Information flow reveals prediction limits in online social activity

James P. Bagrow[1,2,*], Xipei Liu[1,2], and Lewis Mitchell[1,2,3]

[1]Department of Mathematics & Statistics, University of Vermont, Burlington, VT, United States
[2]Vermont Complex Systems Center, University of Vermont, Burlington, VT, United States
[3]School of Mathematical Sciences, North Terrace Campus, The University of Adelaide, SA 5005, Australia
[*]Corresponding author. Email: james.bagrow@uvm.edu, Homepage: bagrow.com

August 15, 2017

Abstract Modern society depends on the flow of information over online social networks, and popular social platforms now generate significant behavioral data. Yet it remains unclear what fundamental limits may exist when using these data to predict the activities and interests of individuals. Here we apply tools from information theory to estimate the predictive information content of the writings of Twitter users and to what extent that information flows between users. Distinct temporal and social effects are visible in the information flow, and these estimates provide a fundamental bound on the predictive accuracy achievable with these data. Due to the social flow of information, we estimate that approximately 95% of the potential predictive accuracy attainable for an individual is available within the social ties of that individual only, without requiring the individual's data.

The flow of information in online social platforms is now a significant factor in protest movements, national elections, and rumor and misinformation campaigns [1, 2, 3]. Data collected from these platforms are a boon for researchers [4] but also a source of concern for privacy, as the social flow of predictive information can reveal details on both users and non-users of the platform [5]. Information flow on social media has primarily been studied *structurally* (for example, by tracking the movements of keywords [2, 6, 7, 8] or adoptions of behaviors [9, 10, 11]) or *temporally*, by applying tools from information theory to quantify the information contained in the timings of user activity, as temporal relationships between user activity reflect underlying coordination patterns [12, 13]. Yet neither approach considers the full extent of information available, both the complete language data provided by individuals and their temporal activity patterns. Here we unify these two primary approaches, by applying information-theoretic estimators to a collection of Twitter user activities that fully incorporate language data while also accounting for the temporal ordering of user activities.

We gathered a dataset of $n = 927$ users of the Twitter social media platform. Users were selected who wrote in English, were active for at least one year, and had comparably sized social networks. We applied both computational tools and human raters to help avoid bots and non-personal accounts. For each user, we retrieved all of their public posts excluding retweets (up to the 3200 most recent public posts, as allowed by Twitter). Examining these texts, we determined each user's 15 most frequent Twitter contacts and gathered the texts of those users as well, providing us ego-alter pairs. See Supporting Material (SM) for full details on data collection, filtering, and processing.

The ability to accurately profile and predict individuals is reflected in the predictability of their written text. The predictive information contained within a user's text can be characterized by three related quantities, the entropy rate h, the perplexity 2^h, and the predictability Π. The entropy rate quantifies the average uncertainty one has about future words given the text one has already observed (Fig. 1A). Higher entropies correspond to less predictable text and reflect individuals whose interests are more difficult to predict. In the context of language modeling, it is also common to consider the perplexity. While the entropy rate specifies how many bits h are needed on average to express subsequent words given the preceding text, the perplexity tells us that our remaining uncertainty about those unseen words is equivalent to that of choosing uniformly at random from among 2^h possibilities. For example, if $h = 6$ bits (typical of individuals in our dataset), the perplexity is 64 words, a significant reduction from choosing randomly over the entire vocabulary (social media users have ≈5000-word vocabularies on average; see SM). Lastly, the predictability Π, given via Fano's inequality [14], is the probability that an *ideal* predictive algorithm will correctly predict the subsequent word given the preceding text. Repeated, accurate predictions of future words indicate that the available information can be used to build profiles and predictive models of a user and estimating Π allows us to fundamentally bound the usefulness of the information present in a user's writing without depending on the results of specific predictive algorithms.

Information theory has a long history of estimating the information content of text [15, 16, 17, 18]. Crucially, information is available not just in the words of the text but in their order of appearance. We applied a nonparametric entropy estimator that incorporates the full sequence structure of the text [18]. This estimator has been proved to converge asymptotically to the true entropy rate for stationary processes and has been applied to human mobility data [19]. See SM for details.

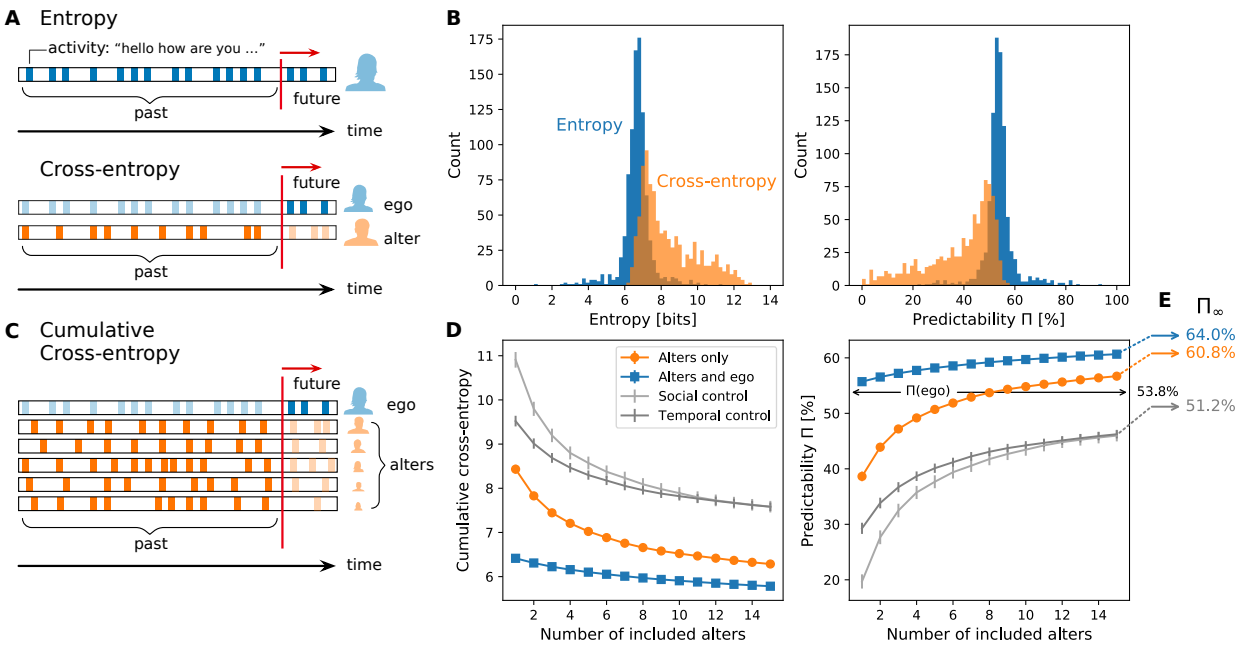

Figure 1: Information and predictability in online social activity. (**A**) A user posts written text over time and we would like to predict their subsequent words given their past writing. Treating each user's posts as a contiguous text stream, the entropy rate tells us how uncertain we are about a user's future writing given their past. To study information flow, the cross-entropy rate tells us how much information about the future text of one user (the ego) is present in the past text of another user (the alter). (**B**) Most users have entropies and predictabilities in a well-defined range, whereas the cross-entropies and associated predictabilities indicate a broad variety of social information flow levels. (**C**) Predictive information may be available in the pasts of multiple alters, so we computed the cumulative cross-entropy as we included more alters in order of most to least frequently contacted. (**D**) As the past activities of more alters are used to predict the ego, more information is available and the entropy drops and predictability rises. Including the ego's past with the alters shows that the alters provided non-redundant predictive information. (**E**) Extrapolating beyond our data window estimates the prediction limit Π_∞ of online activity.

The text streams of the egos were relatively well clustered around $h \approx 6.6$ bits, with most falling between 5.5–8 bits (Fig. 1B). Equivalently, this corresponds to a perplexity range of ≈45–256 words, far smaller than the typical user's ≈5000-word vocabulary, and a mean predictability of ≈53%, quite high for predicting a given word out of ≈5000 possible words on average. We found this typical value of information comparable to other sources of written text, but social media texts were more broadly distributed—individuals were more likely to be either highly predictable or highly unpredictable compared with formally written text (see SM).

Next, instead of asking how much information is present in what the ego has previously written regarding what the ego will write in the future, we ask how much information is present on average in what the ***alter*** has previously written regarding what the ego will write in the future (Fig. 1A). If there is consistent, predictive

information in the alter's past about the ego's future, especially beyond the information available in the ego's own past, then there is evidence of information flow.

