Temporal Walk Based Centrality Metric for Graph Streams

Ferenc Béres[1,2] Róbert Pálovics[3] András A. Benczúr[1]
[1]Institute for Computer Science and Control of the Hungarian Academy of Sciences
[2]Eötvös Loránd University Budapest
[3]Stanford University
{beres, benczur}@sztaki.hu, palovics@stanford.edu

ABSTRACT

Centrality measures account for the importance of the nodes of a network. In the seminal study of Boldi and Vigna (2014), the comparative evaluation of centrality measures was termed a difficult, arduous task. In networks with fast dynamics, such as the Twitter mention or retweet graphs, predicting emerging centrality is even more challenging. Our main result is a new, temporal walk based dynamic centrality measure that models temporal information propagation by considering the order of edge creation. This measure outperforms graph snapshot based static and other recently proposed dynamic centrality measures in assigning the highest time-aware centrality to the actually relevant nodes of the network. One of our main contributions is creating a quantitative experiment to assess temporal centrality metrics. Our data and codes are publicly available[1].

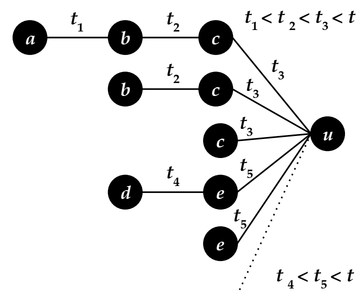


Figure 1: Temporal walks ending at node u before time t.

1 INTRODUCTION

We present **temporal Katz centrality**, an online updateable graph centrality metric for measuring user importance over time. We consider temporal networks where the edges of the network arrive continuously in time. In other words the graph is represented as a sequence of time-stamped edges [20]. Our metric is based on the concept of time-respecting walks containing a sequence of adjacent edges ordered in time. As seen in Figure 1, temporal Katz centrality aggregates each temporal walk ending at node u before time t.

Our new method is a graph algorithm for online machine learning [5]. Our goal is to reduce the delay caused by collecting data only for the range of hours to process as a graph snapshot. We provide online updateability, which poses computational restrictions and challenges to most centrality measures and graph algorithms.

Although many studies tried to identify the best estimates for the importance of a social media user, to the best of our knowledge, there is only one previous study [20] that proposes **data stream updateable** centrality measures. However, their algorithm, which we analyze in Section 3.2, cannot incorporate the actual edge arrival times in its calculations. We believe our method is superior in using the exact time of interaction between two social media users, resulting in better performance in our prediction task.

Another key issue that we address is the difficulty of the timely evaluation of fast changes in social media. Even a static ground truth labeling for a static centrality measure requires tedious human effort. In [6], for example, the Text Retrieval Conference (TREC) topics are used. In a dynamic graph, depending on time granularity, the same human data curation may be required in each time step. For example, in the study most similar to ours [20], only small temporal social network snapshots are collected, and evaluation is mostly based on convergence to static centrality measures.

[1]GitHub repository of our research: https://github.com/ferencberes/online-centrality

In our best effort to provide quantitative evaluation for dynamic centrality, we consider daily granularity and **compile ground truth** based on an external source. We collect tweets about Roland-Garros 2017, the French Open Tennis Tournament (RG17), and US Open 2017, the United States Open Tennis Tournament (UO17). We compute both static and dynamic centrality metrics over the time-aware mention graph that we extract from the tweets. We define the mention graph by adding a time-stamped edge (u, v, t) whenever user u mentions v in a tweet at time t. For ground truth, we consider the Twitter accounts of players participating in daily rounds as relevant. We then hour by hour investigate how mentions of players for the coming day take over the importance of past participants.

