# An Activity-Based Information-Theoretic Annotation of Social Graphs

Arun V. Sathanur
Department of Electrical Engineering
University of Washington
Seattle, WA – 98195
1-206-616-5606
arunsv@uw.edu

Vikram Jandhyala
Department of Electrical Engineering
University of Washington
Seattle, WA - 98195
1- 206-543-6515
vj@uw.edu

## ABSTRACT
The explosion in social media adoption has opened up new opportunities to understand human interaction and information flow at an unprecedented scale. Influence between people represented as nodes of a social graph is best characterized in terms of the direction, the volume and the delay associated with the information flow. In this work we investigate the relatively new information-theoretic measure called *transfer entropy* as a measure of directed causal influence in online social interactions. The classical definition of transfer entropy is extended to a form applicable to activity on social graphs characterized by causal influence through delayed responses. For fixed but arbitrary interaction delays, we show that the swept delayed transfer entropy (DTE) profile peaks at the true delay. By extending the results to discrete and continuous distributions of interaction delays, the efficacy of DTE in recovering the interaction delay distributions between two causally related signals is demonstrated. An information theoretic annotation of social graphs that captures the volume and velocity of information transfer is presented based on the swept DTE.

## Categories and Subject Descriptors
H.1.1 [**Systems and Information Theory**]: Information Theory; J.4 [**Social and Behavioral Sciences**]: Sociology

## General Terms
Measurement, Performance, Reliability, Human Factors

## Keywords
Time Series, Transfer Entropy, Information Theory, Social Networks, Causality, Directed Influence, Delay Distribution

## 1. INTRODUCTION & RELATED WORK

With a significant fraction of the human population partaking in one or more online social networks (OSNs), the volume and velocity of big-social data is leading to information transfer at an extraordinary scale. Data from social media can best be visualized as a collection of a large number of time series with causal interdependence. Treating every node on an OSN as a source of time series data that is causally influenced by that of friends and followees allows mature information-theoretic and signal processing concepts [1, 2] to be applied in the rigorous influence analysis of big social data. Tie-strengths or connectedness measures on social graphs can represent the extent of similarity between the pair of nodes that constitute an edge, over a set of attributes [3]. However in the case of OSN graphs, such measures usually denoted as edge-weights are in general asymmetric and are a reflection of the directed influence of one node on the other in terms of activity [4]. Another distinguishing feature of OSNs is the causality-based influence that is implicit in the mechanism of activity generation. Hence an influence measure that can capture the volume and velocity (causal delay) of the interaction between nodes will be extremely beneficial as a directed influence measure in OSNs that can then be leveraged by a variety of graph-analytics applications. In this paper we investigate transfer entropy [5] as such a measure of directed influence based on causal responses in OSNs.

Related work in the area of influence detection in OSNs comprises of a number of topology and activity-based approaches. Widely used topology-based approaches include metrics such as the number of followers the user has accumulated, PageRank and its variants [6-9]. Recently, activity-based approaches that make use of weighted digraphs have been proposed to identify influential users on OSNs [10, 11]. The above approaches are rooted in the magnitude of the edge weights that characterizes the local influence model between the nodes constituting the edge while neglecting the delay aspect of the influence. Very recently researchers have started applying information-theoretic [12] and in particular transfer entropy concepts [13-15] to study activity and influence in OSNs. Transfer entropy is well studied in neurobiology where its potential in accurately identifying causal connections between individual neurons and between groups of neurons has been demonstrated [16-18].

In the context of OSNs, transfer entropy is particularly useful when only the activity traces are available and a causal relationship is sought. Thus in its native form, transfer entropy is particularly attractive to analyze networks such as Wikipedia and YouTube. Even in the case of additional headers such as retweets or sharing information being available, transfer entropy analysis can assist in uncovering patterns not captured by simple weight-based approaches. For example, a person not responding when a particular connection posts a message, or posting similar content without acknowledging as retweets or shares. Transfer entropy-based analysis can also point to causal influence between two nodes due to factors outside the network. Such an analysis can help uncover covert groups on OSNs that are not connected at the network level. Similar arguments expand the utility of transfer entropy to human networks beyond OSNs where individual


*WebSci '14,* June 23 - 26 2014, Bloomington, IN, USA

http://dx.doi.org/10.1145/2615569.2615673

activity traces can be observed and a causal link is sought including in a number of practical online applications.