Replacing the ego's past writing with the alter's past converts the entropy to the *cross-entropy* (see SM). The cross-entropy is always greater than the entropy when the alter provides less information about the ego than the ego, and so an increase in cross-entropy tells us how much information we lose by only having access to the alter's information instead of the ego's. Indeed, estimating the cross-entropy between each ego and their most frequently contacted alter (Fig. 1B), we saw higher cross-entropies spanning from 6–12 bits (equivalently, perplexities from 64–4096 words or predictabilities spread from 0–60%). While less frequently contacted alters provided less predictive information than alters in close contact (see SM), even for the closest alters there was a broader range of cross-entropies than the entropies of the egos themselves. This implies a diversity of social relationships: sometimes the ego is well informed by the alter, leading to a cross-entropy closer to the ego's entropy, while other times the ego and alter exhibit little information flow.

Thus far we have examined the information flow between the ego and individual alters, but actionable information regarding the future of the ego may be embedded in the combined pasts of multiple alters (Fig. 1C). To address this, we generalized the cross-entropy estimator to multiple text streams (see SM). We then computed the cross-entropies and predictabilities as we successively accumulated alters in order of decreasing contact volume (Fig. 1D). As more alters were considered, cross-entropy decreased and predictability increased, which is sensible as more potential information is available. Interestingly, with 8–9 alters, we observed a predictability of the ego given the alters at or above the original predictability of the ego alone. As more alters were added, up to our data limit of 15 alters, this increase continued. Paradoxically, this indicated that there is potentially more information about the ego within the total set of alters than within the ego itself.

To understand this apparent paradox, we need to address a limitation with the above analysis: it does not incorporate the ego's own past information. It may be that the information provided by the alters is simply redundant when compared with that of the ego. Therefore, we simply included the ego's past alongside the alters, generalizing the estimator to an entropy akin to a transfer entropy [20, 21], a common approach to studying information flow. This entropy is computed in the "Alters and ego" curves in Fig. 1D. A single alter provided a small amount of extra information beyond that of the ego, $\approx$1.9% more predictability. This

provided us a quantitative measure of the extent of information flow between individual users of social media Beyond the most frequently contacted alter, as more alters were added this extra predictability grew: at 15 alters and the ego there was ≈6.9% more predictability than via the ego alone. Furthermore, the information provided by the alters without the ego is strictly less than the information provided by the ego and alters together, resolving the apparent paradox.

However, this extra predictability also appeared to saturate, and eventually adding more alters will not provide extra information. This observation is compatible with Dunbar's number which uses cognitive limits to argue for an upper bound on the number of meaningful ties an ego can maintain (≈150 alters) [22]. The question now becomes, given enough ties what is the upper bound for the predictability?

To extrapolate beyond our data window, we fitted a nonlinear saturating function to the curves in Fig. 1D, (see SM for details and validation of our extrapolation procedure). From fits to the raw data, we found a limiting predictability given the alters of $\Pi_\infty = 60.8\% \pm 0.691\%$ (Fig. 1E). Of course, egos will not have an infinite number of alters, so a more plausible extrapolation point may be to Dunbar's number: $\Pi_{150} \approx 60.3\%$, within the margin of error for Π_∞, indicating that saturation of predictive information has been reached. Similarly, extrapolating the predictability including the ego's past gives $\Pi_\infty = 64.0\% \pm 1.54\%$ ($\Pi_{150} = 63.5\%$).

These extrapolations showed that significant predictive information was available in the combined social ties of individual users of social media. In fact, there is so much social information that an entity with access to all social media data will have only slightly more potential predictive accuracy (≈64% in our case) than an entity which has access to the activities of an ego's alters but not to that ego (≈61%). This may have distinct implications for privacy: if an individual forgoes using a social media platform or deletes her account, yet her social ties remain, then that platform owner potentially possesses $95.1\% \pm 3.36\%$ of the achievable predictive accuracy of the future activities of that individual.

Two issues can affect the cross-entropy as a measure of information flow. The first is that the predictive information may be due simply to the structure of English: commonly repeated words and phrases will represent a portion of the information flow. The second is that of a common cause: egos and alters may be independently discussing the same concepts. This is particularly important on social media with its emphasis on current events [23].

To study these issues, we constructed two types of controls. The first randomly pairs users together by shuffling alters between egos. The second constructed pseudo-alters by assembling, for each real alter, a random set of posts made at approximately the same times as the real alter's posts, thus controlling for temporal confounds. Both controls used real posted text and only varied the sources of the text. As shown in Fig. 1D, the real alters provided more social information than either control. There was a decrease in entropy as more control alters were added, but the control cross-entropy remained above the real cross-entropy. We also observed that for a single alter the temporal control had a lower cross-entropy than the social control and therefore temporal effects provide more information than social effects (underscoring the role of social media as a news platform [23]), although both controls eventually converge to a limiting predictability of $\approx$51%.

Given the importance of temporal information in online activity, to what extent is this reflected in the information flow? Do recent activities contain most of the predictive information or are there long-term sources of information? To estimate recency effects, we applied a censoring filter to the ego's text stream, removing at each time period the text written in the previous ΔT hours and measuring how much the mean predictability decreased compared with the mean predictability including the recent text. Increasing ΔT decreased Π, especially evident when removing the first 3–4 hours (Fig. 2A): we found an average decrease in predictability of $\approx$1.4% at 4 hours. This loss in predictability relative to the uncensored baseline is comparable to the gain from the rank-1 alter we observed in Fig. 1D. In other words, close alters tended to contain a quantity of information about the ego comparable to the information within just a few hours of the ego's own recent past. Beyond 24 hours the predictability loss continued approximately linearly (SM). We then applied this censoring procedure to the alters alone and the alters combined with the ego, excluding their recent text and measuring how the cross predictability changed on average from their respective baselines. We found a similar drop in predictability during the first few hours, but then a more level trend than when censoring the ego alone. This leveling off showed that less long-term information was present in the alters' pasts than within the ego's.

Next we studied recency by the activity frequencies of alters and egos. Individuals who post frequently to social media, keeping up on current events, may provide more predictive information about either themselves or their social ties than other, infrequent posters. We found that the self-predictability of users was actually

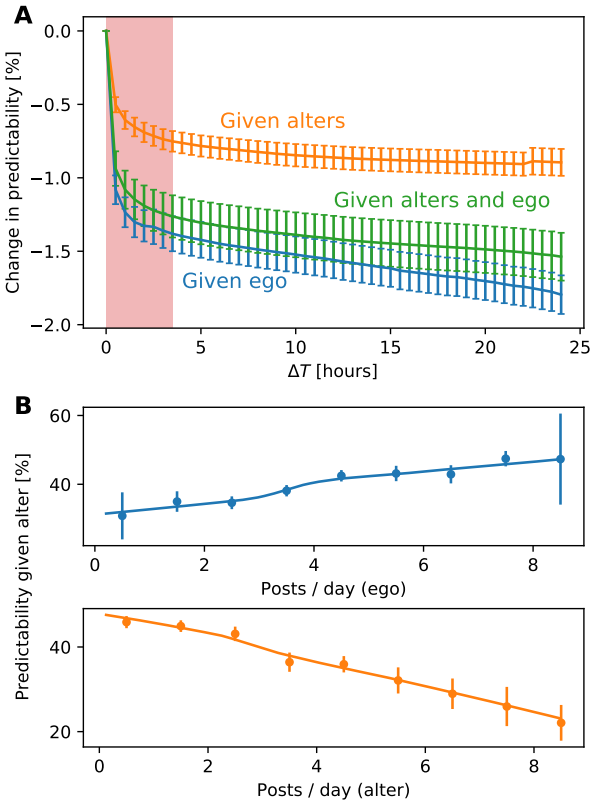

Figure 2: Recency of information. (**A**) Removing the most recent ΔT hours of activity, most predictive information about the ego is contained in the most recent 3–4 hours (shaded region). In all cases information extends backwards beyond these time intervals, but the ego contains more long-range past information than the combined alters alone. (**B**) Egos who post more frequently are more predictable from their alter than egos who post less frequently, whereas frequently posting alters provide less information about their egos than alters who post less often.

independent of activity frequency (SM) but there were strong associations between activity frequency and social information flow: egos who posted 8 times per day on average were ≈17% more predictable given their alters than egos who posted once per day on average (Fig. 2B). Interestingly, this trend reversed itself when considering the activity frequencies of the alters: alters who posted 8 times per day on average were ≈23% less predictive of their egos than alters who posted once per day on average. Highly active alters tended to inhibit information flow, perhaps due to covering too many topics of low relevance to the ego.

