ZU064-05-FPR LRHEM-v0.4 29 September 2021 18:56

Fifteen minutes of fame: A model for labeled relational hyperevents with application to the dynamic network analysis of the Andy Warhol diaries

Moses Boudourides
Haverford College
mboudourid@haverford.edu

Jürgen Lerner
University of Konstanz
juergen.lerner@uni-konstanz.de

Alessandro Lomi
University of the italian Switzerland, Lugano
alessandro.lomi@usi.ch

Abstract

1 Introduction

2 Background and motivation

Our hypotheses are related with three types of effects: (1) recency (difference between long-term and short-term effects of previous events), (2) closure (tendency that two nodes, actors and/or labels, sharing past events with a common third node co-appear in future events), and (3) assortativity (tendency to label events including popular participants with popular terms). Moreover, the first two types of effects can be analyzed separately for persons and terms or by jointly considering persons and terms; the third type of effects by definition jointly considers persons and terms. We are interested whether jointly considering participants and labels adds model expressiveness and model fit over models considering only participants or only labels, respectively.

We find that extending RHEM to labeled RHEM (L-RHEM) brings at least three crucial benefits. (1) Increased model expressiveness due to fundamentally new types of effects, (2) L-RHEM have better model fit to empirical data than models for unlabeled relational hyperevents, and (3) novel empirical findings that are qualitatively different from related findings with unlabeled RHEM (mostly the different signs for the closure effects: actor closure is negative but label closure or mixed closure is positive).

$e_1 = (t_1, [\{a,b,c,d\}, \{x,y\}])$
$e_2 = (t_2, [\{c,d,e\}, \{z\}])$
$e_3 = (t_3, [\{b,c,d\}, \{x,y,z\}])$

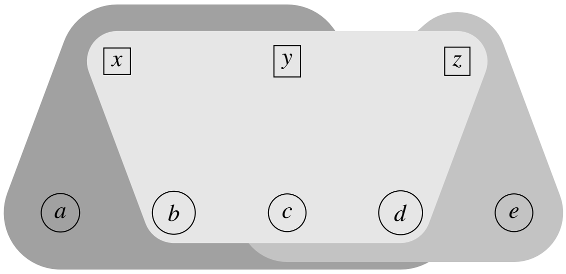

Fig. 1. Stylized example of a list of three labeled hyperevents e_1, e_2, e_3 at event times $t_1 < t_2 < t_3$. Participants are from the set of actors $A = \{a,b,c,d,e\}$ and are displayed as circles; labels are taken from the set of labels $L = \{x,y,z\}$ and are displayed as rectangles. The participants and labels of events are enclosed by gray-shaded areas; older events are in darker color.

3 Models for labeled relational hyperevents (L-RHEM)

3.1 General data format

A *labeled relational hyperevent* is a pair $e = (t_e, h_e)$, where t_e is the time of the event and $h_e = [a_e, \ell_e]$ is the labeled hyperedge of the event, where $a_e \subseteq A_{t_e}$ is the set of participants, taken from a set of actors A_{t_e} that can potentially participate in an event at the event time, and $\ell_e \subseteq L_{t_e}$ is a set of labels (or topics), taken from a set of labels L_{t_e} at the event time. In this paper we analyze only labeled hyperevents containing at least one person and at least one label. Thus, a labeled relational hyperevent is a combination of two undirected hyperevents (**?**; **?**), the hyperevent (t_e, a_e) whose hyperedge is the set of participants a_e and the hyperevent (t_e, ℓ_e) whose hyperedge is the set of labels ℓ_e. A labeled hyperevent can also be seen as a directed hyperevent (**?**; **?**) whose possible sources are targets are selected from disjoint sets, giving rise to a "two-mode hyperevent network". Figure 1 illustrates a list of three labeled hyperevents.

3.2 L-RHEM

Similar to RHEM for undirected hyperevents (**?**), L-RHEM specify point processes on hyperedges. Extending the model from (**?**), L-RHEM specify time-varying intensities $\lambda(t,h)$ for labeled hyperedges $h = [a, \ell]$, that is, for pairs of two sets, a set of participants a and a set of labels ℓ.

Given a sequence $E = (e_1, \ldots, e_N)$ of labeled relational hyperevents and following the framework of Cox proportional hazard models (**?**; **?**), L-RHEM specify the relative intensity λ_0 as a parametric function of hyperedge statistics $s(t,h,E) \in \mathbb{R}^k$ via the formula

$$\lambda_0(t,h,E) = \exp\left(\sum_{i=1}^{k} \theta_i s_i(t,h,E)\right) .$$

The hyperedge statistics $s(t,h,E) \in \mathbb{R}^k$ quantify how the hyperedge h is embedded in the network of past events $E_{<t} = \{e \in E\, ;\, t_e < t\}$. L-RHEM parameters θ are estimated to maximize the likelihood