In this work we investigate some modifications to the original transfer entropy definition that enable deeper insights into the nature of causal influence between nodes on an OSN. Section 2 defines transfer entropy and discusses the treatment of delay models in social networks within the transfer entropy framework. In section 3, the original transfer entropy definition is extended to include arbitrary but fixed delays. Tradeoffs involving the choice of bin size are examined. Section 4 examines parameterized distributions of interaction delay and demonstrates the recovery of delay distributions from the delayed transfer entropy profiles while introducing an associated information-theoretic annotation of social graphs. Section 5 concludes the paper and outlines ongoing and future work.

## 2. TRANSFER ENTROPY DEFINITIONS

One distinguishing feature of OSNs is the directed causality-based influence that is implicit in the mechanism of activity generation. Thus correlation-based measures struggle to capture the directed causal influence, as do symmetric information-theoretic measures such as mutual information [5]. Transfer entropy, first proposed by Schreiber [5], overcomes these limitations by specifically incorporating directedness and causality in the definition that is based on the Kullback relative entropy.

Consider two discrete-valued, discrete-time stochastic processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ over a common alphabet $\Psi$. It is assumed that $\Psi = [0,1]$ where a 0 denotes absence of activity and a 1 denotes presence of activity such as posting a message, uploading a photo or a video, commenting, sharing and re-tweeting. The transfer entropy from $\boldsymbol{X}$ to $\boldsymbol{Y}$ is given by

$$TE_{X\to Y} = H\left(Y_{n+1}\middle|\boldsymbol{Y_n^k}\right) - H\left(Y_{n+1}\middle|\boldsymbol{Y_n^k}, \boldsymbol{X_n^l}\right) \quad (1)$$

Here $\boldsymbol{Y_n^k}$ is the vector $[Y_n\, Y_{n-1} Y_{n-2} \ldots Y_{n-k+1}]$ that denotes the $k$-length history of $\boldsymbol{Y}$ at the time snapshot $n$ upon which $Y_{n+1}$ might depend. Similarly $\boldsymbol{X_n^l}$ is the $l$-length history of $\boldsymbol{X}$ at $n$. The difference of the two conditional entropies as in eq. (1) is non-negative and is a measure of the causal influence from $\boldsymbol{X}$ to $\boldsymbol{Y}$.

The quantity defined in eq. (1) is called the delay-1 transfer entropy since it relates the past values of $\boldsymbol{X}$ and $\boldsymbol{Y}$ at time $n$ to the value of $\boldsymbol{Y}$ at time $(n+1)$. The definition in eq. (1) however cannot be generally applied to time series from social data since it requires extension to the wide range of delays encountered in responses on human networks including OSNs [19, 20]. Ver Steeg and Galystyan [14] address this issue by increasing the length of the history of $\boldsymbol{X}$ and by using non-uniform bin sizes. Bauer *et al.* [13] also increase the length of the history of $\boldsymbol{X}$ but the approach is an empirical one based on testing different values of $l$. It is known that increasing the order, which amounts to conditioning on multiple variables, results in increased computational complexity and leads to accuracy issues when dealing with sparse data because of the exponentially larger alphabet size [5, 17].

