Supplementary Information for

Temporal patterns of reciprocity in communication networks

S. Chowdhary, E. Andres, A. Manna, L. Blagojevic, L. Di Gaetano, G. Iñiguez*

*Corresponding author email: iniguezg@ceu.edu

S1 Temporal network concepts and reciprocity measures

Events

An *event* e_{ijt} is the interaction (communication) between source node i and target node j at time t.

Link

A link l_{ij} exists between nodes i and j if at least one event happens between them in the observation period, i.e. in the underlying static aggregate network (an undirected, simple network where link weights are discarded).

Node event sequence

A sequence of events of node i is, e.g., $S_i = \{e_{ijt_1}, e_{kit_2}, e_{imt_3}...e_{jit_T}\}$, that is, a series of T events where i is always involved, either as target or source node. Events are then out-links or in-links happening at some time t, i.e. the communication interactions that an ego has with its alters.

Link event sequence

A sequence of events in link l_{ij} between nodes i and j is, e.g., $S_{ij} = \{e_{ijt_1}, e_{jit_2}, e_{ijt_3}...e_{jit_T}\}$, that is, a series of T events involving i and j in any direction. Events are then out-links or in-links happening at some time t, i.e. the communication interactions between the pair of individuals.

Reciprocal events

A pair of events is reciprocal if events are consecutive and the direction of the second event is opposite to the first, i.e. (e_{ijt_1}, e_{jit_2}) where $t_2 > t_1$.

Reciprocal link

A reciprocal link contains at least one reciprocation (reciprocal event pair) in its sequence of events. We compute: (i) the number of reciprocations $E_{rec,ij}$ over link l_{ij}, relative to the number of consecutive event pairs on that link, $E_{ij} - 1$. By averaging over links, we obtain the reciprocation probability $p(E_{rec}) = \langle E_{rec,ij}/(E_{ij} - 1)\rangle_{ij}$. We also compute: (ii) the number of links with at least one reciprocation (l_{rec}) relative to the total number of links (L), $p(l_{rec}) = l_{rec}/L$.

Inter-event time

Inter-event time is the time span between consecutive events in a sequence. Inter-event times can be computed for both node and link event sequences.

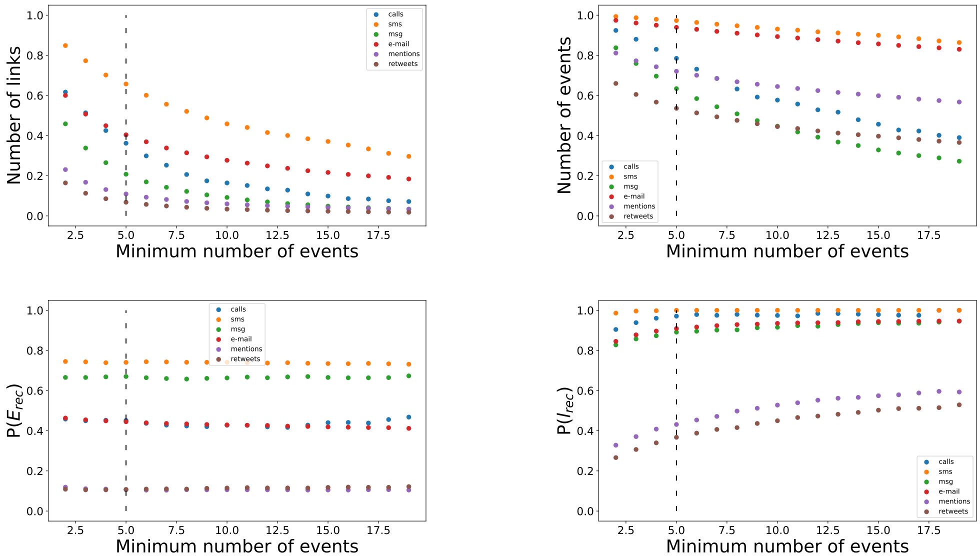


Figure S1. Network and temporal reciprocity measures as a function of filtering parameter. The dashed line represents the chosen value for the filter, where there is relative stability across all datasets.