$$L(\theta) = \prod_{e \in E} \frac{\lambda_0(t_e, h_e)}{\sum_{h \in R_{t_e}} \lambda_0(t_e, h, E)} \quad , \tag{1}$$

in which the relative intensity $\lambda_0(t_e, h_e)$ on the event hyperedge $h_e = [a_e, \ell_e]$ is compared with the relative intensities on hyperedges $h \in R_{t_e}$ that could potentially have been the hyperedge of an event at time t_e. The risk set $R_{t_e} \subseteq \mathscr{P}(A_{t_e}) \times \mathscr{P}(L_{t_e})$ contains pairs $[a, \ell]$ where $a \subseteq A_{t_e}$ is a set of actors and $\ell \subseteq L_{t_e}$ is a set of labels. Similar to (**?**), we consider here only size-constrained risk sets where for all elements $(a, \ell) \in R_{t_e}$, it is $|a| = |a_e|$ and $|\ell| = |\ell_e|$. That is, we compare the event hyperedge h_e only with non-event hyperedges that have the same number of participants and the same number of labels. Moreover, to cope with the sheer size of the risk set, we apply case-control sampling (**?**; **?**; **?**; **?**) by sampling only a fixed number of non-event hyperedges from the risk set. In our empirical analysis we sample 100 non-event hyperedges for each hyperevent.

3.3 Effects in L-RHEM (hyperedge statistics)

Hyperedge statistics are the variables explaining the intensity (event rate, hazard rate) associated with labeled hyperedges and operationalize possible effects in hyperevent networks. The fact that labeled hyperedges are pairs of two sets, the set of participants and the set of labels, gives rise to many more types of hyperedge statistics in L-RHEM than in undirected RHEM.

Most hyperedge statistics are functions of previous events recorded in a hyperedge attribute dubbed *hyperedge degree*. The *(cumulative) hyperedge degree* of a labeled hyperedge $h = (a, \ell)$ and a point in time t is the number of hyperevents that happen strictly before t, whose set of participants is a superset of a, and whose set of labels is a superset of ℓ:

$$hy.deg(t, h, E) = |\{e \in E_{<t}\,;\, a \subseteq a_e \wedge \ell \subseteq \ell_e\}| \quad . \tag{2}$$

The *recent hyperedge degree* lets the influence of past events on hyperedges decay exponentially with a given half life period $T_{1/2}$ (**?**; **?**):

$$rec.hy.deg(t, h, E) = \sum_{e \in E_{<t} : a \subseteq a_e \wedge \ell \subseteq \ell_e} \exp\left(-(t - t_e)\frac{log(2)}{T_{1/2}}\right) \quad .$$

The cumulative and recent hyperedge degrees represent the dynamically changing state of the hyperevent network that shapes the distribution of the next events. All hyperedge statistics defined in the following can alternatively be defined as functions of the cumulative hyperedge degree (assessing the influence of all previous events on the current predicted event rates) or as functions of the recent hyperedge degree (assessing the influence of recent previous events on the current predicted event rates). For brevity we give in the following the definitions for hyperedge statistics based on the cumulative hyperedge degree *hy.deg*.

3.3.1 Labeled subset repetition

For two integers p and q, subset repetition of order (p,q) is defined by

$$sub.rep^{(p,q)}(t,h,E) = \frac{1}{\binom{|a|}{p}\cdot\binom{|\ell|}{q}} \sum_{h'\in\binom{a}{p}\times\binom{\ell}{q}} hy.deg(t,h',E) \ .$$

The sum iterates over all labeled hyperedges $h' = (a',\ell')$, where $a' \subset a$ is a p-element subset of a and $\ell' \subset \ell$ is a q-element subset of ℓ.

In the example displayed in Fig. 1, the third labeled hyperevent $e_3 = (t_3,[\{b,c,d\},\{x,y,z\}])$ takes the value 1/3 in subset repetition of order $(3,2)$ since its three participants b,c,d and two of its three labels, namely $\{x,y\}$, are subsets of the participants, or labels respectively, of the past event $e_1 = (t_1,[\{a,b,c,d\},\{x,y\}])$. As another example from the same figure, the third labeled hyperevent $e_3 = (t_3,[\{b,c,d\},\{x,y,z\}])$ takes the value $4/3$ in subset repetition of order $(2,0)$, since among the three unordered pairs of its participants $\{b,c,d\}$ two, namely $\{b,c\}$ and $\{b,d\}$, have co-participated in one previous event (e_1) and one of the unordered pairs of its participants, namely $\{c,d\}$, has co-participated in two previous events (e_1 and e_2).