## 3. DELAYED TRANSFER ENTROPY

Consider two binary processes $X(t)$ and $Y(t)$ with $Y(t)$ causally dependent on $X(t)$ with a delay of $t_d$. By using an appropriate choice of the bin size, $X(t)$ and $Y(t)$ can be converted to discrete-time processes such that the original delay-1 transfer entropy is still applicable. However, the loss of resolution by using a large bin size can be a serious handicap. To circumvent this the delay-1 transfer entropy can be extended to delay-$d$ transfer entropy by replacing $\boldsymbol{X_n^l}$ with $\boldsymbol{X_{n-d}^l}$ where $d$ is the discrete time equivalent of the delay $t_d$ [17, 21]. Thus

$$TE_{X\to Y}^{d} = H\left(Y_{n+1}\middle|\boldsymbol{Y_n^k}\right) - H\left(Y_{n+1}\middle|\boldsymbol{Y_n^k}, \boldsymbol{X_{n-d}^l}\right) \quad (2)$$

To further clarify the definition of delayed transfer entropy, it can be noted that under stationarity, $TE_{X\to Y}^{d} = H\left(Y_{n+d+1}\middle|\boldsymbol{Y_{n+d}^k}\right) - H\left(Y_{n+d+1}\middle|\boldsymbol{Y_{n+d}^k}, \boldsymbol{X_n^l}\right)$ whereas $TE_{X\to Y}^{d} \neq H\left(Y_{n+d+1}\middle|\boldsymbol{Y_n^k}\right) - H\left(Y_{n+d+1}\middle|\boldsymbol{Y_n^k}, \boldsymbol{X_n^l}\right)$ [21]. The dependence of $Y_{n+1}$ on the history of $X$ is delayed while the dependence on the history of $Y$ is still retained at the previous time-step. $TE_{X\to Y}^{d}$ is called the delayed transfer entropy (DTE). The efficacy of DTE in recovering the interaction delay distribution is investigated via the following steps.

1. A receiver point process $Y(t)$ that is causally dependent on a source point process $X(t)$ according to a known forward delay model is generated. In this section, in order to emphasize the delay aspects, all the spikes in the activity of $X$ appear in the activity of $Y$ according to the various delay models discussed.
2. Working with only the two time series, binning them and sweeping the delay in eq. (2) a profile of the DTE is built to relate to the original model-based delay. The simplicity and efficiency of computation can be retained by choosing $k = l = 1$ while the effect of increasing history length is addressed through the delay parameter. Further independent computations mean that parallel implementations are very straightforward.

Consider a message source $X$ on an OSN generating activity in the form of a Poisson process. Let $Y$ be a follower of $X$ that engages in influenced activity (for example retweeting on Twitter) after a fixed delay of $t_d$. The choice of Poisson process is representative and the analysis can be easily applied to other types of stationary point processes. For further discussion on the stationarity aspects involved in transfer entropy estimation the reader is referred to the work by Kaiser and Schreiber [22]. Figure 1 shows the profile of the swept DTE for the process rate $\lambda = 1\, hr^{-1}$ and $t_d = 6\, hr$ at three different bin widths of $0.05\, hr$, $0.2\, hr$ and $0.5\, hr$. The computation of $TE_{X\to Y}^{d}$ can be considerably simplified by adopting the formula based on joint Shannon entropies [22] namely

$$TE_{X\to Y}^{d} = H\left(\boldsymbol{Y_n^k} \otimes \boldsymbol{X_{n-d}^l}\right) - H\left(Y_{n+1} \otimes \boldsymbol{Y_n^k} \otimes \boldsymbol{X_{n-d}^l}\right) + H\left(Y_{n+1} \otimes \boldsymbol{Y_n^k}\right) - H\left(\boldsymbol{Y_n^k}\right) \quad (3)$$

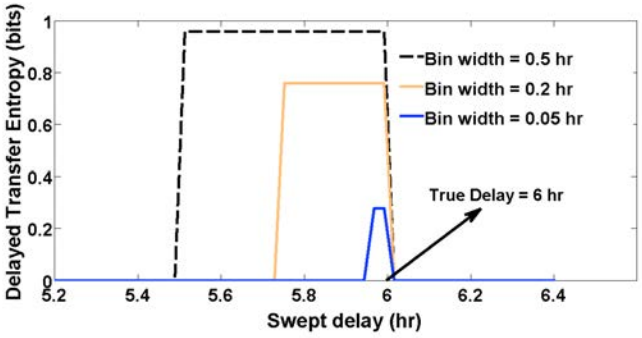


**FIG. 1. For a forward model with a single delay, the swept delayed transfer entropy profile has a maximum at the true delay. The resolution and the profile peak are dependent on the bin width used in the discretization.**

The following points are noteworthy in the above figure.

a. Transfer entropy is maximized at the true delay. Wibral *et al.* [21] provide a theoretical justification of this result based on the data processing inequality.

b. Bin size has an effect on the maximum value of transfer entropy.

c. A larger bin size, while reducing the computational burden, introduces ambiguity in the delay estimation.