Time gap

A time gap is the time elapsed between two successive events comprising a reciprocation or non-reciprocation between a pair of nodes. It is analogous to an inter-event time, but limited to a subset of inter-event times, i.e. those within (non-) reciprocations, instead of to all events.

Time gap burstiness

Time gap burstiness is defined as $B = (\sigma - \mu)/(\sigma + \mu)$, where μ and σ are, respectively, the mean and standard deviation of the time gaps within (non-)reciprocations. Time gap burstiness B ranges between -1 and +1, meaning time gaps are distributed either regularly or broadly in time. It can be computed for both node and link event sequences.

S2 Data description

We analyze several datasets of social contact between individuals from a wide range of studies in the temporal networks literature (see Table 1 in main text). Each dataset includes a time-ordered set of communication events between anonymized individuals i and j (according to hashed timestamps). From a dataset we construct a temporal network where a directed link l_{ij} appears if individual i initiates an event towards individual j at some point in time.

S2.1 List of datasets

Copenhagen Networks Study (calls & sms). Dataset of multi-channel, phone-enabled social interactions from the Copenhagen Networks Study (CNS) [1, 2]. The original study includes activity of roughly 1,000 individuals during 2012-2013 [1]. Data used here is a selected portion of the full dataset as described in [2]. The dataset includes events in two channels (disregarding Bluetooh data): call and short message logs between individuals, with data on timestamps of the call/message, anonymized user IDs, and call duration. We also delete missed calls, making the dataset smaller from the one in [2]. Data is publicly available via *figshare* in [3].

Data\Method	$p(E_{rec})$ NTS	NDS	NTSR	NDSR	$p(l_{rec})$ NTS	NDS	NTSR	NDSR
sms	+ 0.02 *	+ 0.02 *	+ 0.02 *	+ 0.02 *	1.00	1.00	+ 0.02 *	1.00
msg	+ 0.02 *	+ 0.02 *	+ 0.02 *	+ 0.02 *	1.00	1.00	+ 0.20	- 0.02 *
email	+ 0.02 *	+ 0.02 *	+ 0.02 *	+ 0.38	1.00	1.00	+ 0.02 *	- 0.02 *
retweets	- 0.02 *	- 0.02 *	- 0.02 *	- 0.02 *	1.00	1.00	- 0.02 *	- 0.02 *
mentions	- 0.03 *	- 0.03 *	- 0.03 *	- 0.03 *	1.00	1.00	+ 0.03 *	- 0.03 *

Table S1. Signed (one-tailed) p-values of the temporal reciprocity measures $p(E_{rec})$ and $p(l_{rec})$ between the studied datasets and four null models shuffling interaction events. An asterisk ($*$) indicates a statistical significant difference between data and model (significance level $\alpha = 0.03$). The sign of the p-value is chosen with respect to median values: for positive (green) p-values, empirical measures are higher than the median of shuffling results, and viceversa for the negative (red) p-values. Null models are denoted by NTS (Network Shuffling Timestamps), NDS (Node Shuffling Timestamps), NTSR (Network Rewiring and Shuffling Timestamps), and NDSR (Node Rewiring and Shuffling Timestamps). The calls dataset is not included due to its small size after filtering (see Section S3). We see more positive than negative statistically significant p-values, implying that temporal reciprocity is not reproduced by random mechanisms. The NTSR model randomizes the timeline of social interactions of an individual and the identities of its neighbours, erasing its structural and temporal memory. Positive p-values for NTSR thus suggest memory as a potentially relevant mechanism for reciprocal interaction in social communication. There is also a notable difference in p-value sign between conversation (sms, msg, email) and broadcasting (retweets, mentions) channels, pointing to the distinct roles of bidirectional vs. unidirectional exchange.