3.3.2 Triadic effects (closure in 2-mode hyperevent networks)

Triadic closure in hyperevent networks measures whether two nodes (which are possible participants or labels of the next hyperevent) are both connected by previous events to a common third node. In the case of labeled hyperevents, that is, 2-mode hyperevent networks, we get four different closure effects depending on the mode of the two focal nodes and the mode of the intermediate "third" node. The two focal nodes (i. e., nodes that are members of the next hyperevent) can be actors or they can be labels; the intermediate node ("shared node") can be an actor or a label. These combinations give the following four hyperedge statistics for a labeled hyperedge $h = [a,\ell]$ at time t (for brevity we drop the event sequence E in these formulas).

$$\begin{aligned}
actor.closure(t,h) &= \sum_{\{a_1,a_2\}\in\binom{a}{2}\wedge a_3\neq a_1,a_2} \\
&\quad \min[hy.deg(t,[\{a_1,a_3\},\emptyset]),hy.deg(t,[\{a_2,a_3\},\emptyset])]/\binom{|a|}{2} \\
label.closure(t,h) &= \sum_{\{\ell_1,\ell_2\}\in\binom{\ell}{2}\wedge \ell_3\neq \ell_1,\ell_2} \\
&\quad \min[hy.deg(t,[\emptyset,\{\ell_1,\ell_3\}]),hy.deg(t,[\emptyset,\{\ell_2,\ell_3\}])]/\binom{|\ell|}{2} \\
actor.closure.by.label(t,h) &= \sum_{\{a_1,a_2\}\in\binom{a}{2}\wedge \ell_3\in L} \\
&\quad \min[hy.deg(t,[\{a_1\},\{\ell_3\}]),hy.deg(t,[\{a_2\},\{\ell_3\}])]/\binom{|a|}{2} \\
label.closure.by.actor(t,h) &= \sum_{\{\ell_1,\ell_2\}\in\binom{\ell}{2}\wedge a_3\in A} \\
&\quad \min[hy.deg(t,[\{a_3\},\{\ell_1\}]),hy.deg(t,[\{a_3\},\{\ell_2\}])]/\binom{|\ell|}{2}
\end{aligned}$$

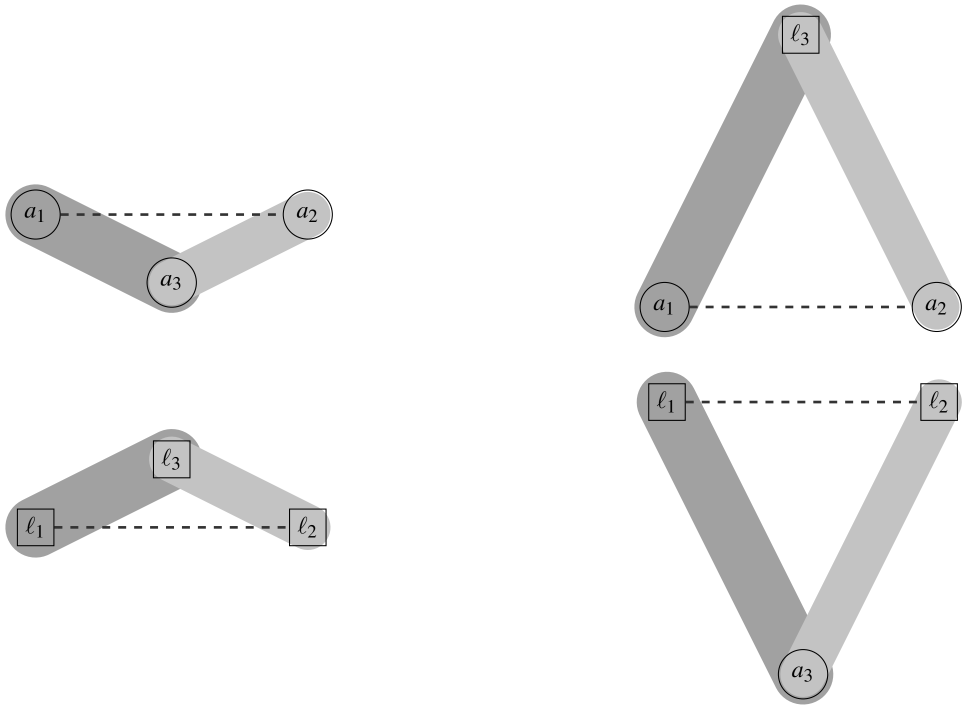

Fig. 2. Illustration of closure. Actors are displayed by circles and denoted by a_1 through a_3; labels are displayed by rectangles and denoted by ℓ_1 through ℓ_3. The dashed line indicates a pair of nodes that are both members of a potential future hyperevent (possibly together with other nodes not shown in the images). Gray-shaded areas indicate past events in which the enclosed nodes participated (possibly together with other nodes not shown in the images). *Top left*: Illustration of *actor.closure*. Actors a_1 and a_2 both have past events with the common third actor a_3. *Top right*: Illustration of *actor.closure.by.label*. Actors a_1 and a_2 both have past events with the common label ℓ_3. *Bottom left*: Illustration of *label.closure*. Labels ℓ_1 and ℓ_2 both have past events with the common third third label ℓ_3. *Bottom right*: Illustration of *label.closure.by.actor*. Labels ℓ_1 and ℓ_2 both have past events with the actor a_3.