Information flow reflects the social network and social interaction patterns (Fig. 3). We measured information flow for egos with more popular alters compared with egos with less popular alters. Alters with more social ties provided less predictive information about their egos than alters with fewer ties (Fig. 3A). The decrease in predictability of the ego was especially strong up to alters with ≈400 ties, where the bulk

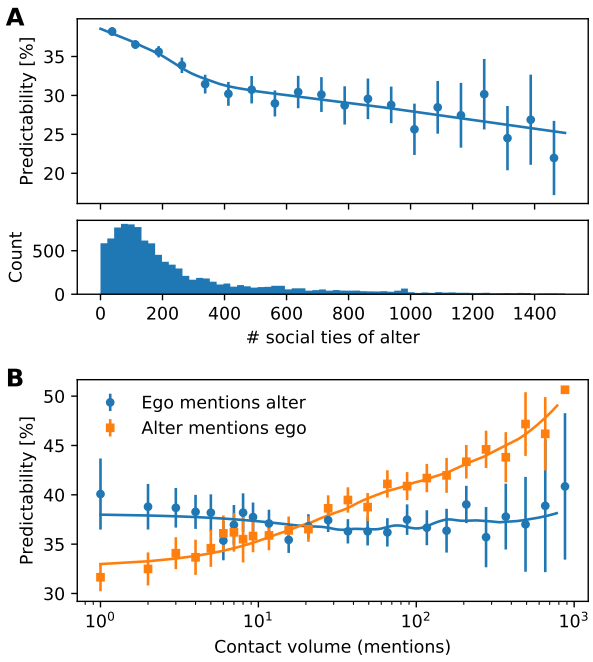

Figure 3: Social interactions are visible in information flow. (**A**) Alters with more social ties of their own provided less information about the ego than less popular alters. (**B**) Information flow captures directionality in relationships, a key factor in social dynamics. Alters who often contact the ego provide more predictive information about the ego than alters who rarely mention the ego. Yet, if the ego frequently mentions the alter, it does not necessarily mean the alter will provide more predictive information about the ego.

of our data lies, but the trend continued beyond this as well. This decreasing trend belies the power of hubs in many ways: while hubs strongly connect a social network topologically [24], limited time and divided attention across their social ties bounds the alter's ability to participate in information dynamics mediated by the social network and this is reflected in the predictability.

Reciprocated contact is an important indicator of social relationships [25], especially in online social activity where so much communication is potentially one-sided [23]. In Fig. 3B we investigated how directionality in contact volume, how often the ego mentions the alter and vice versa, related to information flow. We found that the ego was more predictable given the alter for those dyads where the alter more frequently contacted the ego, but there was little change across dyads when the ego mentioned the alter more or less frequently (Fig. 3B). We also observed a similar trend for information flow but in reverse, when predicting the alter given the ego (SM). These trends captured the reciprocity of information flow: an alter frequently contacting an ego will tend to give predictive information about the ego, but the converse is not

true: an ego can frequently contact her alter but that does not necessarily mean that the alter will be any more predictive, as evidenced by the flat trend in Fig. 3B.

In summary, the ability to repeatedly and accurately predict the text of individuals provides considerable value to the providers of social media, allowing them to develop profiles to identify and track individuals [26, 27] and even manipulate information exposure [28]. That information is so strongly embedded socially underscores the power of the social network: by knowing who are the social ties of an individual and what are the activities of those ties, our results show that one can in principle accurately profile even those individuals who are not present in the data [5].

The time-ordered cross-entropy (Fig. 1A) applied here to online social activity is a natural, principled information-theoretic measure that incorporates all the available textual and temporal information. While weaker than full causal entailment, by incorporating time ordering we identify social information flow as the presence of useful, predictive information in the past of one's social tie beyond that of the information in one's own past. Doing so closely connects this measure with Granger causality and other strong approaches to information flow [29, 20].

Acknowledgments We gratefully acknowledge the resources provided by the Vermont Advanced Computing Core. This material is based upon work supported by the National Science Foundation under Grant No. IIS-1447634. LM acknowledges support from the Data To Decisions Cooperative Research Centre (D2D CRC), and the ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS).

References

[1] C. Shirky. The political power of social media: Technology, the public sphere, and political change. *Foreign affairs*, pages 28–41, 2011. 1

[2] G. Lotan, E. Graeff, M. Ananny, D. Gaffney, I. Pearce, et al. The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International journal of communication*, 5:31, 2011. 1

[3] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, page 201517441, 2016. 1

[4] D. Lazer, A. S. Pentland, L. Adamic, S. Aral, A. L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al. Computational social science. *Science*, 323(5915):721, 2009. 1

[5] D. Garcia. Leaking privacy and shadow profiles in online social networks. *Science Advances*, 3(8), 2017. 1, 9

[6] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. In *Proceedings of the 13th international conference on World Wide Web*, pages 491–501. ACM, 2004. 1

[7] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, pages 519–528. ACM, 2012. 1

[8] E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, 2015. 1

[9] S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009. 1

[10] D. Centola. The spread of behavior in an online social network experiment. *Science*, 329(5996):1194–1197, 2010. 1

[11] S. Aral and D. Walker. Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341, 2012. 1

[12] G. Ver Steeg and A. Galstyan. Information transfer in social media. In *Proceedings of the 21st international conference on World Wide Web*, pages 509–518. ACM, 2012. 1

[13] J. Borge-Holthoefer, N. Perra, B. Gonçalves, S. González-Bailón, A. Arenas, Y. Moreno, and A. Vespignani. The dynamics of information-driven coordination phenomena: A transfer entropy analysis. *Science Advances*, 2(4):e1501158, 2016. 1

[14] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012. 2

[15] C. E. Shannon. Prediction and entropy of printed english. *Bell Labs Technical Journal*, 30(1):50–64, 1951. 2

[16] P. F. Brown, V. J. D. Pietra, R. L. Mercer, S. A. D. Pietra, and J. C. Lai. An estimate of an upper bound for the entropy of english. *Computational Linguistics*, 18(1):31–40, 1992. 2

[17] T. Schürmann and P. Grassberger. Entropy estimation of symbol sequences. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 6(3):414–427, 1996. 2

[18] I. Kontoyiannis, P. Algoet, Y. M. Suhov, and A. Wyner. Nonparametric entropy estimation for stationary processes and random fields, with applications to english text. *IEEE Transactions on Information Theory*, 44(3):1319–1327, May 1998. 00104. 2

[19] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010. 2

[20] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461, 2000. 4, 9

[21] M. Staniek and K. Lehnertz. Symbolic transfer entropy. *Physical Review Letters*, 100(15):158101, 2008. 4

[22] R. I. Dunbar. Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences*, 16(4):681–694, 1993. 5

[23] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM, 2010. 5, 6, 8

[24] R. Albert, H. Jeong, and A.-L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000. 8

[25] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*. Cambridge university press, 1994. 8

[26] Y.-A. De Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3:1376, 2013. 9

[27] Y.-A. De Montjoye, L. Radaelli, V. K. Singh, et al. Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, 347(6221):536–539, 2015. 9

[28] E. Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011. 9

[29] C. W. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969. 9

Supporting Material for "Information flow reveals prediction limits in online social activity"

James P. Bagrow[1,2,*], Xipei Liu[1,2], and Lewis Mitchell[1,2,3]

[1]Department of Mathematics & Statistics, University of Vermont, Burlington, VT, United States
[2]Vermont Complex Systems Center, University of Vermont, Burlington, VT, United States
[3]School of Mathematical Sciences, North Terrace Campus, The University of Adelaide, SA 5005, Australia
[*]Corresponding author. Email: james.bagrow@uvm.edu, Homepage: bagrow.com

August 15, 2017

Contents

List of figures

List of tables

S1 Dataset details

We selected a random sample of individuals for study from the Twitter 10% Gardenhose feed collected during the first week of April 2014. From this we uniformly sampled individuals who had tweeted in English (as reported by Twitter in the metadata for each tweet) during this time period and had 50–500 followers, as reported in the feed metadata. The lower follower cutoff is to avoid inactive and bot accounts while the higher cutoff is to ensure that individuals in our sample have comparably-sized egonetworks and to avoid studying unusually popular outlier accounts such as celebrity accounts. For each user we then collected their complete public tweet history excluding retweets (up to 3200 most recent public messages, as allowed by Twitter's Public REST API limit [1]). As discussed below, we then applied to these users a filtering procedure including both computational tools and human raters to help ensure sufficient data on individual activities and to limit bots and non-individual accounts from our sample (Sec. S1.1). When finished, we retained a final sample of $n = 927$ individual egos and their top-15 alters ($n = 13,905$ total users).

For each initially sampled ego, we collected the user IDs of the account whom the ego "at-mentioned" most frequently in their public tweets, forming the ***rank-1 alter***. Mentions are a stronger signal than simply Twitter following, as it demonstrates active communication on the behalf of at least one of the individuals of a social tie. As was done with the egos, the REST API was then used to retrieve the complete public tweet history of this alter. Examining the messages of the (ego, rank-1 alter) dyad, we retained egos where the ego's tweets covered at least a one-year period, the alter's tweets covered at least a one year period, and the ego at-mentioned at least 15 unique Twitter users (including the rank-1 alter).