In this paper, we design and evaluate an online updateable, dynamic graph centrality measure. Our main contribution is threefold: (1) We propose a new, online updateable path count based centrality measure as a temporal variant of the successful Katz index [12]. Our measure incorporates arbitrary time decay functions that can be adapted to the task in question. (2) We compile a data set with ground truth labels for the quantitative evaluation of dynamic centrality. Our evaluation is based on our Twitter collection about tennis tournaments. For centrality ground truth at a given time, we set the players participating in rounds on given days. (3) We experiment over Twitter tennis tournament data sets and observe that our method outperforms the temporal PageRank of [20].

2 RELATED RESULTS

To quantify the popularity of a node, several graph centrality measures have been proposed [6]. The definitions of centrality vary greatly and incorporate both global and local factors of a node's location within the network. The high variability of centrality scores reflects the nature of popularity observed in real-world networks [17]. Several models have been suggested to explain the emergence of high variability, habitually involving some variation of the preferential attachment mechanism, also extended to the dynamic setting [10].

For temporal networks, a few generalizations of static centrality measures to dynamic settings have been suggested recently [2, 9, 13, 22, 23]. In these works, tracking centrality of a single node and determining its variability play a major role [23], as it has been observed in the literature that centrality of nodes can change drastically from one time period to another [7].

The above results [2, 9, 13, 22, 23], however, cannot be used for computing and updating centrality online. The following results devise methods that are variants of our snapshot baselines: In [23], the spectrum of a set of discrete graph snapshots is analyzed in time; however, the spectrum cannot be dynamically updated with fine time granularity, as required by our application. Similarly, in [9], sequences of snapshots are considered. Finally, in [2, 13, 22], degree, closeness, and betweenness are considered in dynamic graphs, but these measures, with the exception of the degree, cannot be efficiently updated online. Note that online degree, also with time decay, is compared as a baseline method in our experiments.

In this paper we address a practically important variant of dynamic centrality: our goal is to compute online updateable measures that can be computed from a data stream of time-stamped edges. To the best of our knowledge, the only previous such algorithms are temporal PageRank [20] and temporal degree [13]—other measures are inefficient to update online. We show that our algorithm has better performance for assessing centrality in a dynamic graph, which we explain in Section 3 by showing that we can incorporate temporal information while keeping dynamic update computational costs very low. In fact, temporal PageRank is based on PageRank [19], while our method is based on the Katz index [12], both of which are shown to have similar theoretical and practical properties by [6].

To our knowledge, temporal PageRank [20] is the only published work about temporal generalizations of PageRank. Other results focus on coarse, static snapshots [15], or use temporal information to calculate edges of a static graph [11, 16]. Finally, another line of research considers updating PageRank in dynamic or online scenarios [3, 4, 14, 18, 21]; however, in these results PageRank is considered a stationary distribution over the current, static graph. In our experiments, we will show that our temporal Katz centrality outperforms snapshot-based static measures for assessing node importance in a temporally changing environment.

3 CENTRALITY MEASURES

Three axioms of centrality are defined in [6]. There is a single measure, harmonic centrality, that satisfies all three of them. Since the computation of harmonic centrality for a given node u involves all the distances from the node u in question, the measure is computationally challenging even in a static graph.

The starting point of our temporal Katz centrality measure is PageRank [19], which along with the Katz index satisfies the last two axioms defined in [6]. The importance of PageRank in our work has multiple reasons. On the one hand, it is widely used and has favorable properties by the axioms of [6]. On the other hand, temporal PageRank [20] is a modification of PageRank, which to the best of our knowledge is the only temporal ranking metric proposed in the literature prior to our work.

PageRank, Katz index, and temporal PageRank are all based on counting paths in the underlying networks. Next, we review the general properties of the path counting centrality metrics and temporal PageRank [20]. Then in Section 4, we describe our temporal Katz centrality measure.