For instance, in Fig. 1, at the time of the third event $e_3 = (t_3, [\{b,c,d\}, \{x,y,z\}])$ the two labels y and z have no past common event – but they are both connected to common third actors (c and d) via different past events. Thus, the third event could – among others – be explained by an effect predicting that labels sharing a common actor (quantified through *label.closure.by.actor*) have an increased probability to co-label a future event. All triadic closure effects are illustrated in Fig. 2.

3.3.3 Assortativity

Last but not least, we define an assortativity statistic to assess whether popular actors are typically mentioned with popular labels. *Assortativity* of a labeled hyperedge $h = [a, \ell]$ at time t is defined by

$$assortativity(t,h,E) = sub.rep^{(1,0)}(t,h,E) \cdot sub.rep^{(0,1)}(t,h,E) \ .$$

3.4 Models: actor-only, label-only, and joint L-RHEM

We fit three fundamentally different types of L-RHEM to our empirical data from the Andy Warhol Diaries: a "actor-only" model, a "label-only" model, and a joint model. The actor-only model specifies the event rate on a hyperedge $h = [a, \ell]$ only as a function of its set of actors a – in other words in the actor-only model we completely ignore any possible influence of the labels on event rates. Mirroring this model, the "label-only" model specifies event rates only as a function of the hyperedge labels. Finally, the joint model specifies event rates on hyperedges as functions of participants and labels. One of our methodological objectives is to assess whether the joint modeling of participants and labels adds model expressiveness and fit over the two sub-models that are already included in the model for undirected, unlabeled RHEM (**?**).

The selection of hyperedge statistics included in an L-RHEM determines the type of the model. The actor-only model can include subset repetition of order $(p,0)$ for any value of p (that is, the cardinality of the set of repeated labels is fixed at zero) and actor closure. The label-only model can include subset repetition of order $(0,q)$ for any value of q (that is, the cardinality of the set of repeated participants is fixed at zero) and label closure. The joint models can additionally include subset repetition of order (p,q) where both p and q are non-zero, they can contain the mixed closure statistics "actor-closure by label" and "label-closure by actor", and they can contain assortativity statistics assessing whether events including popular participants are labeled by popular terms.

4 Empirical analysis

4.1 Empirical setting: The Andy Warhol Diaries

The *Andy Warhol Diaries* is the posthumously published memoirs of the American artist Andy Warhol, edited by his collaborator and long-time friend, Pat Hackett. Warner Books first published the book in 1989 (**?**) with an introduction by Pat Hackett and it was republished as a Penguin Classic in 2010.

The book is 807 pages long (containing 450,333 words and 2,401,603 characters/symbols) and it comprises 2024 diaries written in a period of more than eleven years, starting from November 24, 1976, and ending on February 17, 1987, just five days before Andy Warhol's death. It is a condensed version by Hackett of Warhol's more than 20,000 page hand-written personal diary. Each diary (daily entry in the Andy Warhol Diaries) gives rise to a labeled hyperevent by extracting actors (persons mentioned in the diary) and labels (prominent terms appearing in the diary's text).

A person is included as a participant of the hyperevent corresponding to a diary d if the person's name appears in the text of d. To extract person names in the diary's text by

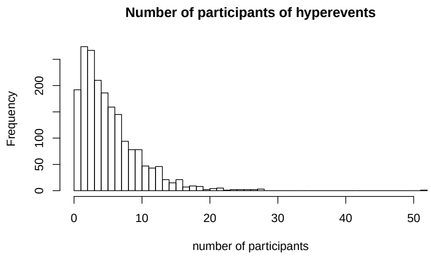

Fig. 3. Histogram of the number of participants of hyperevents.

natural language processing (**?**), we applied named entity recognition to identify named entities mentioned in unstructured text into a certain pre-defined category of person names (**?**; **?**), implemented in the Python module spaCy (**?**). For persons that were referred with their first name, we identify the full name resorting to the two unauthorized indexes which were published by *Spy* (**?**) and *Fame* (**?**) magazines. As mentioned above, we analyze the list of all hyperevents containing at least one person and at least one label, which applies to 1,924 hyperevents. There are 1,793 unique actors labeling at least one of these hyperevents. The number of participants of a given hyperevent ranges from one to 52 with a mean equal to 5.73. The distribution of the number of participants of the hyperevents is displayed in Fig. 3.

To extract the terms used in the diary's text, we have been using the Python module Gensim (**?**) to tokenize the corpora of the text of diaries and removing stop-words. Subsequently, using part of speech tagging (**?**), we were collecting words (terms) identified as nouns or verbs. Since terms appearing in nearly all diaries are not characterizing for the diaries' topics, we retain only terms that are sufficiently specific, as measured by tf-idf scores. To determine the diary's labels we use a variant of tf-idf scores where we adjust term frequency for the number of terms contained in documents. Thus, a term in a longer document has to be more specific (i. e., appear in a smaller fraction of documents) to reach the same tf-idf score than a term in a shorter document. We include a term as a label of the hyperevent corresponding to a given diary if the term's tf-idf score for that diary is at or above the median over all tf-idf scores.