For dyads who satisfied these criteria we collected the full public messages of the remaining 14 most at-mentioned alters, giving us the full public activities of the ego and his or her top-15 most mentioned alters.

S1.1 Filtering and rating procedure

To limit the effects of bots and non-personal accounts, we moved beyond the basic filtering criteria listed above and employed both computational tools and human raters to examine the accounts of the egos in our dataset. These tools were applied in April 2017. A small number of accounts in our sample were suspended or deleted after our data collection period and were not available to be examined, so we simply retained these unrated accounts in our sample. We used the ***botometer***[1] API [2, 3, 4, 5] to score the probability that an ego account was a bot, and eliminated $n = 46$ accounts that scored above 50%. This tool examines

[1] https://botometer.iuni.iu.edu/

Twitter accounts along a number of dimensions to estimate the likelihood the account belongs to a bot. Next, we asked human raters to examine the accounts and report whether the account appeared to belong to a real person or a non-personal entity such as a corporation or a bot. Two independent raters examined each account's Twitter homepage if available. We removed $n = 84$ accounts where both raters agreed the account did not belong to an individual, beyond those already flagged by the botometer scores. Raters were recruited on Amazon Mechanical Turk and compensated at a rate of $0.10 per three Twitter accounts. Lastly, we also removed a small number of accounts ($n = 31$) showing convergence issues with our entropy estimators, as inferred by negative KL-divergences from the ego to the alter or vice versa. (See also our full convergence analysis in Sec. S3.4.) This gave our final sample size of $n = 927$ egos and their top-15 associated alters.

S1.2 Text processing

To apply the entropy estimators discussed in Sec. S3, we first need to process and tokenize the texts of users. The UTF-8 encoded text of each user was processed by removing casing, punctuation (except for twitter specific "@" and "#" symbols), and URLs (identified as words beginning with "http://" or "https://"). All tweet texts were concatenated into a single text string in time order (based on the tweet timestamps), except for "retweets" which were all excluded in order to focus on the effect of shared language and avoid artificially inflating predictability scores. The text was then tokenized into words by segmenting on whitespace.

S2 Control procedures

We performed two controls for the cross-entropy experiments: random tweets or "temporal control" and random alters or "social control". For the temporal control we constructed proxy tweet streams for the alters which preserved the approximate times at which alters had written messages. To do this we substituted for each real alter tweet a randomly sampled English-language tweet posted during the same hour as the real alter tweet. The randomly sampled replacement tweets were taken from the 10% Gardenhose feed.

In the social control we randomized the ego networks, swapping the tweet text streams of true alters with those of randomly chosen alters. This control does not preserve the times at which the original alters had authored tweets, hence the use of the previous temporal control.

S3 Estimators for entropy, cross-entropy, and cumulative cross-entropy

S3.1 Entropy (with and without correlations)

The entropy (rate) h of a sequence of words is the number of bits needed to encode the next word, on average, given past words. Kontoyianni et al. [6] proved convergence for a nonparametric estimator $\hat{h}$ for h:

$$\hat{h} = \frac{N \log N}{\sum_{i=1}^{N} \Lambda_i}, \tag{1}$$

where N is the length of the sequence of words and Λ_i is the match length of the prefix at position i, i.e., it is the length of the shortest subsequence (of words) starting at i that has not previously appeared. (All logarithms are base 2.) If the sequence of words were randomly shuffled, breaking any long-range structure, this estimator converges to the traditional Shannon entropy on unigrams (Fig. S1).

The ideas underlying estimators such as Eq. (1) play an important role in the mathematics of data compression algorithms. Indeed, some authors have used practical compression software to estimate the

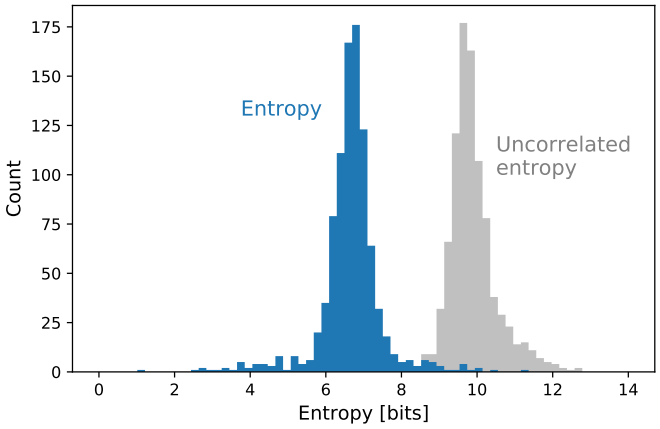

Figure S1: Correlations in the text account for ≈3 additional bits of information. The uncorrelated entropy (considering only the relative frequencies of words posted by Twitter users) is approximately 3 bits higher than the correlated entropy as estimated from Eq. 1.

information content of a text. However, such estimates tend to be biased, as specific compression implementations (such as gzip) tend to sacrifice small amounts of extra compression in order to run much more efficiently. Due to these approximations it is important to work directly with the theoretical estimator to more accurately estimate h, as we have when we applied Eq. (1).

S3.2 Cross-entropy

To generalize Eq. (1) to a cross-entropy between two sequences A and B, define the **cross-parsed match length** $\Lambda_i(A|B)$ as the length of the shortest subsequence starting at position i of sequence A not previously seen in sequence B. If sequences A and B are ***time-aligned***, as in timestamped social media posts, then 'previously' refers to all the words of B written prior to $t_i(A)$, the time when the ith word of A was posted, according to the timestamp of the respective tweet. The estimator for the cross-entropy rate is then

$$\hat{h}_\times(A \mid B) = \frac{N_A \log N_B}{\sum_{i=1}^{N_A} \Lambda_i(A \mid B)}, \tag{2}$$

where N_A and N_B are the lengths of A and B, respectively. An estimator of the relative entropy (or KL-divergence) similar to Eq. (2) was introduced by Ziv and Merhav [7]. The log term in Eq. (2) has changed to $\log N_B$ because now B is the "database" (or window, in Lempel-Ziv terms) we are searching over when we compute the match lengths; the N_A factor is due to the average of the Λ_i's taking place over A. The cross-entropy tells us how many bits on average we need to encode the next word of A given the information previously seen in B. Further, $\hat{h}_\times(A \mid A) = \hat{h}$. The cross-entropy can be applied directly to an ego-alter pair by choosing B to be the text stream of the alter and A the text stream of the ego.

The cross-entropy can also be associated with the predictability by applying Fano's Inequality [8]. Fano's Inequality relies on both the entropy and the cardinality of the random variable; here we take the size of the ego's unique vocabulary as this is the variable we are trying to predict. In Fig. S2 we present the relationship between cross-entropy and predictability for our data compared with solid lines denoting constant vocabulary-size curves. The predictability of the ego given the alter is lower than the predictability of the ego given the ego because the cross-entropy is greater than the entropy, capturing the increased uncertainty (decreased information) we have by trying to predict the ego given the alter instead of the ego.

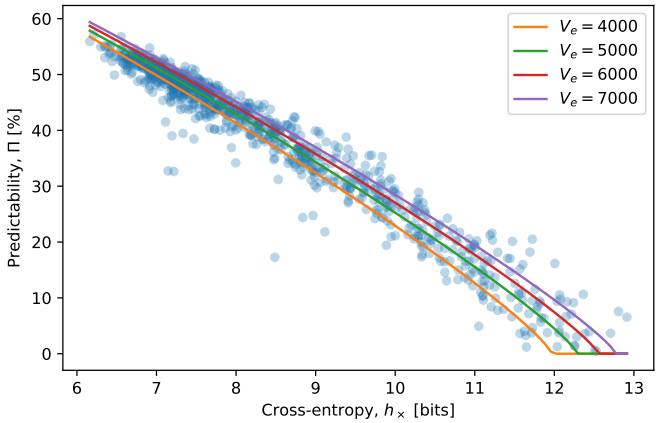

Figure S2: Cross-entropy $\hat{h}_\times$(ego | alter) and predictability Π across different ego vocabulary sizes.