Twitter (retweets & mentions). The Twitter dataset has been collected through the Twitter API from 2018-02-18 to 2018-12-18. The request was limited to tweets that contained at least one selected keyword (*vaccin*[1], vax, libertàdiscelta, libertadiscelta, ddl770, trivalente, #mmr), with no limitation to the location nor to the language of the tweet. From the list of tweets collected, two directed temporal networks have been created according to the type of action (retweets and mentions). While the reply network would be particularly interesting in terms of reciprocal interactions, the low number of replies in the dataset prompted us to disregard this network, especially after filtering (see Section S3). In all Twitter networks, nodes represent users, while a link goes from i to j if user i mentioned j or retweeted a post of user j.

College messages (msg). This dataset is comprised of private messages sent on an online social network at the University of California, Irvine [4] from April to October 2004. Users join an online community, intended to help students communicate with their friends, and to meet new people. First, members have to create an account by filling in some information; then they can search the network for others and then initiate conversation based on profile information. There is a total of 1,899 users, who have exchanged 59,835 online messages.

EU research institution (email). The network was generated using email data from a large European research institution [5] from October 2003 to May 2005. It comprises 3038531 emails, sent from 287755 different email addresses. All information has been anonymized. Emails only represent communication between institution members (the core), and the dataset does not contain incoming messages from or outgoing messages to the rest of the world.

[1]The string *'*vaccin*'* allows us to capture every possible declination and compounding of the Italian word *vaccino* (vaccine) and the verb *vaccinare* (vaccinate).

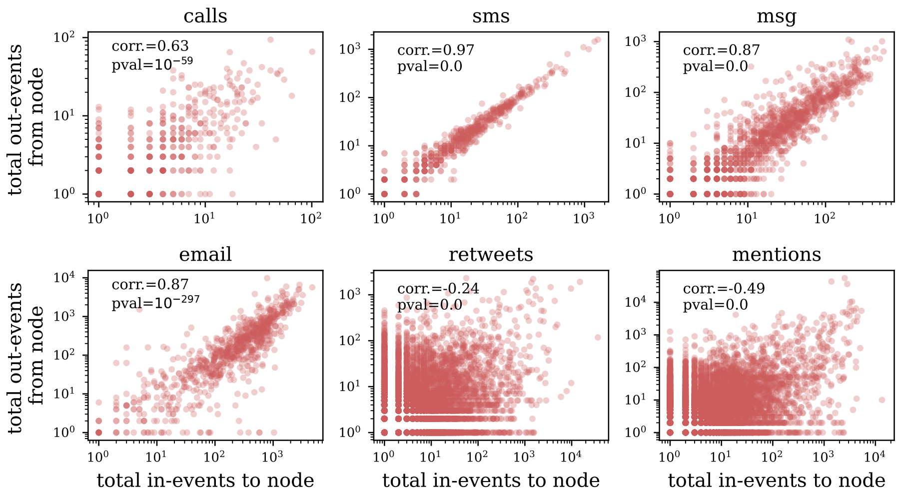


Figure S2. Scatter plot of number of in- and out-events per node (i.e. node in- and out-strengths in aggregated data) for all networks before filtering. We show the associated Spearman rank correlation coefficients (corr) and p-values (pval; small values reject the null hypothesis that the correlation is zero).

S3 Event filtering

To ensure statistically significant results, we filtered the original networks in such a way that each remaining link has a minimum number of events. This decision is mainly motivated by previous works [6–8], where the authors claim that the quantity of interactions between people is a good proxy for social tie strength. We then focus on significant ties, meaning that the two involved individuals have exchanged some minimum number of messages (events) between themselves.

We have to pick a value equal to or higher than 3, since we need at least 2 inter-event times to compute their standard deviation, which we use in our exploratory results. Based on our sensitivity analysis, where we varied the minimum number of events per link (Fig. S1), we fix this filtering parameter to 5, marked by the vertical dashed line. This is the minimum value for which most measures stabilize, meaning that their rate of change is relatively low, compared to the smaller values of other filtering parameter values (3 or 4 minimum events per link).

S4 Null models

We utilize null models to assess how much our measures differ when computed in the original empirical networks, compared to their randomized versions. By combining two shuffling types and two resolution levels, we employ four different null models (randomized reference models):

- **NTS (Network Shuffling Timestamps)**. All events occur at their original links, between the same nodes, while each time of occurrence for every event is sampled without replacement from the set of all times of occurrence for the network. This means that all events occur between the same nodes, but each of them at a different, randomly sampled time.