3.1 Path counting centrality metrics

As perhaps the first centrality metric based on path counting, Katz introduced his index [12] as the summation of all paths coming into a node, but with an exponentially decaying weight based on the length of the path:

$$\vec{\text{Katz}} = \mathbf{1} \cdot \sum_{k=0}^{\infty} \beta^k A^k, \tag{1}$$

where $\vec{\text{Katz}}$ is the Katz index vector, A is the directed adjacency matrix, and $\beta < 1$ is a constant. Hence the Katz index of a node is the weighted sum of the number of paths of different lengths k terminating in u, where the weight is β^k:

$$\vec{\text{Katz}}(u) := \sum_{v} \sum_{k=0}^{\infty} \beta^k |\{\text{paths of length } k \text{ from } v \text{ to } u\}|, \tag{2}$$

The Katz index is finite only if $\beta < 1/|\lambda_1|$, where λ_1 is the eigenvalue of A with largest absolute value [12]. Since $1/|\lambda_1|$ is often very small, around 0.05 in our graphs, the relative weight of a length two path stays very small compared to a single edge. In order to be able to use larger values of β, we introduce the truncated Katz index as

$$\vec{\text{Katz}}^{[K]} = \mathbf{1} \cdot \sum_{k=0}^{K} \beta^k A^k. \tag{3}$$

By the basic definition, PageRank is normally considered to be the static distribution of a random walk with damping [19]. In order to compare PageRank and the Katz index, and to motivate online update rules, we use the result of [8], who show—and use as an efficient algorithm—that PageRank is equal to the path counting formula

$$\vec{\text{PageRank}} = \mathbf{1} \cdot \frac{c}{N} \cdot \sum_{k=0}^{\infty} (1-c)^k M^k, \tag{4}$$

where c is the damping constant and M is the random walk transition matrix. In other words, M is the outdegree normalized adjacency matrix: $M = (K^{-1}A)^T$ where K is a diagonal matrix with the outdegrees in the diagonal.

3.2 Temporal PageRank

In [20], temporal PageRank, a dynamic variant of PageRank, is defined as follows. In a dynamic graph, edges are time-stamped and can appear multiple times. The main idea is to aggregate **time respecting temporal walks**

$$z = (u_0, u_1, t_1), (u_1, u_2, t_2), \cdots, (u_{j-1}, u_j, t_j); \qquad t_{i-1} \le t_i. \tag{5}$$

ending in a certain node to compute its temporal centrality. In such a walk, they model an information flow from the start node u_0 to the destination u_j by passing along edges that arrive subsequently in time.

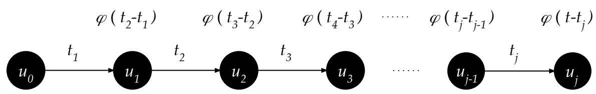


Figure 2: Edge weights along a temporal walk at time t.

For each edge (u_{i-1}, u_i, t_i) in walk z, they assign the transition weight as β^k, where $\beta < 1$ is a decay constant and k is the number of edges (u_{i-1}, y, t') that appear after the previous edge but not later than the present edge in the walk, that is, $t_{i-1} < t' < t_i$. They incorporate this weight assignment in formula (4), for full details, see [20].

Intuitively, their notion of edge transition weight decays exponentially with the number of possible continuations of the temporal walk at node u_{i-1}. The more edges appear before (u_{i-1}, u_i, t_i), in their model it is exponentially less likely that the information is sent along the given edge—and not another edge that appears earlier.

The main problem with the above path counting algorithm is that it overvalues nodes with low activity. Consider a node that communicates to ten contacts in a few minutes. The tenth contact will only receive a propagated score proportional to β^{-10}. By contrast, if another node sends only one message per day, the neighbor receives the full score even though the information may already be highly outdated.

One key motivation of the above definition for temporal PageRank is that it possesses a computationally low cost update algorithm. While it is tempting to modify the weight formula to incorporate the actual time elapsed, the stream-based computation of such a modified temporal PageRank becomes unclear.