While point (c) is apparent, the rest can be understood by the following analysis. Let $h(\lambda, b_W)$ denote the Shannon entropy of a binned Poisson point process $\boldsymbol{X}$. Let $t_{d_1}$ be the swept delay used in the transfer entropy computation and let $d_1$ and $d$ be the discrete time versions $t_{d_1}$ and $t_d$ respectively. In the above scenario $Y_{n+1} = X_{n-d_1}\, for\;\; d_1 = d$. Otherwise $Y_{n+1}$ is independent of $X_{n-d_1}$. From eq. (3), it is apparent that

$$TE_{X\to Y}^{d_1} = TE_{X\to Y}^{max} = h(\lambda, b_W)\,; d_1 = d \tag{4a}$$

$$TE_{X\to Y}^{d_1} = 0\,; d_1 \neq d \tag{4b}$$

For a Poisson point process with rate $\lambda$ characterized by an exponential inter-arrival density function of $\lambda e^{-\lambda t}$, the probability that there is at-least one spike in the time interval corresponding to $b_W$ is $\int_0^{b_W} \lambda e^{-\lambda t} = \left(1 - e^{-\lambda b_W}\right)$ leading to

$$TE_{X\to Y}^{d} = h(\lambda, b_W) = \left(\left(1 - e^{-\lambda b_W}\right) log_2\left(1 - e^{-\lambda b_W}\right) + e^{-\lambda b_W}\, log_2(e^{-\lambda b_W})\right) \tag{5}$$

$TE_{X\to Y}^{d}$ is maximized at $b_W = \left(\frac{ln\,2}{\lambda}\right)$. Assuming that the rate can be estimated from the data, the choice of optimum bin size for sampling boils down to the tradeoff between acceptable delay uncertainty, the computational complexity and the height of the DTE profile especially in the presence of noise. The optimum delay step itself is the same as the bin width and a bin width of about a third of the optimum works well in practice given the various considerations.

Figure 2 shows the tradeoff by sweeping the bin sizes for various $\eta$ where $\eta$ is the parameter of the Bernoulli noise process added to the receiver signal $Y(t)$. The rate of the source Poisson process is $2.3hr^{-1}$. It must be noted that when the unknown delay is not an integral multiple of the bin width, during the process of digitizing the signals, there can be leaks into the adjacent bins due to rounding off errors. In such cases the transfer entropy peak may not be sharp as predicted by eqs. (4a) and (4b). This can be remedied by searching in the local space around a given bin width for the maximum transfer entropy.

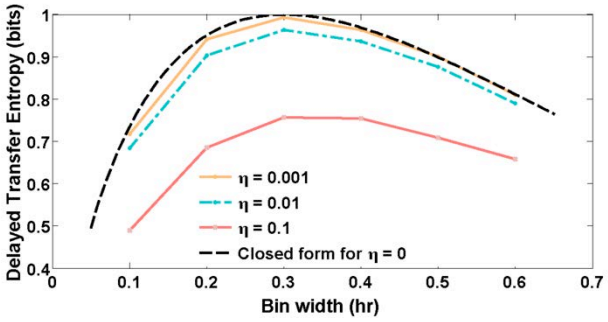


**FIG. 2. The peak of the swept delayed transfer entropy (Fig. 1) depends on the bin width and has an optimum with respect to it. The peak transfer entropy also reduces in the presence of noise.**