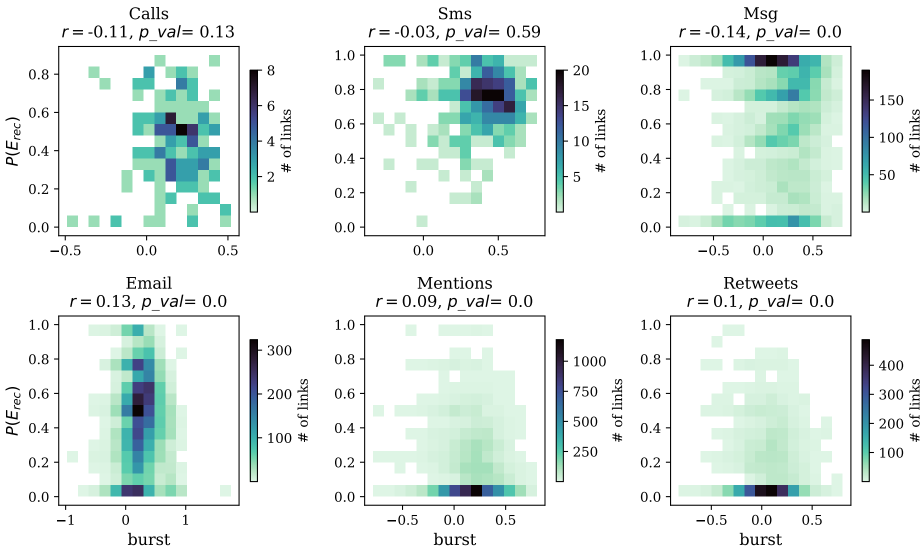


Figure S3. Joint distribution of number of links with given values of standard burstiness (denoted $burst$) and probability of reciprocation $p(E_{rec})$ for all networks considered. We include the Spearman rank correlation coefficient (r) and the associated p-value (p_val; small values reject the null hypothesis that the correlation is zero).

- **NDS (Node Shuffling Timestamps)**. All events occur at their original links, between the same nodes, but each event randomly obtains a time of occurrence sampled with a replacement from the set of events that belong to its neighboring nodes. This means that all events occur between the same nodes, but at different times, randomly determined from their initial neighborhood.

- **NTSR (Network Rewiring and Shuffling Timestamps)** Network links are randomly reassigned, with in- and out-degrees mostly conserved for each node (unless self-loops occur, which happens in the configuration model). Additionally, the time of occurrence of each event is randomly sampled without replacement from the set of all events in the network. This means that the same number of events occurs, but between randomly sampled nodes within a network and at randomly sampled times.

- **NDSR (Node Rewiring and Shuffling Timestamps)** All links are randomly reassigned, with in- and out-degrees conserved for each node, and each node randomly obtains a link sampled with replacement from the set of its initial neighbors. Additionally, the time of occurrence for each event is randomly sampled without replacement from the set of all neighbor's events. This means that the same number of events occurs, but between randomly sampled nodes from their original neighborhoods and at times randomly sampled from their neighbors.

All shuffling methods are implemented either $N_{trials} = 50$ times for all conversation channels, or 39 times for Twitter (due to the larger size of the dataset), leading to probability distributions of both temporal reciprocity measures [$p(E_{rec})$ and $p(l_{rec})$] compatible with the null hypothesis of each shuffling method. Results of a Lilliefors test for normality on these probability distributions show that roughly 64% of the distributions are not normal with some level of statistical significance, barring us from comparing data and null models via, e.g., z-scores. We can, however, calculate p-values directly as the probability that the null hypothesis introduced by each shuffling method produces temporal reciprocity