4 TEMPORAL KATZ CENTRALITY

We define our temporal Katz centrality measure over the stream of edges arriving in time from a dynamic network. Our goal is to specify a metric that is based on the weighted sum of time respecting walks, updateable by the edge stream, and that can incorporate the actual elapsed time in the weights of the walks.

To motivate our new method, we reconsider the temporal PageRank [20] edge transition weight rule that involves time decay in an indirect way through a combination with the activity of the nodes. Hence here the notion of time is difficult to directly capture as the more time elapses before the next edge appears, the more other edges have the chance to appear in between.

We define **temporal Katz centrality** by introducing a natural, purely time-dependent edge transition weight $\varphi(\tau)$, which is an arbitrary function of the time elapsed since the previous edge in a path. Temporal Katz centrality is the weighted sum of all time respecting walks that end in node u,

$$r_u(t) := \sum_{v} \sum_{\substack{\text{temporal paths } z \\ \text{from } v \text{ to } u}} \Phi(z, t) \tag{6}$$

where $\Phi(z, t)$ is the weight of walk z at time t. Truncated temporal Katz centrality is defined similar to equation (3) by restricting to walks of length at most K. For a temporal walk as in equation (5) where edges appeared at $(t_1, t_2, ..., t_j)$, we define weight $\Phi(z, t)$ as

$$\Phi(z, t) := \prod_{i=1}^{j} \varphi(t_{i+1} - t_i). \tag{7}$$

where φ is a time-aware weighting function, and for $i = j$ we let $t_{j+1} := t$. Hence $\Phi(z, t)$ is the product of individual edge transition weights $\varphi(t_{i+1} - t_i)$ as seen in Figure 2. The last term of the product $\varphi(t - t_j)$ captures the delay between present time t and the appearance of the last edge in the path. In other words, temporal Katz centrality is the variant of the Katz index (1) in which time respecting paths are weighted by $\Phi(z, t)$. By using different edge weight functions, we cover two important special cases.

- If $\varphi(\tau) := \beta$ is constant, we obtain a variant of the Katz equation (2) with summation for temporal paths instead of all paths irrespective of time.
- In another special case, $\varphi(\tau) := \beta \cdot \exp(-c\tau)$. Since φ is an exponential function, $\varphi(a) \cdot \varphi(b) = \varphi(a + b)$. Hence the path weight in (7) becomes

$$\begin{aligned}\Phi(z, t) &= \beta \exp(-c[t - t_j])...\beta \exp(-c[t_2 - t_1]) \\ &= \beta^{|z|} \exp(-c[t - t_1]),\end{aligned} \tag{8}$$

  that is, it involves a Katz-style decay proportional to the length of the path, combined with an exponential decay depending on the time elapsed since the first interaction t_1 over the path occurred. This weight is capable of capturing the temporal decay of information spreading and propagation.

Online update. We base our analysis below on the fact that temporal Katz centrality, which is the sum of all temporal paths to u can be derived by using the number of temporal paths ending at the in-edges of u. If edge vu appears at time t_{vu}, the future centrality of node u at time t increases as (1) a new time respecting walk appears that starts from v and has weight $\varphi(t - t_{vu})$, (2) for each time respecting walk that ended in v at t_{vu}, a new walk with the new edge vu appears. As the total weight of paths that ended in v is $r_v(t_{vu})$, by adding up the weight of the two types of new walks we get

$$r_u(t) = \sum_{vu \in E(t)} (1 + r_v(t_{vu}))\, \varphi(t - t_{vu}), \tag{9}$$

where $E(t)$ is the multi-set of edges appearing no later than t.

Based on the above recursive formula, if edge vu appears at time t_{vu}, it increases the future centrality of node u by $(1 + r_v(t_{vu}))\, \varphi(t - t_{vu})$. The increase of the centrality of u can be computed by maintaining the values t_{vu} and $w_{vu} := 1 + r_v(t_{vu})$. When edge vu appears, first we calculate the current value of r_v as

$$r_v := \sum_{zv \in E(t)} w_{zv} \cdot \varphi(t - t_{zv}). \tag{10}$$

Then we add a new edge vu to the multi-set of edges $E(t)$ with $w_{vu} := r_v + 1$ to propagate the centrality score along edge vu, and set $t_{vu} := t$.