More precisely, let $D = (d_1, \ldots, d_N)$ be the list of all "documents" (corresponding to daily diary entries), let $f_{w,d}$ denote the raw frequency (counting multiplicities) of term w in diary d, let n_w denote the number of diaries in D containing term w at least once, and let $N = |D|$ denote the number of all diaries. The adjusted term frequency of term w in diary d is defined by

$$tf_{w,d} = \frac{f_{w,d}}{\sum_{w' \in d} f_{w',d}}$$

and the resulting tf-idf score of a term w in diary d is

$$tf.idf_{w,d} = tf_{wd} \cdot \log\left(\frac{N}{n_w}\right).$$

The distribution of the logarithmized tf-idf scores over all terms in all diaries is shown in Fig. 4 (*left*) and loosely resembles a normal distribution. We include a term w as a label of

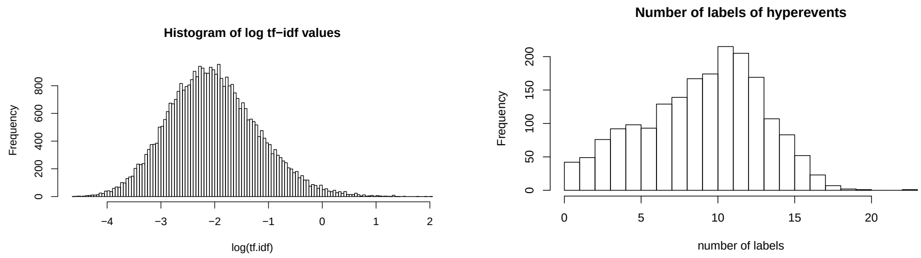

Fig. 4. *Left*: histogram of log tf-idf scores over all terms in all diaries. *Right*: histogram of the number of labels of hyperevents.

the hyperevent corresponding to a diary d if the score $tf.idf_{w,d}$ is at or above the median over all tf-idf scores. This results in 3,309 unique labels for the hyperevents having at least one actor and at least one label. The number of labels of a given hyperevent ranges from one to 23 with a mean equal to 9.48. The distribution of the number of labels of the hyperevents is displayed in Fig. 4 (*right*).

4.2 Results

Below we report L-RHEM parameters of various models fitted to the Andy Warhol Data. We compute hyperedge statistics with an extension of `eventnet`[1] (**?**) to labeled hyperevents. We standardize statistics (zero mean and standard deviation equal to one) to get standardized parameters. Given hyperedge statistics, parameters are estimated with the R package `survival`[2] (**?**).

4.2.1 Actor-only model

Table 1 reports estimated parameters of three "actor-only" models. The first model assesses the effect of all previous events without any decay (cumulative statistics). The second model assesses the effect of recent previous events where we let the influence of past events decay exponentially with a half life of 30 days. The third model combines cumulative and recent statistics. Looking at the model fit indicator AIC (lower values reveal better fit), we find that the short-term-only model fits better than the long-term-only model and that the joint model is the best of all three.

Subset repetition of order $(1,0)$ measures the effect of the past popularity of individual participants. We find that both, the long-term and the short-term effects of this statistic are positive. This means that if an actor has been mentioned more often by Andy Warhol in past diary entries then it is more likely that this same actor will participate in future events. If an actor has recently been mentioned more often, then this has an additional effect increasing the likelihood that he/she will be mentioned in further diary entries in the near future.

[1] `https://github.com/juergenlerner/eventnet`
[2] `https://CRAN.R-project.org/package=survival`

Table 1. Actor-only models. The statistics starting with "rec" assess the effect of recent previous events (exponential decay with half life 30 days). Those without that prefix assess the effect of all previous events without decay.

	actor effects (long-term)	actor effects (short-term)	long-term and short-term
sub.rep.1.0	0.393 (0.014)***		0.154 (0.016)***
sub.rep.2.0	1.037 (0.052)***		0.390 (0.046)***
sub.rep.3.0	0.731 (0.055)***		0.804 (0.070)***
actor.closure	−0.258 (0.039)***		−0.002 (0.033)
rec.sub.rep.1.0		0.923 (0.020)***	0.620 (0.022)***
rec.sub.rep.2.0		0.755 (0.042)***	0.410 (0.041)***
rec.sub.rep.3.0		−0.040 (0.022)$^{\cdot}$	−0.270 (0.029)***
rec.actor.closure		−0.249 (0.027)***	−0.154 (0.025)***
AIC	9689.520	9390.113	8031.553
Num. events	1,924	1,924	1,924
Num. obs.	194,224	194,224	194,224

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Subset repetition of order $(2,0)$ measures the effect of the past joint popularity of pairs of actors. We also find that both the long-term and short-term effect of this statistic are consistently positive. This means that if two actors have been co-mentioned more often, then their likelihood to be co-mentioned in the future increases. If two actors have been recently co-mentioned then this has an additional effect, further increasing their likelihood to be co-mentioned in the near future.