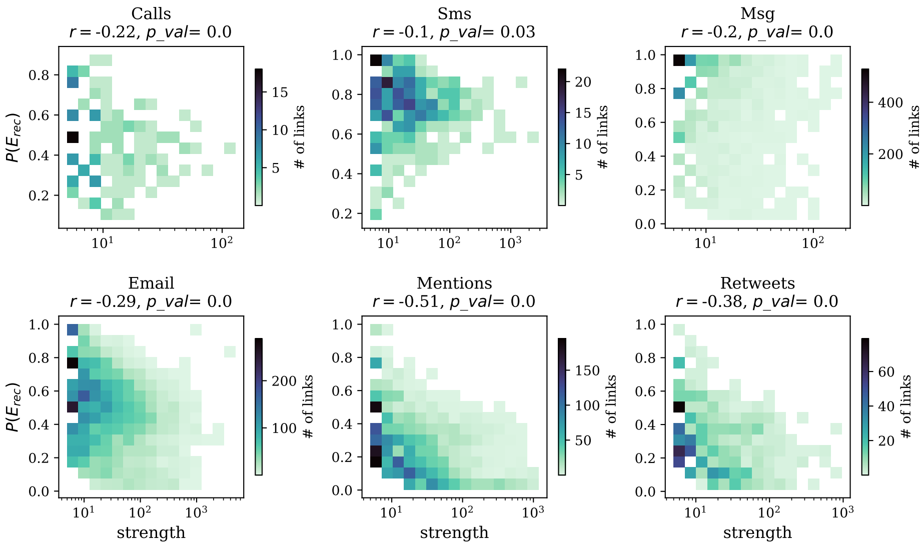


Figure S4. Joint distribution of number of links with given values of communication frequency (number of events over link, denoted strength) and probability of reciprocation $p(E_{rec})$ for all networks considered after having filtered for links with $p(E_{rec}) > 0$. We include the Spearman rank correlation coefficient (r) and the associated p-value (p_val; small values reject the null hypothesis that the correlation is zero).

values at least as extreme as the empirical value in each dataset. Explicitly, a signed (one-tailed) p-value is calculated as the fraction of realizations produced by a shuffling method that are either larger (positive sign) or smaller (negative sign) than the empirical value (where the positive or negative sign is chosen if the median of the null model distribution is smaller or larger than the empirical value, respectively). For example, if a link in a dataset has $p(E_{rec}) = 0.2$ and the median of the corresponding null distribution for a given shuffling method is 0.5 [i.e. the median is larger than $p(E_{rec})$], then the p-value is the fraction of realizations with values smaller than $p(E_{rec})$ and its chosen sign is negative. As the number N_{trials} of realizations is relatively small (50 and 39), for very low p-values we choose instead the conservative upper bound $1/N_{trials}$. Resulting p-value estimates are shown in Table S1.

S5 Correlation between node in- and out-strengths

Fig. S2 shows a scatter plot of the number of in- and out-events for all nodes in each of the studied networks (i.e. the node in- and out-strengths in the aggregated weighted network). We compute the Spearman rank correlation coefficient r, which shows positive and significant correlations between these two quantities for conversation channels, thus revealing the presence of reciprocal relationships. Apparently, r is higher for conversation channels than for Twitter (where it is negative instead), further highlighting the use of Twitter as a broadcasting platform, not a place for one-to-one communication.