The time complexity of the above algorithm is linear in the degree of u. It is possible to further improve the online update complexity to constant time per update if φ satisfies $\varphi(a + b) = \varphi(a) \cdot \varphi(b)$.

Convergence properties. Let us assume that we sample a sequence of T edges from a graph with edge set of size E. We can

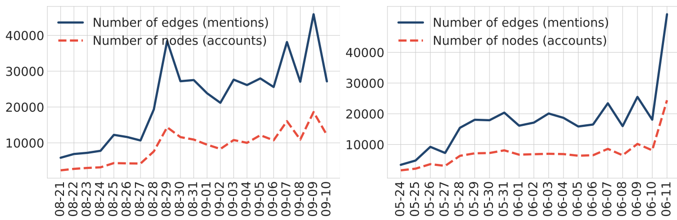


Figure 3: Number of nodes and edges. Left: UO17; Right: RG17. Low activity during the qualifiers and peak near semi-finals and finals.

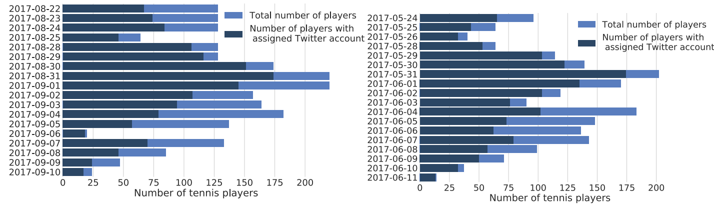


Figure 4: The number of players active on a given day and the number of of them with identified Twitter accounts. Left: UO17; Right: RG17.

compute the expected value of temporal Katz centrality over the sampled edge stream. We assume that sampling is done in a uniform way over time, hence in what follows, time t corresponds to the number of sampled edges in the process. The proof of the theorem below is left for an extended version of the paper.

Theorem 4.1. *Let us sample a sequence of T edges from an edge set of size E. Let us compute (truncated or normal) temporal Katz centrality with exponential weighting, $\varphi(\tau) := \beta \exp(-c\tau)$. Then as $T \mapsto \infty$, the limit of the expected value of temporal Katz centrality is*

$$\vec{TemporalKatz} = \mathbf{1} \cdot \sum_{k=0}^{K} A^k \left(\frac{\beta}{E}\right)^k \left(\frac{1}{e^c - 1}\right)^k. \qquad (11)$$

5 TWITTER TENNIS DATA SETS

We compiled two separate tweet collections, *RG17* for Roland-Garros 2017, the French Open Tennis Tournament, and *UO17* for US Open 2017, the United States Open Tennis Championship. The events took place between May 22 and June 11 as well as August 21 and September 10, respectively. We assessed the temporal relevance of centrality measures by using the list of players of different days as ground truth. We gathered data with the Twitter Search API, by using the following two separate sets of keywords:
{@rolandgarros, #RolandGarros2017, #rolandgarros2017, #RG17, #RolandGarros, #rolandgarros, #FrenchOpen, #frenchopen, #rg17}
{#usopen, #Usopen, #UsOpen, #USOPEN, #usopen17, #UsOpen17, #Usopen2017, @usopen, #WTA, #wta, #ATP, #atp, @WTA, @ATPWorldTour, #Tennis, #tennis, #tenis, #Tenis}
The RG17 data covers the events of the championship starting May 24 with 444, 328 tweets, 815, 086 retweets, and 336, 234 time-stamped mentions. The UO17 data consists of 636, 810 tweets, 1, 048, 786 retweets, and 482, 061 mentions. The daily distribution of mentions is shown for both tennis events in Figure 3. Note that we imposed no language restrictions on the text of the tweets during the data collection process.