Subset repetition of order $(3,0)$ measures the effect of the past joint popularity of triples of actors by considering how often three actors have been jointly mentioned in the same diary entry. We find that the long-term effect of this statistic is consistently positive but the corresponding short-term effect is negative ($p < 0.1$ in the short-term-only model). Thus, if three actors are jointly mentioned in a diary entry then, controling for all other effects, these three actors have a lower probability to be co-mentioned in a diary entry in the near future – but in the long term their co-mentioning probability increases.

Finally, the effect of actor closure is mostly negative (the corresponding long-term effect is not significant in the model combining long-term and short-term effects). This means that if two actors a_1 and a_2 have been mentioned in (potentially different) past diary entries jointly with a common third actor a_3, then the probability that a_1 and a_2 will be co-mentioned in a future diary entry increases. This effect implies a tendency that overlapping dense groups of co-appearing actors are kept separtate over time.

Taken together, the findings from the actor-only models are suggesting that Andy Warhol tends to keep separate different but partially overlapping dense groups of actors.

Table 2. Label-only models. The statistics starting with "rec" assess the effect of recent previous events (exponential decay with half life 30 days). Those without that prefix assess the effect of all previous events without decay.

	label effects (long-term)	label effects (short-term)	long-term and short-term
sub.rep.0.1	0.893 (0.049)***		0.755 (0.050)***
sub.rep.0.2	0.229 (0.022)***		0.237 (0.024)***
sub.rep.0.3	−0.021 (0.011)$^{\cdot}$		−0.012 (0.011)
label.closure	0.623 (0.058)***		0.641 (0.058)***
rec.sub.rep.0.1		1.286 (0.025)***	0.252 (0.031)***
rec.sub.rep.0.2		0.320 (0.051)***	0.010 (0.060)
rec.sub.rep.0.3		−0.220 (0.097)*	−0.347 (0.115)**
rec.label.closure		−0.413 (0.067)***	−0.099 (0.076)
AIC	9593.453	14203.575	9521.823
Num. events	1,924	1,924	1,924
Num. obs.	194,224	194,224	194,224

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

4.2.2 Label-only model

Table 2 reports estimated parameters of three "label-only" models. The first model assesses the effect of all previous events without any decay (cumulative statistics). The second model assesses the effect of recent previous events where we let the influence of past events decay exponentially with a half life of 30 days. The third model combines cumulative and recent statistics. Looking at the model fit indicator AIC (lower values reveal better fit), we find that the long-term-only model fits better than the short-term-only model and that the joint model is the best of all three.

Subset repetition of order $(0,1)$ measures the effect of the past popularity of individual labels. We find that both, the long-term and the short-term effects of this statistic are positive. This means that if a term has been used more often by Andy Warhol to label diary entries then it is more likely that this same term will be used to label future entries. If a term has recently been used more often, then this has an additional effect increasing the likelihood that it labels further diary entries in the near future.

Subset repetition of order $(0,2)$ measures the effect of the past joint popularity of pairs of labels. The long term effect of this statistic counts how often two terms have been jointly used to label the same diary entry. We find that this long-term effect is positive so that pairs of terms co-labeling past diary entries more often are more likely to co-label future diary entries. The corresponding short-term effect is also positive in the model including only short-term effects but not significant in the model including long-term and short-term effects. Thus, recent joint use of two labels also has a positive effect on co-labeling probability in the near future – but when we control for long-term effects, the recency of past co-labeling has no additional effect.

Subset repetition of order $(0,3)$ measures the effect of the past joint popularity of triples of labels by considering how often three terms have been jointly used to label the same diary entry. We find that the long-term effect of this statistic in the long-term-only model is slightly negative ($p < 0.1$) but not significant in the model including both long-term and short-term effects. The corresponding short-term effect is negative in both models in which it is included. Thus, if three terms are jointly used to label a diary entry then, controling for all other effects, these three terms have a lower probability to co-label a diary entry in the near future – but in the long term their co-appearance probability increases.

Finally, the long-term effect of label closure is consistently positive. This means that if two terms w_1 and w_2 have been used to label (potentially different) past diary entries jointly with a common third term w_3, then the probability that w_1 and w_2 will jointly label a future diary entry increases. This effect implies a tendency that overlapping dense groups of co-appearing labels tend to merge over time. In contrast, the short-term effect of label closure is negative in the model including only short-term effects (and not significant in the third model in Table 2). This means that the merging of overlapping dense groups of co-appearing labels takes time – in the short-term, such different but related (overlapping) topics are kept separate.