S3.3 Cumulative cross-entropy

We now wish to generalize the cross-entropy to $\hat{h}_\times(A \mid \mathcal{B})$, estimating the average amount of information needed to encode the next word of sequence A given the information in a *set* of sequences $\mathcal{B}$. The cross-parsed match length for a set of databases is $\Lambda_i(A \mid \mathcal{B}) = \max\{\Lambda_i(A \mid B), B \in \mathcal{B}\}$, i.e. we take the longest match length over any of the sequences in $\mathcal{B}$. This cross-parsing implies a new $\log N_{A\mathcal{B}}$ factor in the estimator, where $N_{A\mathcal{B}}$ is the average of the lengths N_B ($B \in \mathcal{B}$), weighted by the number of times matches were found in each sequence $B \in \mathcal{B}$. (If the same match length occurs for more than one sequence $B \in \mathcal{B}$ then each such sequence receives a weight in the average.) The estimator is

$$\hat{h}_\times(A \mid \mathcal{B}) = \frac{N_A \log N_{A\mathcal{B}}}{\sum_{i=1}^{N_A} \Lambda_i(A \mid \mathcal{B})}, \tag{3}$$

where $N_{A\mathcal{B}} = \sum_{B\in\mathcal{B}} w_B N_B / \sum_{B\in\mathcal{B}} w_B$ and w_B is the number of times that matches from A are found in $B \in \mathcal{B}$. Note that $\sum_B w_B \geq N_A$ due to possible ties, with equality holding if no ties occur. Note that Eq. (3) reduces to Eq. (2) when $|\mathcal{B}| = 1$.

Equation (3) lets us build the cumulative cross-entropy by appropriate choices of $\mathcal{B}$. In the main text we sequentially added alters to the set $\mathcal{B}$ in order of decreasing contact volume (i.e., $\mathcal{B} = \{\text{alters}\}$), to understand how information grows as more alters are made available. Likewise, Eq. (3) lets us build the transfer entropy-like measures[2] discussed in the main text by additionally including the ego within the set $\mathcal{B}$ (i.e., $\mathcal{B} = \{\text{ego}\} \cup \{\text{alters}\}$).

S3.4 Estimator convergence on our data

The estimator given by Eq. (1) has been proven to converge asymptotically under stationarity assumptions [6]. However, our data are finite, and so we investigated the convergence properties of the estimator empirically (Fig. S3). In general, we observed that the entropy saturates after around 1000 tweets (approximately 10,000 words).

For the cross-entropy estimator $h_\times(A \mid B)$, we examined the convergence over the lifespan or time window within which the ego has authored tweets. Figure S4 shows the convergence of the cross-entropy for the

[2]Note that this is not truly a transfer entropy because transfer entropy is defined based on the conditional entropy [9] whereas here we still construct a cross-entropy.

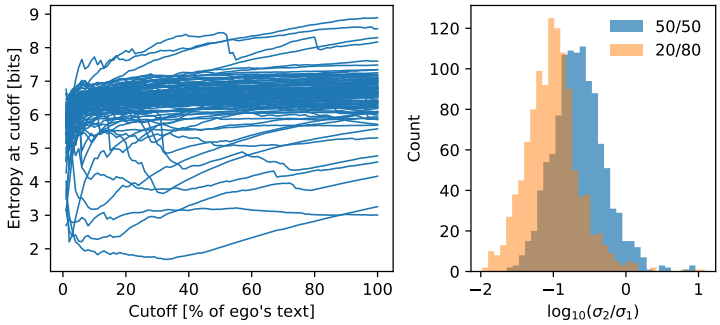

Figure S3: Convergence of the entropy estimator. (**left**) The estimator of egos generally saturates well within our data window, as evidenced by the flattening of the entropy estimate as we examine more of the ego's text. (**right**) Here we compute the variance of each ego's entropy over two portions of the curves shown to the left. In one distribution we compare the variance of the final 50% of the data to the initial 50%, while in the other distribution we compare the variance of the final 20% of the data to the initial 80%. The latter shows a significantly smaller variability than the former, underscoring how much the entropy estimates have settled down by the end of our data window. In the left plot we show a random selection of egos, while the distributions on the right cover all dyads in our dataset.

rank-1 alter $h_\times$(ego|alter 1) (left panel), where we truncate both the ego and alter's tweets after some fraction of the ego's lifespan. In general, we found that the cross-entropy estimator saturates within around 50% of the ego's lifespan. The right panel shows a histogram of the slopes of the convergence curves for all users over the final 25% of the ego's lifespan, as a fraction of the final value of $H_\times$(ego|alter 1). These slopes were computed via linear regressions, and many of the slopes are close to zero.

S4 Extrapolating cross-entropy and predictability

We are limited by our data to a window of the top-15 most frequently contacted alters per ego. To address a limit of entropy or predictability as more alters are added, we used a saturating function to extrapolate beyond alter rank $r = 15$.

Specifically, we extrapolated the cross-entropy using the function

$$h(r) = h_\infty + \frac{\beta_0}{\beta_1 + r}, \tag{4}$$

with the goal of identifying the value of h_∞ and, perhaps more realistically, to estimate $h(r_{\text{dunbar}})$, where $r_{\text{dunbar}} \approx 150$ [10]. Using Levenberg-Marquardt for nonlinear regression, we found best fit parameters of (value ± 95% CI):

$$\begin{aligned} h_\infty &= 5.761978 \pm 0.089699, \\ \beta_0 &= 9.455984 \pm 1.358027, \\ \beta_1 &= 2.553345 \pm 0.444479, \end{aligned}$$

for the cross-entropy of the ego given the alters.

In Fig. S5 we show the mean cross-entropy as a function of alter rank and compare it with the results of the fitted function. The fit is reasonable. Similarly, fits of the same functional form were applied to the

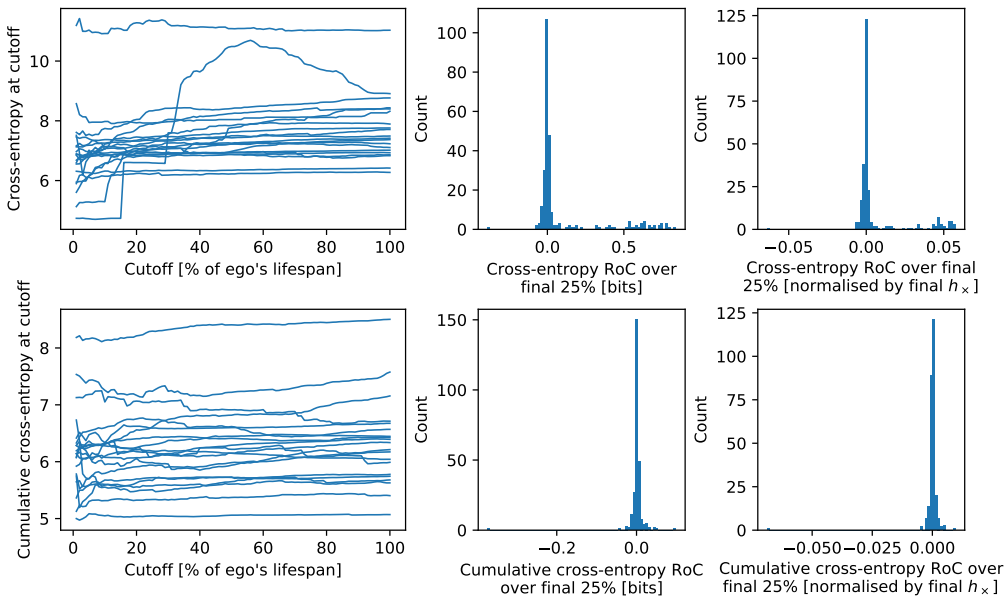

Figure S4: Convergence of the cross-entropy estimator. (**left**) The estimator saturates well within the lifespan of the ego's tweets, generally within 50% of the lifespan. (**right**) The distributions of the slope (RoC: rate-of-change) over the final 25% of the curves. The majority of egos have very flat RoCs at the end of their data windows. In the left plots we show a randomly selection of egos, while the distributions on the right curve all dyads in our dataset.

predictability (ego given alter) curves:

$$\Pi(r) = \Pi_\infty + \frac{\beta_0}{\beta_1 + r}, \tag{5}$$

and here we found best fit parameters

$$\begin{aligned} \Pi_\infty &= 0.608219 \pm 0.006914, \\ \beta_0 &= -0.734410 \pm 0.100195, \\ \beta_1 &= 2.320039 \pm 0.398486. \end{aligned}$$

We also experimented with a second form of extrapolating function:

$$h(r) = h_\infty + \beta_0 r^{-\beta_1}, \qquad \Pi(r) = \Pi_\infty + \beta_0 r^{-\beta_1}. \tag{6}$$

This function, referred to as **Function 2**, also fits the data well (Fig. S6) but is a bit less conservative in its extrapolation prediction when extrapolating for $r \to \infty$. To further compare Function 2 and the original function (Function 1), we plotted the residuals between the fits and the data in Fig. S6.

Taken together, we see that Function 1 (Eqs. (4) and (5)), the more conservative estimator, has consistently smaller residuals than Function 2. Both functions' residuals were statistically independent of the exogenous variable r ($p > 0.05$). We concluded that Function 1 is a better choice since it has smaller residuals and is more conservative than Function 2.