We measure the performance of centrality measures by means of comparison with the official schedule of the tournaments. The daily timetables are accessible in HTML file format and contain the following information for each tennis game:

- Full names of the participating players (two for singles and four for doubles games)
- Approximate time of the game during the day (e.g.: after 11:00, not before 15:00, etc.)
- Category and round identifier of the game (e.g. Women's Singles—Round 1, Men's Singles—Final)
- Court name, where the game took place (e.g. Grandstand, Arthur Ashe Stadium, etc.)
- Information about whether the game was canceled, resumed from a previous day, or the final result if completed.

Based on the approximate time of the games, we consider a player *active* for a given day if he or she participated in a *completed game*, a *canceled game*, or a *resumed game* on the same day. All of these events are expected to cause a social media burst.

One of the most time-consuming parts of our measurement was to assign Twitter accounts to tennis players. The total number of professional participants is 798 for US Open and 698 for Roland-Garros. Unfortunately, many of the players have no Twitter accounts.

We assigned players to accounts by the Twitter Search API's people endpoint; however, the API was sometimes unable to identify the accounts of the active players.

In case the people API endpoint failed to return the account of a player, we considered the *account name* (e.g. @rogerfederer, @RafaelNadal) and *name* (e.g. "RafaNadal" for the account @RafaelNadal). Using edit distance, for each player we automatically selected accounts where the *account name* or the displayed *name* is very similar to the full name. Note that the same player often has multiple Twitter accounts, especially the popular players, who usually have official sites and distinct accounts for fans with different nationalities. As a last step, we excluded fake assignments such as @AndyMurray and @DominicThiem by manual verification.

To match accounts and player names, we first listed the accounts that have minimum edit distance from a given player's name. We removed whitespaces and transformed all characters to lower case. Since name matching can lead to false player-account pairs, we manually searched the lists of different edit distance values to find valid player account matches. We first considered screen names, and in case there was no match, we continued with account names.

Using the above semi-automatic procedure, we managed to find Twitter accounts for 58.4% of the US Open and 64.2% of the Roland-Garros players, as seen in Figure 4.

6 UNSUPERVISED EVALUATION

In addition to the data with ground truth of the previous section, we used the data sets of [20] for unsupervised analysis (see Table 1).

	Edges	**Nodes**	**Days**
Students	10,000	1,654	121
Facebook	10,000	4,752	104
Enron	6,251	1,944	892
Tumblr	7,645	1,757	89
UO17	482,061	106,920	21
RG17	336,234	78,095	19

Table 1: Summary of the data sets used.

These small temporal networks (Students, Facebook, Enron, Tumblr) have no more than 10,000 edges[2], as seen in Table 1.

6.1 Stability vs. changeability

We assess the amount of variability of temporal Katz centrality in time, depending on the parameters β and the time decay exponent to exhibit the speed of focus shift in daily interactions. We use the weight function $\varphi(\tau) = \beta \cdot 2^{-c\tau}$; c can be considered as the half-life of the information sent over an edge. We update temporal Katz centrality after each edge arrival, and compute the top 100 nodes with highest centrality scores for each snapshot. We generate the lists at the beginning of each day for the small data sets of [20], and each hour for our Twitter collections RG17 and UO17. Spearman correlation is calculated between lists of adjacent snapshots, for different values of c and β, as shown in Figure 5.

Our measurements show that the similarity between adjacent lists depends on two different factors. We can turn temporal Katz centrality more static by using longer half-life in the decay. If the half-life is short, we even get negative correlations as the number of nodes present in both lists decreases. Another option is to use larger β. By increasing β, the contribution of long walks will be more relevant, which cannot be dominated by recently added edges as easily as for a small β. The two approaches can also be used in combination. We observed the highest similarity using $\beta = 1.0$ with large half-life value.