S6 Correlation between temporal reciprocity and other network properties

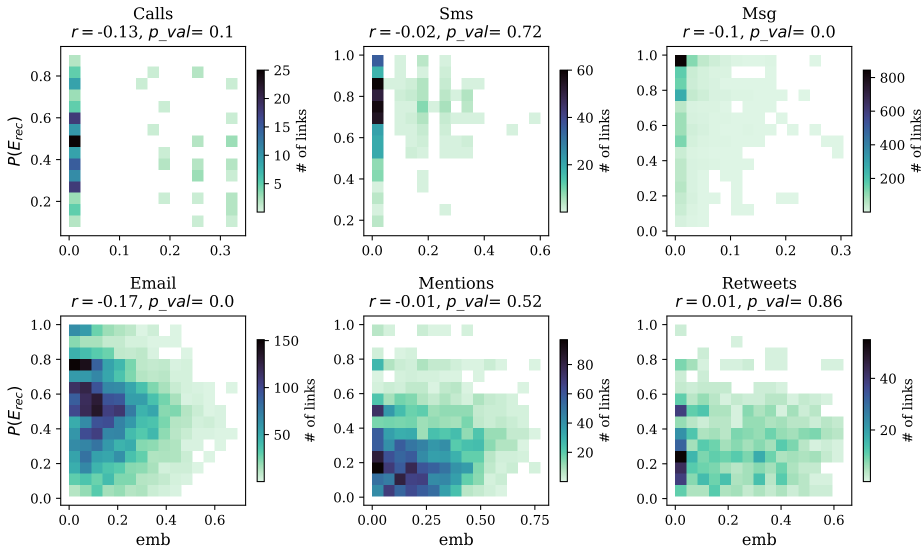


Figure S5. Joint distribution of number of links with given values of tie embeddedness [denoted *emb*; see Eq. (S1)] and probability of reciprocation $p(E_{rec})$ for all networks considered after having filtered for links with $p(E_{rec}), emb > 0$. We include the Spearman rank correlation coefficient (r) and the associated p-value (p_val; small values reject the null hypothesis that the correlation is zero).

S6.1 Reciprocity vs standard burstiness

Fig. S3 shows the joint distribution of the number of links with given values of standard burstiness and probability or reciprocation $p(E_{rec})$ for each network considered. The Spearman correlation coefficient r indicates no significant correlation between these two quantities.

S6.2 Reciprocity vs communication frequency and tie embeddedness

The reciprocity of communication between two individuals may in principle be related to other properties of the aggregated communication network. For instance, we may expect that temporal reciprocity is related to the number of common friends between them, or to the intensity of their communication activity. We measure the correlation between $p(E_{rec})$ and two network properties: (i) communication frequency, the number of interaction events between nodes i and j, and (ii) tie embeddedness, the number of common neighbors of i and j, namely

$$emb = \frac{|S_i \cap S_j|}{|S_i \cup S_j|} \tag{S1}$$

where S_i and S_j are respectively the set of individuals i and j have ever been in contact with.

Fig. S4 and Fig. S5 show respectively the joint distributions of the number of links with given values of $p(E_{rec})$ and either communication frequency or tie embeddedness. In both cases we find no significant correlation between each pair of quantities. In other words, high and low levels of temporal reciprocity can be found regardless of the amount of communication activity and structural cohesion around a link, implying that our measures of temporal reciprocity give additional information on the properties of social ties beyond known measures like communication frequency and tie embeddedness.

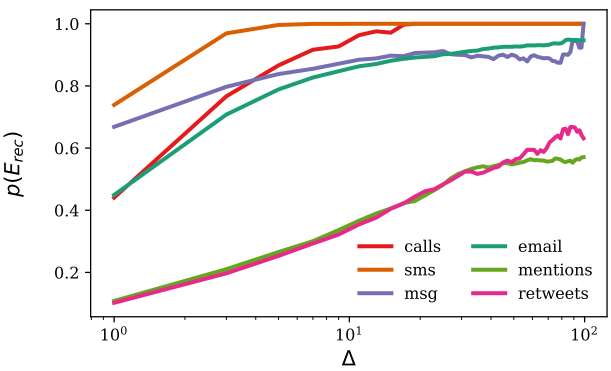


Figure S6. Probability of reciprocation $P(E_{rec})$ as a function of time window length Δ. As the number of consecutive evens considered simultaneously $(\Delta + 1)$ increases from $\Delta = 1$ (strict reciprocity) to $\Delta > 1$ (relaxed reciprocity), the probability of finding at least one reciprocal event between them increases as well. The rate of change is also faster in conversation channels, implying the existence of reciprocal (but non-consecutive) event pairs.