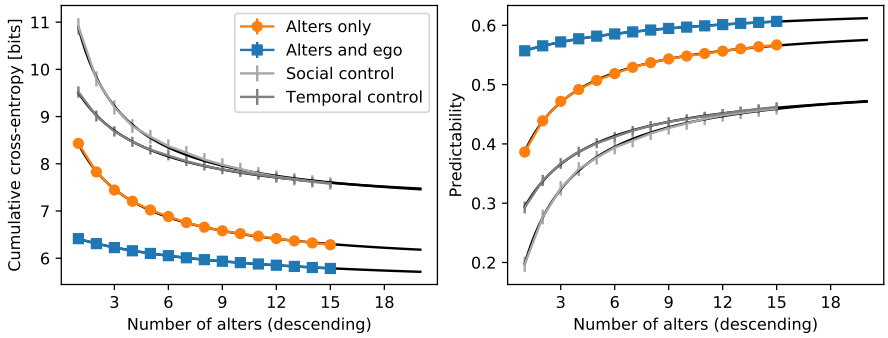

Figure S5: Extrapolating cross-entropy and predictability. The fitted functions (Eqs. (4) and (5), solid lines) compared with the original cross-entropy data (averaged for each alter rank). Note that the function was fitted to the original and not averaged cross-entropy values.

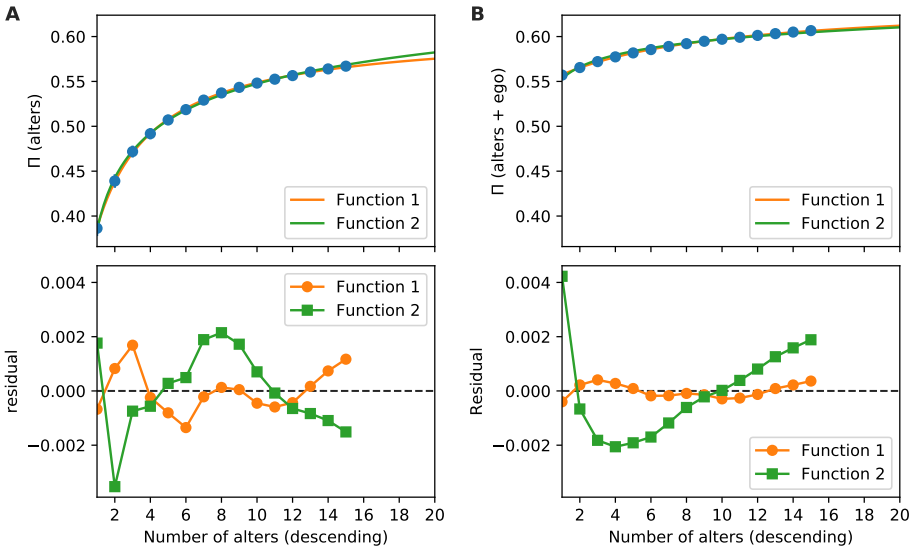

Figure S6: Extrapolations and residuals for the predictability functions (Function 1: Eq. (5); Function 2: Eq. (6)). (**A**) Comparison of the measured cross-entropies (for the top-15 alters) with the extrapolation functions and mean residuals between function fit and original data. (**B**) Same as panel A but for fits of the same form as Eqs. (5) and (6) to $\Pi(r)$ including the past of the ego along with the alters. Overall, Function 1 was slightly more conservative than Function 2, extrapolating to a slightly smaller value of Π, and had lower residuals.

S5 Supporting results

S5.1 Vocabulary sizes on social media

In Fig. S7 we present the distributions of the total number of words written per ego and the number of unique words (the vocabulary size) per ego, for the users in our dataset. Egos had an average of ≈26800[3] total numbers of words and ≈5200 numbers of unique words. The latter quantity, the vocabulary size, was used

[3]For context, this is about the typical length of a novella, defined by the Science Fiction and Fantasy Writers of America as 17500–39999 words [11].

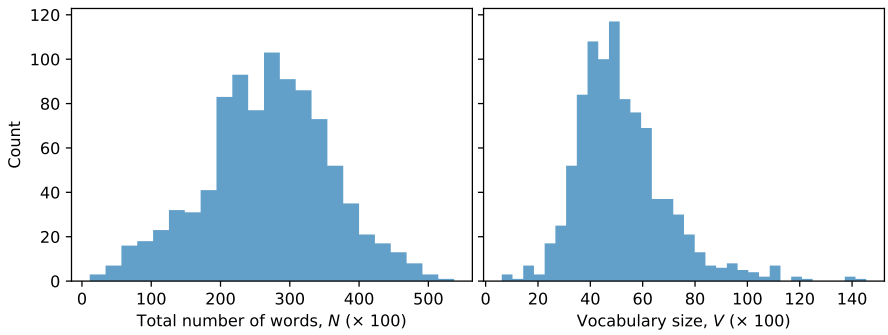

Figure S7: Distributions of Twitter ego vocabulary size: the total number of words written (left) and the vocabulary size or number of unique words written (right).

Text	Author	$\hat{h}$ (sample 1) [bits]	$\hat{h}$ (sample 2) [bits]
For Whom the Bell Tolls	Ernest Hemingway	5.870953	5.910003
Gravity's Rainbow	Thomas Pynchon	5.881336	5.881336
The Fellowship of the Ring	J.R.R. Tolkien	6.439215	6.340354
Ulysses	James Joyce	7.067339	7.227677

Table S1: Entropy rates of some example texts. Samples 1 and 2 were the first and second 38,000 words of each text, respectively (a bit longer than the typical Twitter user's text stream). Hemingway is known for his simple writing style while Joyce is famous for the opposite; this is well reflected in their respective entropy rates.

in Fano's Inequality to compute the predictability.

S5.2 Information content on social media compared with formal written text

To contextualize the entropy rates we estimated for our dataset (most egos had entropies of $5.5 < h < 8$ bits), we compared the entropy rates of formal text with the rates of Twitter users to better understand the information content of social media writings[4]. First, we considered the entropies of some famous example texts (Table S1). We considered writers who were famous for being very simple in style (Hemingway) and very complex (Joyce) and found this was reflected in the entropy rates (5.87 bits for Hemingway compared with 7.06 bits for Joyce). The higher entropies reflect that Joyce's word choices are less regular and less predictable than Hemingway's. These formally written and edited texts are very different from social media posts, and yet the range of entropies values we observed was compatible to some extent.

We also took the standard ***Brown corpus*** [12], a benchmark text set used in natural language processing and computational linguistics research, as a large-scale baseline of formal text. The corpus consists of approximately 1M words and covers 500 writing samples across 15 fiction and non-fiction categories. Each category was broken into 10-thousand-word chunks and the entropies of these chunks were computed. Individual chunks did not span multiple categories and if a chunk at the end of a category was less than 10 thousand words it was discarded to ensure all entropy estimates were computed using the same amount of data. This gave $n = 93$ samples.

[4]The texts were processed by removing punctuation and casing.

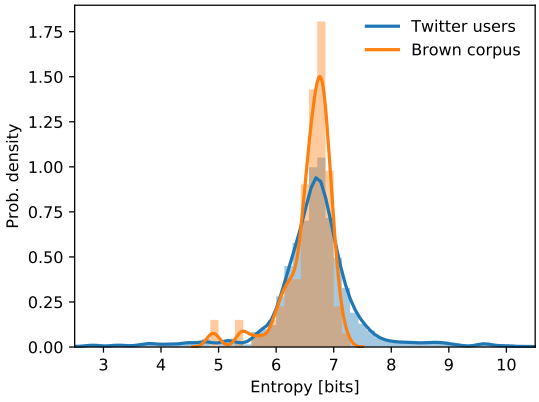

Figure S8: Entropy distributions for social and formal written text corpora. We found that the distributions have the same central tendency (Mann-Whitney U test: $U = 39797$, $p > 0.1$) but different dispersions (Fligner-Killeen test on variances, $\chi^2 = 15.580$, $p < 10^{-4}$) Brown corpus was taken from NLTK v3.2.1 corpora [13].

We found that formal and social text have the same average value but that the variation across the Twitter sample was significantly greater than across the formal text (Fig. S8).

S5.3 A censoring filter to determine long-range information in the egos and alters

To study the recency of information we applied the (cross-)entropy estimators to censored text, where we removed the recent past of the text and asked how much if any information is lost. If most predictive information is in the recent past, by removing it we should see a significant change in the cross-entropy, although there should always be some loss in information, as the sequences being matched across are always getting shorter.

Specifically, to compute the original cross-entropy (Eq. (2)) between two sequences A and B, we need the cross-parsed match length at position i, $\Lambda_i(A \mid B)$, giving us the shortest subsequence of words in A beginning at position i not seen previously in B. This last part, the past of B, is based on the timestamps of the words: we search all words in B written before the ith word w_i in A: $[w_j \in B \mid t_j(B) < t_i(A)]$, where $t_i(A)$ is the time when the ith word in A was posted (taking all words in a single tweet to be posted at the time the tweet was posted).