# 4. SWEPT DELAYED TRANSFER ENTROPY FOR DISTRIBUTIONS OF DELAY

We now introduce a multiple-delay model before moving to distribution of delays. First, three different delays $t_{d_1}, t_{d_2}, t_{d_3}$ are introduced with the delays being selected in the forward model with probabilities $w_1, w_2, w_3$ respectively and $w_1 + w_2 + w_3 = 1$. The swept DTE profiles are shown in figure 3 where it can be seen that the strongest delay components are easily recovered. The DTE peaks corresponding to the true delays depend in a non-linear fashion on the delay selection probabilities $w_1, w_2, w_3$. This can be understood by the following analysis.

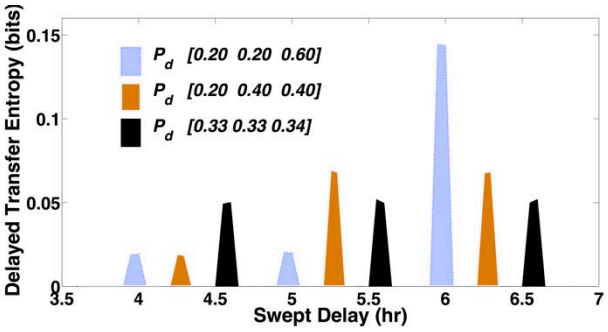


**FIG. 3. The delayed transfer entropy response for the three-delay model for different probability distributions of the delays. The delays are purposely misaligned for the three different cases to improve visualization. $P_d$ refers to the vector of probabilities $[w_1\ w_2\ w_3]$. The nonlinear dependence of the DTE peak values on the probability of the delay occurring is evident.**

Consider a Poisson process $\boldsymbol{X}$ with rate $\lambda$ and a process $\boldsymbol{Y}$ that is influenced by $\boldsymbol{X}$ such that a portion $w$ $(< 1)$ of the spikes in $\boldsymbol{X}$ appear in $\boldsymbol{Y}$ with a delay $t_d$. By treating a binned Poisson process (with bin size $b_W$) as a Bernoulli process with $p(1) = p(\lambda, b_W) = (1 - e^{-\lambda b_W})$, using eq. (3), after some algebra the peak DTE value corresponding to a delay of $t_d$ becomes

$$TE_{X\to Y}^{d} = \xi(w,p) = 2g(wp) + g\big((1-w)p\big) + g(1-wp) - g(w^2p) - 2g\big((1-w)wp\big) - g((1-w)^2p) \tag{6}$$

Here $g(x) = -xlog_2(x)$ and $p = p(\lambda, b_W) = (1 - e^{-\lambda b_W})$. It can be seen that $\xi(w,p)$ is approximately linear in $w$ for small $p(\lambda, b_W)$. However in the presence of multiple delays as in the case of figure 3, when computing the DTE, at any given physical delay ($t_{d_1}$ *or* $t_{d_2}$ *or* $t_{d_3}$ in the above example), the interference due to other spikes that suffer a different delay but still appear within the bin corresponding to the delay under consideration reduces the height of the transfer entropy peak. It can be shown that under these circumstances with the total delay probabilities summing to 1, the DTE for the delay under consideration can be approximated by $\xi\left(w^2, p(\lambda, b_W)\right)$ which by prior considerations is close to quadratic in $w$. The same is verified in figure 4. Here, a Bernoulli process with probability of the symbol $'1'$ given by $p = 0.2$ undergoes four different delays with probabilities given by $\left[w\ \frac{2(1-w)}{3}\ \frac{(1-w)}{6}\ \frac{(1-w)}{6}\right]$ respectively. The figure shows the analytical and simulated DTE profiles for delay 1 and delay 2 as $w$

is varied. The simulated curves were obtained from averaging over a 25 Monte Carlo run.

