambda}_t(i,|J|)\exp\left\{\beta_0^{\mathrm{T}} x_t(i,J)\right\}\mathbf{1}\{J\subseteq\mathcal{J}_t(i)\} \quad ,\end{aligned}$$

if the covariates $x_t(i,J)$ admit the decomposition:

$$x_t(i,J) = \sum_{j\in J} x_t(i,j) \quad .$$

Suitability of j as a receiver is assumed to be independent of other receivers $j' \in J$.

RHEM do not impose that condition and allow more general **hyperedge covariates** $x_t(i,J)$.

RHEM for directed hyperevents.

$$\lambda_t(i,J) \;=\; \overline{\lambda}_t(i,|J|)\exp\left\{\beta_0^{\mathrm{T}} x_t(i,J)\right\}\mathbf{1}\{J \subseteq \mathcal{J}_t(i)\} \; .$$

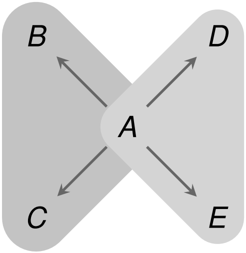
$$\log L_t(\beta) \;=\; \sum_{t_e \leq t}\left(\beta^{\mathrm{T}} x_{t_e}(i_e,J_e) - \log\left[\sum_{J \in \binom{\mathcal{J}_{t_e}(i_e)}{|J_e|}} \exp\left\{\beta^{\mathrm{T}} x_{t_e}(i_e,J)\right\}\right]\right) .$$

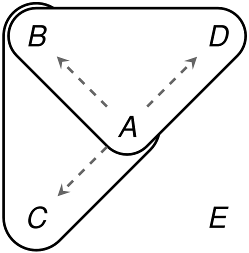
Hyperedge covariates $x_t(i,J)$ do not necessarily decompose into dyadic covariates $x_t(i,j),\ j \in J$.

Usually: sample from the risk set (case-control sampling).

Insufficiency of dyadic effects (I).

Actor A sent two messages: $(A, \{B, C\})$ and $(A, \{D, E\})$.

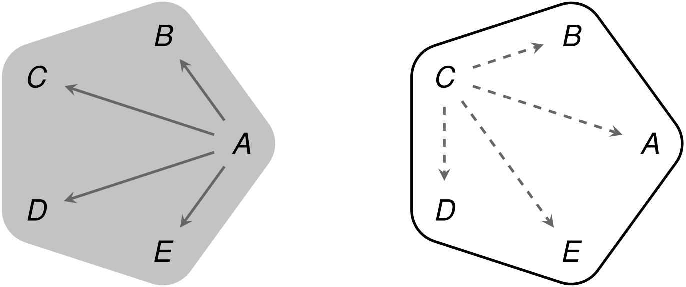




Purely dyadic effects would consider a future message $(A, \{B, C\})$ as likely as a message $(A, \{B, D\})$.

Insufficiency of dyadic effects (II).

"Reply-to-all" in email communication:



Such patterns cannot be captured with purely dyadic covariates.

Objectives of this study.

- **Demonstrate the potential of higher-order effects.**
- Experimentally tackle the following research questions with given empirical data:
  - Is there evidence for higher-order dependencies?
  - Can findings on dyadic effects be distorted by the failure to control for higher-order dependencies?
  - Do hyperedge covariates increase model fit?
  - Do hyperedge covariates help to distinguish observed events from hyperedges that could have experienced an event, but did not?

Argue that higher-order dependencies should not be considered merely an annoyance to be controlled away.

Argue that they allow to develop and test additional theories.

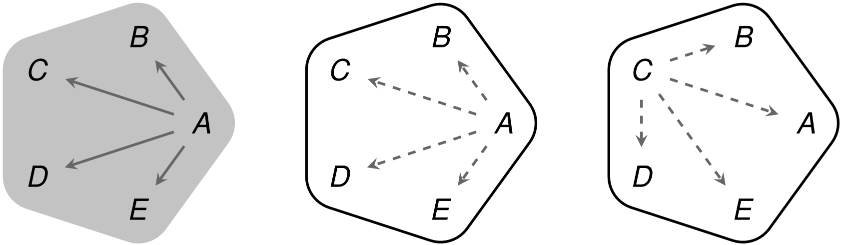
RHEM effects: directed hyperedge covariates.

Here: only for events with a single sender and arbitrary number of receivers.