S6.3 Reciprocity vs balance

The reciprocation probability $p(E_{rec})$ between nodes i and node j has an upper bound related to balance b. Taking n_{ij} and n_{ji}, respectively, as the number of events from i to j and from j to i, we assume $n_{ij} < n_{ji}$ without loss of generality. The temporal configuration of events that maximizes $p(E_{rec})$ is the one that maximizes the change of direction in interactions. In that case, the events from i to j, the minority, are surrounded by events from j to i, which creates two reciprocations by each event from i to j. Then, the maximum probability of reciprocation is

$$p_{max}(E_{rec}) = \frac{2\ min(n_{ij}, n_{ji})}{n_{ij} + n_{ji}}. \tag{S2}$$

Since balance is defined as

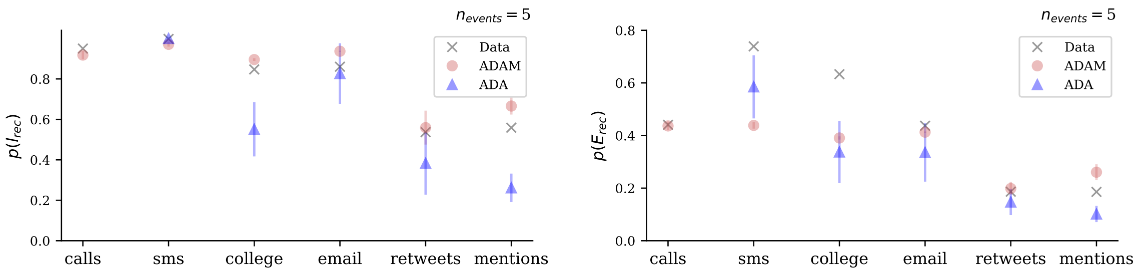
$$b = \frac{max(n_{ij}, n_{ji})}{n_{ij} + n_{ji}} = 1 - \frac{min(n_{ij}, n_{ji})}{n_{ij} + n_{ji}}, \tag{S3}$$

then the relation between the maximum probability of reciprocation and balance is

$$p_{max}(E_{rec}) = 2(1 - b). \tag{S4}$$

S7 Relaxing the reciprocity condition

The definition of temporal reciprocity considered in the present work is particularly strict as it only refers to pairs of consecutive interactions. Here we relax this condition by looking at the reciprocity of sets of n consecutive interactions. We introduce a parameter Δ controlling the number of consecutive events between nodes i and j that we take into account simultaneously. For each link event sequence S_{ij} (see Section S1), we consider a sliding window of length Δ that we slide in steps of one event. We consider a set of $\Delta + 1$ consecutive interactions as reciprocal if at least two of its events have opposite directions, i.e. if there is at least one reciprocal interaction in the considered set. Then, $\Delta = 1$ corresponds to our strict definition of temporal reciprocity used throughout this work, while $\Delta > 1$ implies a more relaxed condition for reciprocity, where neighboring events may appear.



]

Figure S7. Fraction of links having at least one reciprocation, $p(l_{rec})$ (left), and fraction of reciprocations $p(E_{rec})$ (right), for several empirical communication channels (Data), as well as in synthetic temporal networks fitted by the ADAM and ADA models. The filter threshold is higher than in Fig. 4 in the main text ($n_{events} = 5$).

Fig. S6 shows how $p(E_{rec})$ changes as a function of Δ, i.e. as the reciprocity condition is relaxed. The probability to have a reciprocal interaction increases as we consider more events simultaneously. Interestingly, the rate at which $p(E_{rec})$ increases is different for the datasets under analysis. Namely, in conversation channels like sms and msg, $P(E_{rec})$ grows faster than in the broadcasting channels mentions and retweets. A fast growth of temporal reciprocity with Δ indicates that potential reciprocal events may not be located next to each other in a link event sequence, but that a relative short time window can detect them. Conversely, slow growth implies that reciprocity mostly happens between close events.

S8 Effect of filtering in ADAM model

Increasing the filter threshold to $n_{events} = 5$ does not change qualitatively the results obtained in Fig. 4 of the main text (see Fig. S7).

S9 Effect of observation time scale on reciprocity levels in data and performance of activity-driven models

The levels of temporal reciprocity seen in empirical data of human communication activity may in principle depend on the length of the observation period during which data is recorded. In Fig. S8 we investigate the robustness of our measures of temporal reciprocity [$p(E_{rec})$ and $p(l_{rec})$] by artificially considering increasing amounts of time-ordered data for a given communication channel. We vary the number of events considered from 10000 to 30000 (in steps of 5000) for each dataset and find that such time scale of observation does not qualitatively change our results. To give a sense of the change in time scales across communication channels, in Table S2 we additionally show the average time (in days) it takes for such 5000 events to occur.