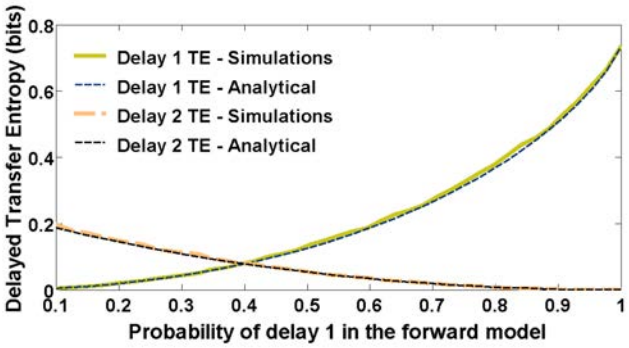


**FIG. 4. A Bernoulli process undergoes four different delays according to a probability mass function characterized by one parameter namely the probability of the first delay. The simulated and analytical results for the delayed transfer entropy at two different delays show excellent agreement enabling generalization to arbitrary delay distributions.**

In the most general case of the interaction delays following a probability distribution, the possibility of estimating the underlying distribution of the delay from DTE is examined next. Let the delay distribution density be given by $f_{t_d}(T_D)$ where $T_D$ is the random variable representing the delay $t_d$. When the delay is scanned as a parameter in the swept DTE computation, for a given $\Delta t_d$ denoting the spacing between the adjacent delays scanned and an index $k$ for the delay under consideration, the probability of the forward model delay being in the interval $[k\Delta t_d \;\; (k+1)\Delta t_d]$ is given by

$$w(k) = \int_{k\Delta t_d}^{(k+1)\Delta t_d} f_{t_d}(T_D)dt_d \approx f\left(\left(k+\frac{1}{2}\right)\Delta t_d\right)\Delta t_d \qquad (7)$$

This amounts to discretizing the density function $f(\,)$ to a sum of impulses as $\sum_k w(k)\delta\left(t_d - \left(k+\frac{1}{2}\right)\Delta t_d\right)$. From the earlier discussion on the multiple delay case, the profile of the DTE (at delay $d_k$) is therefore a discretized, nonlinear version of the underlying forward delay profile is approximately given by

$$TE_{X\to Y}^{d_k} \approx \xi\left(\left(f\left(\left(k+\frac{1}{2}\right)\Delta t_d\right)\Delta t_d\right)^2, p(\lambda, b_W)\right) \qquad (8)$$

This is best illustrated for the pathological case of a triangular delay distribution as in figure 5 where simulated and analytical results are shown for a Poisson process with $\lambda = 2.3\, hr^{-1}$, $b_W = 0.1\, hr$ and the delay follows a triangular distribution with minimum and maximum values of 2.4 $hr$ and 5.6 $hr$ respectively. The simulations involved averaging over a 25 run Monte Carlo. To emphasize the nonlinear transformation, the underlying triangular distribution of the delay is also shown. To facilitate the comparison, the DTE curves are normalized such that the area under the respective curves is unity.

Figures 6a and 6b show the recovered delay distribution for a delay with uniform distribution and a Gaussian distribution. In both the cases, two different standard deviations are shown to illustrate the efficacy of the method. The underlying process was a binned Poisson process with $\lambda = 2.3\, hr^{-1}$, $b_W = 0.1\, hr$. The $w(k)^2$ dependence of the DTE magnitude can be perceived from the fact that the standard deviation inferred from the DTE profile is $\frac{\sigma}{\sqrt{2}}$ where $\sigma$ is the standard deviation of the underlying delay distribution in the forward model. Thus the swept DTE profile is a *signature* of the underlying delay distribution in the forward model. The inference of the delay distribution can therefore be accomplished by in numerous ways including inverting eq. (8) by leveraging the square root of the DTE profile, lookup tables and formal machine learning methods.