Exact repetition and undirected exact repetition.

$$repetition_t(i, J) = \sum_{e \in E_{<t}} w(t - t_e) \cdot \mathbf{1}(i_e = i \wedge J_e = J) \ .$$

$$undir.rep_t(i, J) = \sum_{e \in E_{<t}} w(t - t_e) \cdot \mathbf{1}(\{i_e\} \cup J_e = \{i\} \cup J) \ .$$

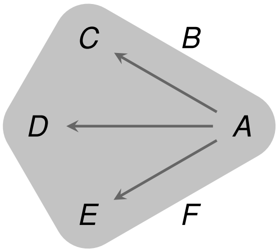


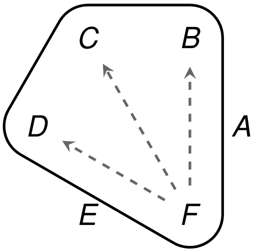
$$w(t - t_e) := \exp\left(-(t - t_e)\frac{\log 2}{T_{1/2}}\right).$$

Partial receiver set repetition.

clustering in space of possible receivers

$$r.sub.rep_t^{(p)}(i,J) \quad = \quad \sum_{J' \in \binom{J}{p}} \frac{hy.deg_t^{(in)}(J')}{\binom{|J|}{p}} \enspace .$$



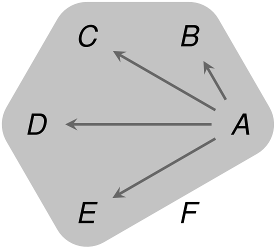


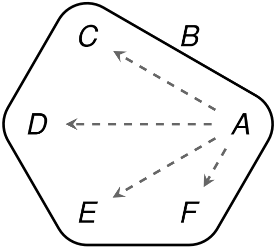
$$hy.deg_t^{(in)}(J') \quad = \quad \sum_{e \in E_{<t}} w(t-t_e) \cdot \mathbf{1}(J' \subseteq J_e) \enspace .$$

Sender-specific partial receiver set repetition.

sender-specific clustering in space of possible receivers

$$s.r.sub.rep_t^{(p)}(i,J) \quad = \quad \sum_{J' \in \binom{J}{p}} \frac{hy.deg_t(i,J')}{\binom{|J|}{p}} \quad .$$



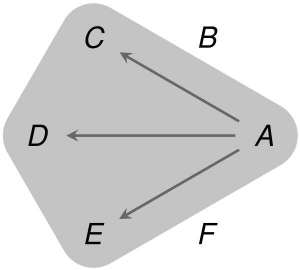


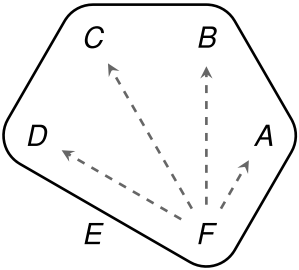
$$hy.deg_t(i,J') \quad = \quad \sum_{e \in E_{<t}} w(t-t_e) \cdot \mathbf{1}(i = i_e \wedge J' \subseteq J_e) \quad .$$

Interaction among receivers.

for instance, citing a paper and some of its references

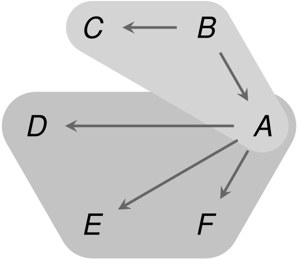
$$\mathit{interact.receivers}_t^{(p)}(i, J) \quad = \sum_{j \in J,\, J' \in \binom{J \setminus \{j\}}{p}} \frac{\mathit{hy.deg}_t(j, J')}{|J| \cdot \binom{|J|-1}{p}} \quad .$$

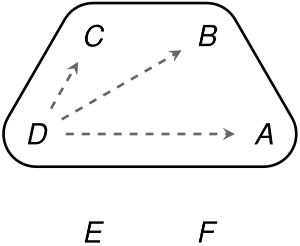




(Generalized) reciprocation.

$$
\begin{aligned}
reciprocation_t(i, J) &= \sum_{j \in J} hy.deg_t(j, \{i\}) / |J| \\
gen.recip_t(i, J) &= \sum_{j \in J} deg_t^{(out)}(j) / |J|
\end{aligned}
$$





$$
deg_t^{(out)}(i') = \sum_{e \in E_{<t}} w(t - t_e) \cdot \mathbf{1}(i' = i_e)
$$