The censoring filter simply truncates the past of B at each position i. Instead of searching all of the past of B we instead search the past older than an amount ΔT: $[w_j \in B \mid t_j(B) < t_i(A) - \Delta T]$. By censoring B as we sweep forward in the computation of the cross-entropy, we can estimate how much information is recent versus long-term on average by the change in the cross-entropy rate as a function of ΔT. The same calculation holds for the "self" entropy, simply by setting $B = A$.

We measured the loss of information in the main text out to 24 hours. Here we complement that calculation with Fig. S9 which presents the information loss out to 1 week. We see in both curves that long-range information is lost by the increasing trend. However, the trend is more shallow for the alters than the ego: taking away more of the ego's past removes more information about the ego than taking away the less recent pasts of the alters.

S5.4 Posting frequency and predictability

Here we present in Figs. S10 and S11 the association between the posting frequencies of the egos and alters with the predictability of the ego (Fig. S10: top row), predictability of the alter (Fig. S10: bottom row), predictability of the ego given the alter (Fig. S11: top row), and the predictability of the alter given the ego (Fig. S10: bottom row). We found that the predictabilities of the egos and alters are roughly independent of

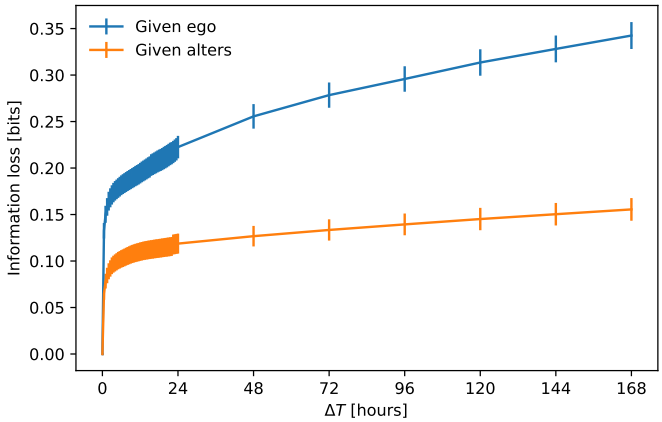

Figure S9: Alters provided less long-range information about the ego than the ego itself. This plot complements the loss in predictability shown in the main text and extends ΔT beyond the 24-hour window to a one-week period.

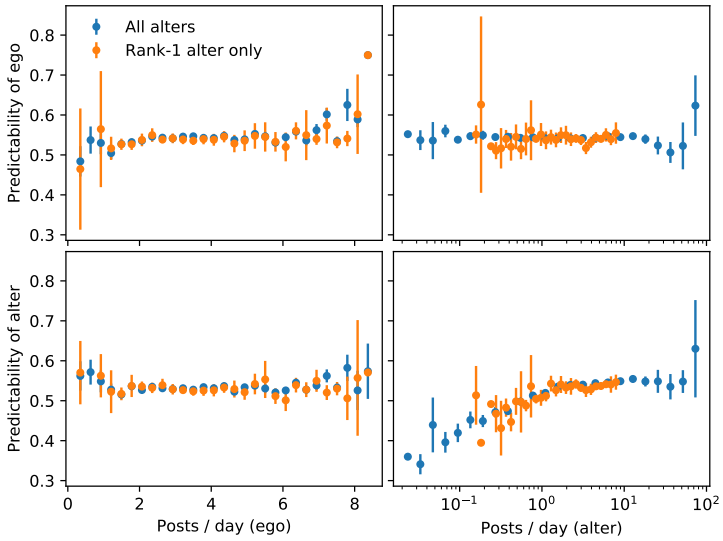

Figure S10: Self-predictabilities are approximately independent of activity frequency, with the exception of predictability of the alter as a function of the alter's activity frequency (lower right). This is primarily due to insufficient data: alters who post very infrequently have low predictability, but the trend levels off for alters who post more than ≈1 time per day.

their posting frequency, except for very infrequently posting alters (Fig. S11: bottom right), which is likely a result of insufficient data.

The associations between posting frequency and the cross-predictability of the ego given the alter hold even when considering all alters not just the rank-1 alter, as we did in the main text (Fig. S11: top row). Likewise, the trends also hold (in reverse) when considering the predictability of the alter given the ego (Fig. S11: bottom row).

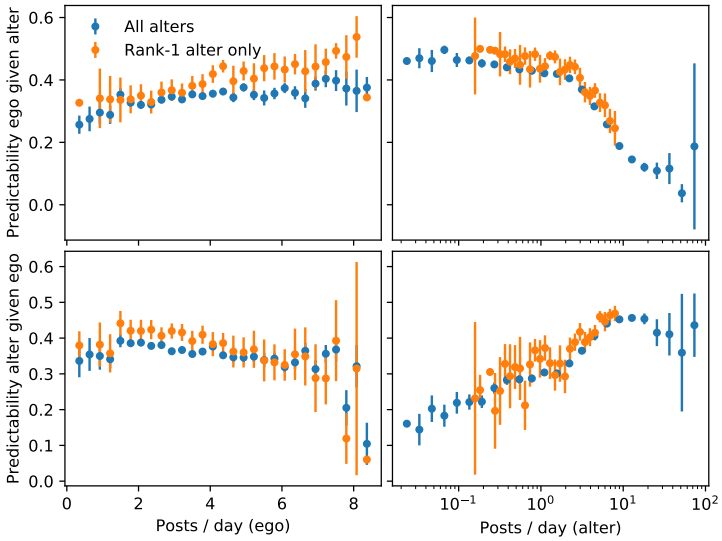

Figure S11: Association between cross-predictability and posting frequency holds for all alters. Here we repeated the trends shown in the main text where we considered the rank-1 alter only, but now we included all alters as well. Due to alters who post very frequently and very infrequently, we used a logarithmic scale on the right column.

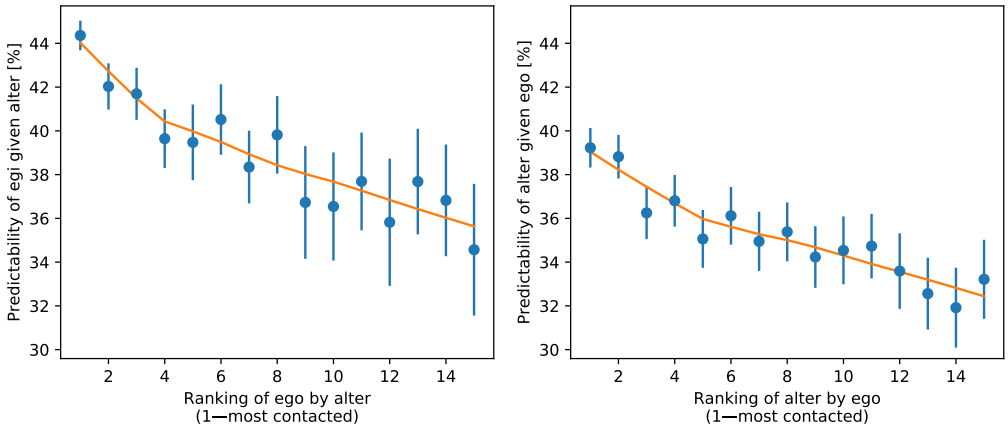

Figure S12: Less frequently contacted ties provide less predictive information

S5.5 Contact volumes and predictability

Here we present in Fig. S12 the predictability across social ties as a function of how often those social ties contact one another. We ranked the ties of individuals in descending order. Working with ranks helps to account for the variability in contact volumes and overall activity levels across users of social media. Across ranks we found a significant decrease in predictive information, in both directions (predicting the ego given the alter and predicting the alter given the ego).

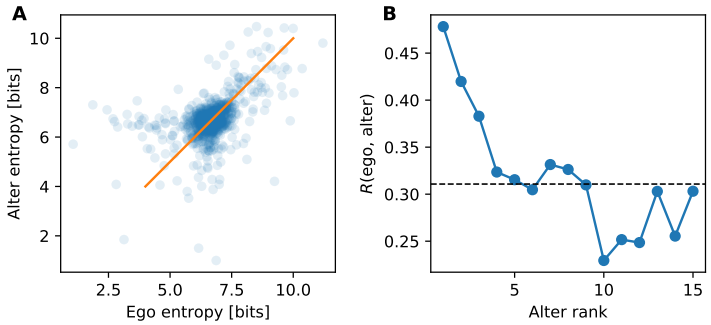

Figure S13: An "information homophily" between egos and alters. The entropies of egos and their alters are strongly correlated, indicating a ***homophily effect***. (**A**) The entropy rate h of egos compared with their rank-1 (most frequently contacted) alter. The straight line $y = x$ provides a guide for the eye. (**B**) The (Pearson) correlation coefficient R between ego and alter entropies as a function of alter rank. Correlation decreases with rank. The dashed line indicates the correlation coefficient over all ranks.