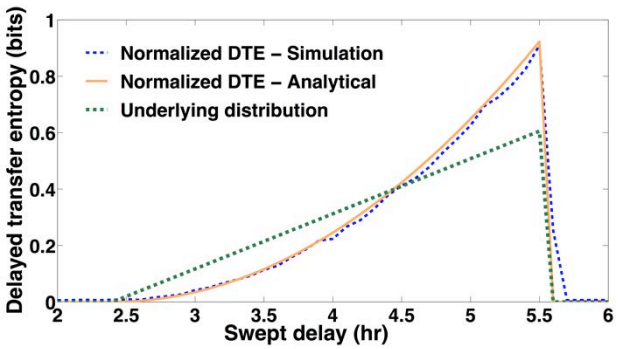


**FIG. 5. A triangular delay distribution in the forward model is considered in order to illustrate the non-linear map from the probability density function of the delay to that of the swept profile of the delayed transfer entropy. To facilitate comparison the delayed transfer entropy curves are normalized so that the corresponding areas under the curves are each unity.**

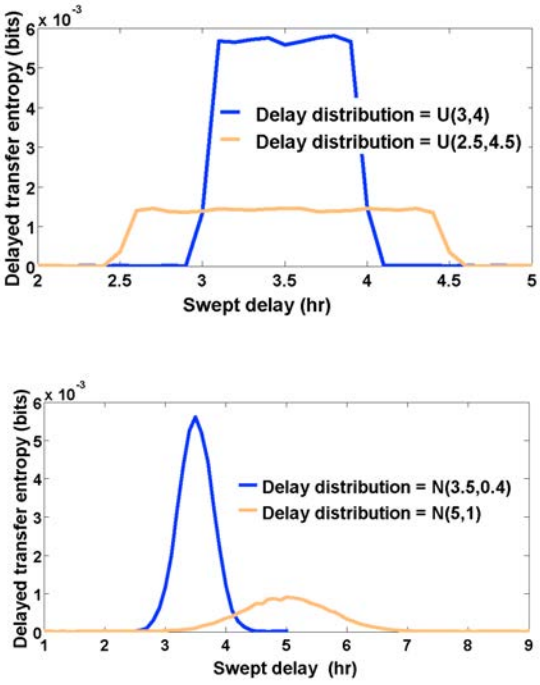


**FIG. 6. The swept delayed transfer entropy profiles for two different uniform (a) and Gaussian (b) delay distributions are illustrated. The swept delayed transfer entropy profiles enable the recovery of causal interaction delay distributions between two time series denoting activity.**

The swept DTE profile also allows for an *information-theoretic annotation* of an edge in a social graph that takes into account both the magnitude and delay involved in the causal influence. One such annotation is illustrated in figure 7. Here $TE^{\mu}$ refers to the average height of the DTE profile and $[t_d^{min} \;\; t_d^{max}]$ refers to the interval corresponding to the DTE above a given threshold.

Annotations based on other statistical measures around the DTE profile are conceivable. Such an annotation allows for complex optimizations to be performed on social graphs to achieve specific outcomes in diverse application areas such as social marketing campaigns, project management in organizations and mobilization for collective action where time-sensitive information transfer plays a major role.

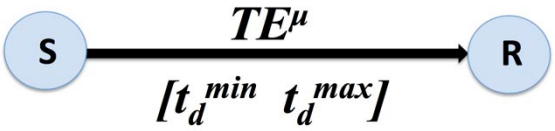


**FIG. 7. The swept DTE profile leads to an information theoretic annotation of an edge in a social graph. The annotation captures the magnitude of the causal influence in addition to the delay bounds**

## 5. CONCLUSIONS AND FUTURE WORK

In this work the effectiveness of swept DTE in recovering arbitrary distributions of delays in causal interactions characteristic of social networks was demonstrated. The swept DTE profile also allows an information-theoretic annotation on social graphs that capture the volume and velocity of causal influence and hence can be used in a variety of applications involving time-sensitive information transfer. Current and future work involve extending the swept DTE concepts to facilitate discovery of more complex activity mechanisms, application to real-world social data and exploring the effect of sampling.

## 6. REFERENCES


[1] T. M. Cover and J. A. Thomas, *Elements of information theory*: John Wiley & Sons, 2012.

[2] J. D. Hamilton, *Time series analysis* vol. 2: Princeton university press Princeton, 1994.