Taken together, the findings from the label-only models are suggesting that Andy Warhol tends to keep separate different but related (partially overlapping) topics in the short term – but in the long term these related topics have a tendency to merge. It is striking that in the label-only models, the long-term model has a better fit than the short-term model – while this is reversed for the actor-only models. This seems to suggest that actors' popularity in the Andy Warhol diaries is short-lived ("*fifteen minutes of fame*") but topics seem to have long-term popularity.

4.2.3 Joint models

Independent combination of participant-effects and label-effects Table 3 reports parameters of models combining actor-effects and label-effects independently. This means that each of the included statistics is either defined only dependent on event participants or it is defined only dependent on terms labeling events. These models do not yet consider interaction effects such as whether certain actors are more likely to appear in events with certain labels. Table 3 reports findings on three models, the first containing only long-term effects, the second containing only short-term effects, and the third combining long-term and short-term effects. Considering the model fit indicator AIC we find that these combined models dependent on actors and labels improve the model fit of both, actor-only and label-only models, by a large margin. Thus, actors and labels seem to provide complementary information enabling models to specifiy the likelihood of future events.

Regarding the sign of effects there are only few parameters that significantly switch their sign compared to their values in the actor-only or label-only models.

Interaction of participant-effects and label-effects Table 4 reports parameters of models including actor-effects, label-effects, and effects depending on the interaction of effects and labels. These interaction effects are subset repetition of order $(1,1)$, $(2,1)$, and $(1,2)$, the mixed closure effects "actor closure by label" and "label closure by actor", as well as

Table 3. Models including an independent combination of actor-effects and label-effects. The statistics starting with "rec" assess the effect of recent previous events (exponential decay with half life 30 days). Those without that prefix assess the effect of all previous events without decay.

	actor+labels (long-term)	actor+labels (short-term)	long-term and short-term
sub.rep.1.0	0.511 (0.021)***		0.189 (0.020)***
sub.rep.2.0	0.852 (0.063)***		0.274 (0.026)***
sub.rep.3.0	0.659 (0.059)***		0.742 (0.044)***
actor.closure	−0.317 (0.053)***		0.054 (0.025)*
sub.rep.0.1	0.696 (0.048)***		0.603 (0.033)***
sub.rep.0.2	0.211 (0.026)***		0.232 (0.022)***
sub.rep.0.3	−0.005 (0.019)		0.031 (0.023)
label.closure	0.680 (0.061)***		0.502 (0.037)***
rec.sub.rep.1.0		0.861 (0.020)***	0.519 (0.019)***
rec.sub.rep.2.0		0.628 (0.044)***	0.315 (0.023)***
rec.sub.rep.3.0		−0.011 (0.024)	−0.237 (0.018)***
rec.actor.closure		−0.180 (0.031)***	−0.097 (0.016)***
rec.sub.rep.0.1		1.055 (0.030)***	0.165 (0.034)***
rec.sub.rep.0.2		0.276 (0.063)***	0.154 (0.041)***
rec.sub.rep.0.3		−0.256 (0.117)*	−0.116 (0.091)
rec.label.closure		−0.395 (0.089)***	−0.414 (0.054)***
AIC	5470.761	7770.960	4632.723
Num. events	1,924	1,924	1,924
Num. obs.	194,224	194,224	194,224

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

assortativity which is the product of subset repetition of order $(1,0)$ with subset repetition of order $(0,1)$. Table 4 reports findings on three models, the first containing only long-term effects, the second containing only short-term effects, and the third combining long-term and short-term effects. Considering the model fit indicator AIC we find that these interaction terms further improve the model fit over the models based on an independent combination of actor-effects and label-effects reported in Table 3.

Those effects that have been already included in previous models mostly seem to keep their signs and we focus our discussion on the newly included interaction effects.

Subset repetition of order $(1,1)$ measures how often a combination of one actor with one label co-appeared in a previous event. The long-term and short-term effects of this statistic are consistently positive, suggesting that Andy Warhol tends to associate specific persons with specific labels.

Subset repetition of order $(2,1)$ measures how often a combination of two actors with one label co-appeared in a previous event and subset repetition of order $(1,2)$ measures how often a combination of one actor with two labels co-appeared in a previous event. The long-term effects of these statistics are mostly insignificant (sub.rep.2.1 is significantly negative in the long-term-only model, though) but their short-term effects are consistently

Table 4. Joint models including actor-effects, label-effects, and effects dependent on a combination of actors and labels. The statistics starting with "rec" assess the effect of recent previous events (exponential decay with half life 30 days). Those without that prefix assess the effect of all previous events without decay.