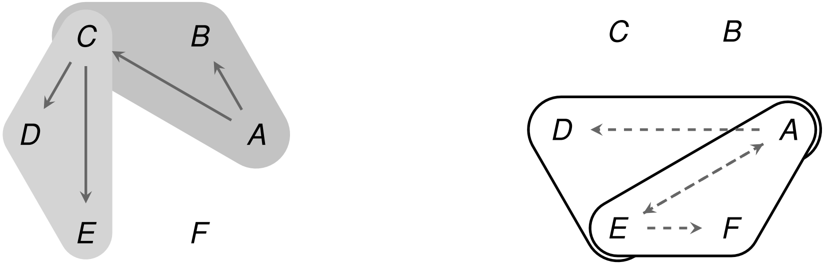
Closure.

$$
\begin{aligned}
trans.closure_t(i, J) &= \sum_{j \in J,\, a \neq i,j} \frac{\min\{deg_t(i, \{a\}), deg_t(a, \{j\})\}}{|J|} \\
cyclic.closure_t(i, J) &= \sum_{j \in J,\, a \neq i,j} \frac{\min\{deg_t(a, \{i\}), deg_t(j, \{a\})\}}{|J|} \\
shared.sender_t(i, J) &= \sum_{j \in J,\, a \neq i,j} \frac{\min\{deg_t(a, \{i\}), deg_t(a, \{j\})\}}{|J|} \\
shared.receiver_t(i, J) &= \sum_{j \in J,\, a \neq i,j} \frac{\min\{deg_t(i, \{a\}), deg_t(j, \{a\})\}}{|J|} \quad .
\end{aligned}
$$

$deg_t(i', J')$ is shorthand for $hy.deg_t(i', J')$, etc.

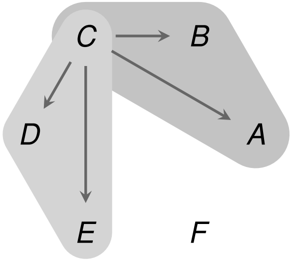
Closure: visual illustration (I).

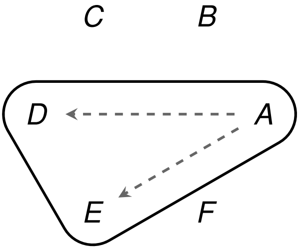
transitive closure and cyclic closure



Closure: visual illustration (II).

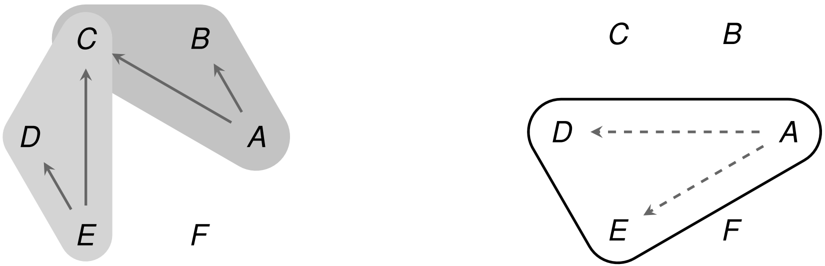
shared sender (source)





Closure: visual illustration (III).

shared receiver (target)



Actor attribute effects.

Given actor attribute $x\colon \mathcal{I} \cup \mathcal{J} \to \mathbb{R}$.

Hyperedge covariates $x_t(i, J)$ dependent on x can measure

- attribute value of the sender $x(i)$
  (not in sender-conditional models)
- summary measure of the distribution of attribute values of the receivers, e. g., $mean_{j\in J}[x(j)]$, $sd_{j\in J}[x(j)]$
- summary measure of the distribution of attribute values of the receivers in relation to the sender, e. g., $mean_{j\in J}[|x(j) - x(i)|]$,

Case study: email network.

Enron email corpus.

https://www.cs.cmu.edu/~enron/

Collection of 21,635 emails among 156 employees of Enron Corporation, cleaned and compiled by Zhou et al. (2007).

Emails (hyperevents) have one sender and between one and 57 receivers.

Actor-level attributes:

gender, seniority, and department (legal, trading, other).

https://github.com/patperry/interaction-proc/tree/master/data/enron

https://github.com/juergenlerner/eventnet/tree/master/data/enron

Receiver set size distribution.

Number of receivers between 1 and 57.

About 30% have more than one receiver.