S5.6 Information homophily

We found a homophily effect between egos and alters in terms of their (self) information. The entropy rates of the ego and alter on a given dyad were correlated (Fig. S13A). Figure S13A covers the correlation between the ego and the rank-1 alter. In Fig. S13B we plot the correlation coefficient between $\hat{h}(\text{ego})$ and $\hat{h}(\text{alter})$ as a function of alter rank. The correlation drops consistently over the first five or so alters, implying that the homophily effect is a connected with contact volume. This information homophily is worth exploration in further work.

On the other hand, the cross-entropies between the egos and alters are less well correlated, either with the cross-entropy in the opposite direction, or with the entropies themselves. The correlations (for the rank-1 alters) are:

$$
\begin{aligned}
R\left(\hat{h}(\text{ego}), \hat{h}(\text{alter})\right) &= 0.478, \\
R\left(\hat{h}_{\times}(\text{ego} \mid \text{alter}), \hat{h}_{\times}(\text{alter} \mid \text{ego})\right) &= -0.122, \\
R\left(\hat{h}(\text{ego}), \hat{h}_{\times}(\text{ego} \mid \text{alter})\right) &= 0.240, \\
R\left(\hat{h}(\text{ego}), \hat{h}_{\times}(\text{alter} \mid \text{ego})\right) &= 0.227, \\
R\left(\hat{h}(\text{alter}), \hat{h}_{\times}(\text{ego} \mid \text{alter})\right) &= 0.247, \\
R\left(\hat{h}(\text{alter}), \hat{h}_{\times}(\text{alter} \mid \text{ego})\right) &= 0.300.
\end{aligned}
$$

While significant in all cases, the correlations between the (self) entropies $\hat{h}(\text{ego})$ and $\hat{h}(\text{alter})$ are stronger than between any of the cross-entropies, demonstrating that the effects captured by the cross-entropies over a dyad are different than that captured by the entropies of the individuals in that dyad.

S5.7 Reciprocity and information flow

In the main text we reported on the relationship between contact volume and information flow, as measured by the cross-entropy. A closely related quantity often employed in this context is the Kullback-Leibler divergence, or KL-divergence, $KL(\text{ego} \parallel \text{alter}) \equiv \hat{h}_{\times}(\text{ego} \mid \text{alter}) - \hat{h}(\text{ego})$ [8]. In our data the correlation

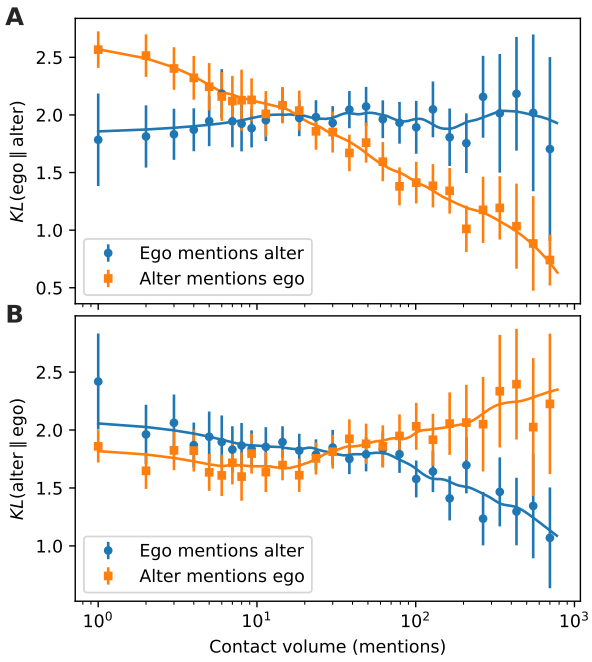

Figure S14: Reciprocated information flows are captured in both directions. (**A**) In the main text we reported on the trend between contact volume and the cross-entropy from the alter to the ego. We repeat that figure here but with the KL-divergence. (**B**) In comparison, if we consider the opposite divergence, from the ego to the alter, we see a similar trend but reversed: egos which more frequently mention their alter give more predictive information (lower divergence) than egos which mention their alter less often.

between $\hat{h}_\times(\text{ego} \mid \text{alter})$ and $KL(\text{ego} \parallel \text{alter})$ is quite high (Table S2) and so they are effectively the same measure.

We showed that alters who more frequently mention their ego provide more predictive information (lower cross-entropy/KL-divergence) than alters who less frequently mention their ego. Meanwhile, the converse was not true: the ego can mention the alter more or less, and there was not an association with the predictive information possessed by the alter about the ego.

Here we supplement that result by reversing the perspective—instead of asking about the predictive information about the ego possessed by the alter we ask about the predictive information about the alter possessed by the ego. We measure this with the reversed KL-divergence, $KL(\text{alter} \parallel \text{ego}) \equiv \hat{h}_\times(\text{alter} \mid \text{ego}) - \hat{h}(\text{alter})$. With this reversal we should expect to also see a reversal in the association of contact volume, and we found this to be the case (Fig. S14). In Fig. S14 we compared both KL-divergences and saw that the trends approximately reverse, as expected.

S5.8 Interrelations between information-theoretic quantities

In Table S2 we present the Spearman rank correlation coefficients between the primary information-theoretic quantities we computed, including the KL-divergence: $KL(\text{ego} \parallel \text{ater}) \equiv \hat{h}_\times(\text{ego} \mid \text{alter}) - \hat{h}(\text{ego})$. The cross-entropy and KL-divergence are strongly correlated.

References

[1] Twitter REST APIs. Available from: `https://dev.twitter.com/rest/public`, 2016. Accessed: 2016-07-07. S2

[2] O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini. Online human-bot interactions: Detection, estimation, and characterization. In *ICWSM*, 2017. S2

	$\hat{h}$(ego)	$\hat{h}$(alter)	$\hat{h}_\times$(e | a)	$\hat{h}_\times$(a | e)	KL(e || a)	KL(a || e)
$\hat{h}$(ego)	1.000000	0.439867	0.302718	0.201535	-0.142122	0.000338
$\hat{h}$(alter)	0.439867	1.000000	0.257699	0.236107	0.048474	-0.176895
$\hat{h}_\times$(e | a)	0.302718	0.257699	1.000000	-0.281961	**0.858274**	-0.395016
$\hat{h}_\times$(a | e)	0.201535	0.236107	-0.281961	1.000000	-0.384678	**0.881891**
KL(e || a)	-0.142122	0.048474	**0.858274**	-0.384678	1.000000	-0.409757
KL(a || e)	0.000338	-0.176895	-0.395016	**0.881891**	-0.409757	1.000000

Table S2: Cross-entropy and KL-divergence are strongly correlated (bold). Here we present the nonparametric Spearman correlation between information-theoretic quantities computed over the $n = 927$ ego-(rank-1 alter) dyads.

[3] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer. BotOrNot: A system to evaluate social bots. In *WWW Developers Day*, 2016. S2

[4] E. Ferrara, O. Varol, C. A. Davis, F. Menczer, and A. Flammini. The rise of social bots. *Communications of the ACM*, 59(7), 2016. S2

[5] V. S. Subrahmanian, A. Azaria, S. Durst, V. Kagan, A. Galstyan, K. Lerman, L. Zhu, E. Ferrara, A. Flammini, and F. Menczer. The DARPA Twitter Bot Challenge. *Computer*, 49(6):38–46, 2016. S2

[6] I. Kontoyiannis, P. Algoet, Y. M. Suhov, and A. Wyner. Nonparametric entropy estimation for stationary processes and random fields, with applications to english text. *IEEE Transactions on Information Theory*, 44(3):1319–1327, May 1998. 00104. S3, S5

[7] J. Ziv and N. Merhav. A measure of relative entropy between individual sequences with application to universal classification. *IEEE Transactions on Information Theory*, 39(4):1270–1279, Jul 1993. S4

[8] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012. S4, S13

[9] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461, 2000. S5

[10] R. I. M. Dunbar. The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5):178–190, 1998. S6

[11] Science Fiction and Fantasy Writers of America, Inc. Nebula rules. http://nebulas.sfwa.org/about-the-nebulas/nebula-rules/. Accessed: 2017-08-05. S8

[12] W. N. Francis and H. Kucera. Brown corpus manual. *Brown University*, 2, 1979. S9

[13] S. Bird, E. Klein, and E. Loper. *Natural language processing with Python*. O'Reilly Media, Inc., 2009. S10