References

[1] A. Stopczynski, V. Sekara, P. Sapiezynski, A. Cuttone, M. M. Madsen, J. E. Larsen, and S. Lehmann, "Measuring large-scale social networks with high resolution," *PloS ONE*, vol. 9, no. 4, p. e95978, 2014.

Dataset	Time (days) for 5000 events to occur
calls	50
sms	6
msg	14
email	8
retweets	6
mentions	2

Table S2. Average time (in days) taken for 5000 events to occur in each communication channel considered in our analysis, which quantifies the relative time scales of our datasets. Values are approximated to the nearest integer.

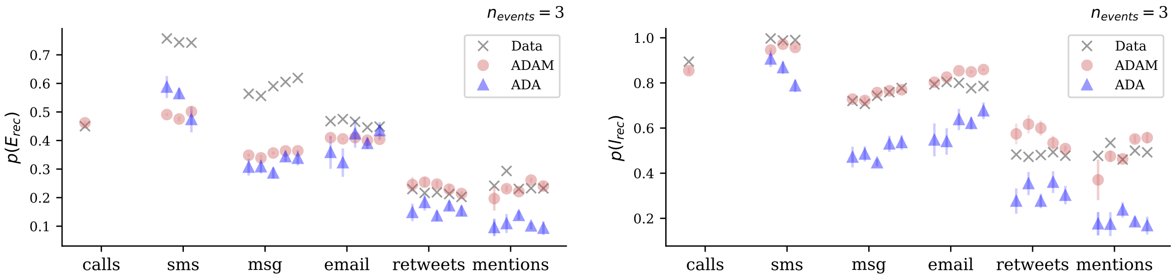


Figure S8. Fraction of reciprocations $p(E_{rec})$ (left) and fraction of links having at least one reciprocation, $p(l_{rec})$ (right), for several empirical communication channels (Data), as well as in synthetic temporal networks fitted by the ADAM and ADA models. For each dataset, we consider an increasing number of events from 10000 to 30000 in steps of 5000 (plotted left to right in ascending order). Some plots (calls, sms) have less than 5 points because the total number of events there are less than 30000. For instance, calls only has 3234 events (< 5000), so we show one point, and sms has 24333 events, so we show points corresponding to 10000, 15000 and 20000 events. Reciprocity levels do not qualitatively change for data or models, signaling that our reciprocity measures are robust to the time scale of the observation period used to record data.

[2] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, "Interaction data from the Copenhagen Networks Study," *Sci. Data*, vol. 6, no. 1, pp. 1–10, 2019.

[3] "The copenhagen networks study interaction data." `https://figshare.com/articles/dataset/The_Copenhagen_Networks_Study_interaction_data/7267433/1`.

[4] P. Panzarasa, T. Opsahl, and K. M. Carley, "Patterns and dynamics of users' behavior and interaction: Network analysis of an online community," *J. Am. Soc. Inf. Sci. Tech.*, vol. 60, no. 5, pp. 911–932, 2009.

[5] A. Paranjape, A. R. Benson, and J. Leskovec, "Motifs in temporal networks," in *Proceedings of the tenth ACM international conference on web search and data mining*, pp. 601–610, 2017.

[6] M. S. Granovetter, "The strength of weak ties," *Am. J. Soc.*, vol. 78, no. 6, pp. 1360–1380, 1973.

[7] P. V. Marsden and K. E. Campbell, "Measuring tie strength," *Social Forces*, vol. 63, no. 2, pp. 482–501, 1984.

[8] C. Wang, O. Lizardo, D. Hachen, A. Strathman, Z. Toroczkai, and N. V. Chawla, "A dyadic reciprocity index for repeated interaction networks," *Netw. Sci.*, vol. 1, no. 1, pp. 31–48, 2013.