Mean number of receivers is 1.77.

num. receivers $\lvert J \rvert$	frequency
1	14,985
2	2,962
3	1,435
4	873
5	711
6	180
7	176
8	61
9	24
10	29
$>$ 10	199
all	21,635

	RHEM	Dyadic REM
r.avg.female	0.220 (0.024)***	0.261 (0.024)***
s.r.abs.diff.gender	−0.184 (0.023)***	−0.232 (0.023)***
r.pair.abs.diff.gender	−0.253 (0.065)***	
r.avg.seniority	0.294 (0.024)***	0.417 (0.024)***
s.r.abs.diff.seniority	−0.424 (0.022)***	−0.496 (0.022)***
r.pair.abs.diff.seniority	−0.795 (0.068)***	
r.avg.in.legal	0.057 (0.032)	0.095 (0.032)**
r.avg.in.trading	−0.074 (0.028)**	−0.180 (0.029)***
s.r.frac.diff.department	−0.761 (0.023)***	−0.922 (0.023)***
r.pair.frac.diff.department	−1.152 (0.066)***	
repetition	−0.221 (0.011)***	
undirected.repetition	0.391 (0.013)***	
r.sub.rep.1	0.089 (0.018)***	0.053 (0.018)**
r.sub.rep.2	0.110 (0.009)***	
r.sub.rep.3	0.139 (0.020)***	
r.sub.rep.4	0.252 (0.054)***	
s.r.sub.rep.1	0.674 (0.012)***	0.888 (0.007)***
s.r.sub.rep.2	0.515 (0.024)***	
s.r.sub.rep.3	1.225 (0.166)***	
receiver.outdeg	0.049 (0.016)**	0.101 (0.015)***
reciprocation	0.062 (0.009)***	0.227 (0.006)***
interact.receivers.1	0.164 (0.007)***	
interact.receivers.2	0.290 (0.037)***	
interact.receivers.3	0.630 (0.145)***	
shared.sender	0.352 (0.016)***	0.384 (0.015)***
shared.receiver	−0.009 (0.016)	0.001 (0.014)
transitive.closure	−0.025 (0.019)	0.120 (0.017)***
cyclic.closure	−0.121 (0.014)***	−0.185 (0.013)***
AIC	74,670.703	85,999.084

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Qualitative findings.

- Found relevant effects that do not admit a dyadic decomposition (higher-order effects).

- Higher-order effects are typically significant.
- Effect sizes of dyadic effects typically decrease when controling for higher order effects.
- In some cases: effect significant in dyadic model but not in RHEM.
- RHEM have better model fit.

Some structural effects.

Negative repetition and positive undirected repetition.

- Turn-taking within emergent conversation groups.
- Alternatively: effect of reply-to-all functionality.

Partial repetition of receiver sets.

- Clustering in space of actors: subsets of actors likely to receive joint messages.

Sender-specific partial repetition of receiver sets.

- Subsets of actors likely to receive joint messages from a given sender (sender-specific clustering).

Predictive performance (within sample).

Fit models to all events. For each event e: how many associated non-events are predicted a higher rate than e?

Results for all emails and emails with given number of receivers.

	all	$\lvert J \rvert = 1$	$\lvert J \rvert = 2$	$\lvert J \rvert = 3$	$\lvert J \rvert = 4$	$\lvert J \rvert \geq 5$
num.emails	21,635	14,985	2,962	1,435	873	1,380
RHEM # first	13,129	7,292	2,382	1,302	832	1,321
RHEM % first	60.68	48.66	80.42	90.73	95.30	95.72
RHEM avgrank	3.76	4.95	1.91	0.68	0.25	0.24
dyad # first	12,580	7,169	2,142	1,228	786	1,255
dyad % first	58.15	47.84	72.32	85.57	90.03	90.94
dyad avgrank	4.11	5.06	2.66	1.36	1.27	1.55

Predictive performance (out-of-sample).

training/test data split 90/10 by time

Fit models to 90% of events.
For each event *e* in the remaining 10%: how many associated non-events are predicted to have a higher rate than *e*?

Results for all emails in the test data and emails with given number of receivers.

	all	$\lvert J\rvert = 1$	$\lvert J\rvert = 2$	$\lvert J\rvert = 3$	$\lvert J\rvert = 4$	$\lvert J\rvert \geq 5$
num.emails (test)	2,164	1,530	308	139	66	121
RHEM # first	1,320	764	247	130	61	118
RHEM % first	61.00	49.93	80.19	93.53	92.42	97.52
RHEM avgrank	3.97	5.18	1.91	0.17	0.29	0.31
dyad # first	1,251	738	237	112	57	107
dyad % first	57.81	48.24	76.94	80.58	86.36	88.43
dyad avgrank	4.44	5.51	2.49	1.35	0.45	1.54

Note: not a clean split between training and test data since predictions are based on some information from the test data.

Conclusion.

Higher-order effects can be found in empirical data.

Ignoring them can decrease model fit and yield potentially spurious findings.

Higher-order dependencies should not be considered merely an annoyance to be controlled away.

They allow to develop and test additional theories.

`https://github.com/juergenlerner/eventnet`