[3] M. Gupte and T. Eliassi-Rad, "Measuring tie strength in implicit social networks," in *Proceedings of the 3rd Annual ACM Web Science Conference*, 2012, pp. 109-118.

[4] S. Hangal, D. MacLean, M. S. Lam, and J. Heer, "All friends are not equal: Using weights in social graphs to improve search," *Workshop on Social Networking Mining and Analysis, ACM KDD*, 2010

[5] T. Schreiber, "Measuring information transfer," *Physical review letters,* vol. 85, p. 461, 2000.

[6] B. Bahmani, A. Chowdhury, and A. Goel, "Fast incremental and personalized PageRank," *Proceedings of the VLDB Endowment,* vol. 4, pp. 173-184, 2010.

[7] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Computer networks and ISDN systems,* vol. 30, pp. 107-117, 1998.

[8] B. Hajian and T. White, "Modelling influence in a social network: Metrics and evaluation," in *Privacy, security, risk and trust (PASSAT), Third IEEE international conference on social computing (socialcom)*, 2011, pp. 497-500.

[9] Y. Yamaguchi, T. Takahashi, T. Amagasa, and H. Kitagawa, "Turank: Twitter user ranking based on user-tweet graph analysis," in *Web Information Systems Engineering–WISE 2010*, ed: Springer, 2010, pp. 240-253.

[10] A. V. Sathanur, V. Jandhyala, and C. Xing, "PHYSENSE: Scalable sociological interaction models for influence estimation on online social networks," in *Intelligence and Security Informatics (ISI), 2013 IEEE International Conference on*, 2013, pp. 358-363.

[11] J. Weng, E.-P. Lim, J. Jiang, and Q. He, "Twitterrank: finding topic-sensitive influential twitterers," in *Proceedings of the third ACM international conference on Web search and data mining*, 2010, pp. 261-270.

[12] C. Wang and B. A. Huberman, "How random are online social interactions?," *Scientific reports,* vol. 2, 2012.

[13] T. L. Bauer, R. Colbaugh, K. Glass, and D. Schnizlein, "Use of transfer entropy to infer relationships from behavior," in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*, 2013, p. 35.

[14] G. Ver Steeg and A. Galstyan, "Information transfer in social media," in *Proceedings of the 21st international conference on World Wide Web*, 2012, pp. 509-518.

[15] G. Ver Steeg and A. Galstyan, "Information-theoretic measures of influence based on content dynamics," in *Proceedings of the sixth ACM international conference on Web search and data mining*, 2013, pp. 3-12.

[16] B. Gourévitch and J. J. Eggermont, "Evaluating information transfer between auditory cortical neurons," *Journal of Neurophysiology,* vol. 97, pp. 2533-2543, 2007.

[17] S. Ito, M. E. Hansen, R. Heiland, A. Lumsdaine, A. M. Litke, and J. M. Beggs, "Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model," *PloS one,* vol. 6, p. e27431, 2011.

[18] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, "Transfer entropy—a model-free measure of effective connectivity for the neurosciences," *Journal of computational neuroscience,* vol. 30, pp. 45-67, 2011.

[19] A. Vázquez, J. G. Oliveira, Z. Dezsö, K.-I. Goh, I. Kondor, and A.-L. Barabási, "Modeling bursts and heavy tails in human dynamics," *Physical Review E,* vol. 73, p. 036127, 2006.

[20] Z. Yang, J. Guo, K. Cai, J. Tang, J. Li, L. Zhang*, et al.* (2010, Understanding retweeting behaviors in social networks. *Proceedings of the 19th ACM international conference on Information and knowledge management,* 1633-1636.

[21] M. Wibral, N. Pampu, V. Priesemann, F. Siebenhühner, H. Seiwert, M. Lindner*, et al.*, "Measuring information-transfer delays," *PloS one,* vol. 8, p. e55809, 2013.

[22] A. Kaiser and T. Schreiber, "Information transfer in continuous processes," *Physica D: Nonlinear Phenomena,* vol. 166, pp. 43-62, 2002.