	actor*labels (long-term)	actor*labels (short-term)	long-term and short-term
sub.rep.1.0	$0.151 (0.033)^{***}$		$-0.102 (0.047)^{*}$
sub.rep.2.0	$0.850 (0.057)^{***}$		$0.225 (0.084)^{**}$
sub.rep.3.0	$0.650 (0.051)^{***}$		$0.686 (0.072)^{***}$
actor.closure	$-1.173 (0.085)^{***}$		$-0.626 (0.099)^{***}$
sub.rep.0.1	$0.504 (0.052)^{***}$		$0.430 (0.050)^{***}$
sub.rep.0.2	$0.179 (0.029)^{***}$		$0.190 (0.031)^{***}$
sub.rep.0.3	$0.005 (0.030)$		$0.047 (0.015)^{**}$
label.closure	$0.166 (0.098)^{\cdot}$		$-0.110 (0.083)$
sub.rep.1.1	$0.552 (0.047)^{***}$		$0.187 (0.051)^{***}$
sub.rep.2.1	$-0.216 (0.049)^{***}$		$0.107 (0.071)$
sub.rep.1.2	$-0.051 (0.037)$		$0.047 (0.049)$
actor.closure.by.label	$1.589 (0.117)^{***}$		$1.329 (0.123)^{***}$
label.closure.by.actor	$0.559 (0.084)^{***}$		$0.780 (0.081)^{***}$
assortativity	$0.051 (0.019)^{**}$		$0.143 (0.020)^{***}$
rec.sub.rep.1.0		$0.811 (0.022)^{***}$	$0.522 (0.032)^{***}$
rec.sub.rep.2.0		$0.808 (0.063)^{***}$	$0.295 (0.065)^{***}$
rec.sub.rep.3.0		$0.055 (0.027)^{*}$	$-0.152 (0.039)^{***}$
rec.actor.closure		$-0.180 (0.033)^{***}$	$-0.044 (0.032)$
rec.sub.rep.0.1		$1.055 (0.034)^{***}$	$0.168 (0.046)^{***}$
rec.sub.rep.0.2		$0.424 (0.063)^{***}$	$0.346 (0.088)^{***}$
rec.sub.rep.0.3		$-0.165 (0.121)$	$-0.252 (0.144)^{\cdot}$
rec.label.closure		$-0.168 (0.082)^{*}$	$-0.149 (0.109)$
rec.sub.rep.1.1		$0.219 (0.050)^{***}$	$0.145 (0.059)^{*}$
rec.sub.rep.2.1		$-0.346 (0.064)^{***}$	$-0.210 (0.081)^{**}$
rec.sub.rep.1.2		$-0.365 (0.083)^{***}$	$-0.382 (0.155)^{*}$
rec.actor.closure.by.label		$-0.178 (0.062)^{**}$	$-0.010 (0.057)$
rec.label.closure.by.actor		$-0.815 (0.098)^{***}$	$-1.130 (0.131)^{***}$
rec.assortativity		$0.053 (0.015)^{***}$	$0.026 (0.018)$
AIC	4937.124	7578.435	4117.257
Num. events	1924	1924	1924
Num. obs.	194224	194224	194224

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

negative. This means that short-term repetition of a pair of actors in combination with a specific label, as well as short-term repetition of a pair of labels with a specific actor are rather unlikely. Note that short-term repetition of a pair of actors, independent of any specific label has an increased probability and the same holds for short-term repetition of a pair of labels independent of any specific actor.

Turning to the mixed closure effects, we find that their long-term effects are positive but their short term effects are negative. This means that two actors a_1 and a_2 who both have previous events labeled with a common term w_1 have an increased long-term probability to be co-mentioned in a future event. This contrasts with the negative effect of actor closure. Thus dense overlaping groups of actors are kept separate but might be merged through common labels.

The short-term effects of the mixed closure effects are negative. Andy Warhol seems to have a certain desire for novelty (novel combinations of actors or labels) in the short term.

Finally assortativity is positive. This means that events including popular persons are labeled with popular labels (assortativity).

5 Discussion

Most forms of subset repetition of the non-decaying statistics (long-term effects) are positive.

Actor-closure is negative but label-closure is positive. Actor-closure through label and label-closure through actor are also positive in the long-term. Thus, Andy Warhol maintains overlapping but separate groups of acquaintances. In contrast, labels have a connecting or bridging effect joining overlapping groups – at least in the long term.

Any kind of closure is rather negative in the short term. This means that in the short term Andy Warhol maintains disjoint groups of actors and labels – but in the long term, overlapping groups of labels have a tendency to merge. Overlapping groups of actors are typically kept separate (negative actor closure) but actors can be closed by labels (meaning that two actors who both have previous events labeled with a common term have an increased probability to be co-mentioned in a future event).

Assortativity is positive. This means that events including popular persons are labeled with popular labels.

Joint model (with interaction among persons and labels) is best model.

It is striking that in the label-only models, the long-term model has a better fit than the short-term model – while this is reversed for the actor-only models. This could suggest that actors' popularity in the Andy Warhol diaries is short-lived ("*fifteen minutes of fame*") but topics (clusters of terms labeling diary entries) seem to have more long-term effects.

6 Conclusion