6.2 Adaptation to concept drift

Rozenshtein et al. [20] showed that temporal PageRank can adapt to the changes in the edge sampling distribution over semi-temporal networks. We conducted similar measurement for temporal Katz centrality on the same data sets.

In our experiment for concept drift adaptation, we randomly selected 500 nodes as a base graph and formed three overlapping subsamples of 400 nodes each. Similar to the approach in [20], we formed a temporal edge stream of three segments corresponding to the three subsamples, in each segment selecting 10,000 random edges from the corresponding subsample. We compute temporal PageRank and temporal Katz centrality by assuming that a new edge in the stream appears in each time unit. In other words, we measure the elapsed time τ by the number of edges in the stream.

We computed weighted Kendall tau [24] rank distance between temporal Katz centrality and static Katz index restricted to the nodes of the actual subsample. By using weighted Kendall tau, we put more emphasis on nodes with high centrality compared to (unweighted)

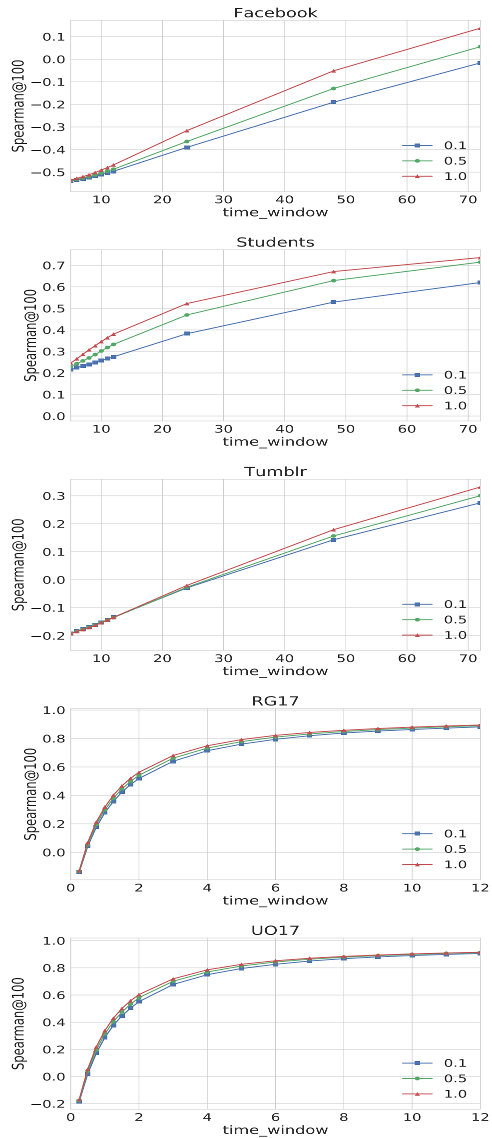


Figure 5: Average Spearman correlation between temporal Katz centrality scores of adjacent snapshots. Daily snapshots are used for Facebook, Students and Tumblr, and hourly for RG17 and UO17. The correlation is presented for β values 0.1, 0.5, 1.0 **and several time decay intensity.**

Kendall tau. For the same reason, we use the asymmetric version as in [24, Section 5.1] by using the weight of 1/rank for the static Katz index and zero for the online methods. By this choice, Kendall tau measures the distance from the Katz index acting as ground truth.

In Figure 6, we evaluated our model for various values of the exponential decay against the Katz index with $\beta = 0.01$. The results show that in case of weak decay $c = \frac{1}{|E|}$, temporal Katz centrality converges to static Katz index. On the contrary, strong decay shifts the focus of temporal centrality towards the recently sampled edges, thus correlation decrease for $c = \frac{10}{|E|}$ and $c = \frac{100}{|E|}$. Temporal Katz

[2]GitHub repository of the temporal PageRank research: https://github.com/polinapolina/temporal-pagerank

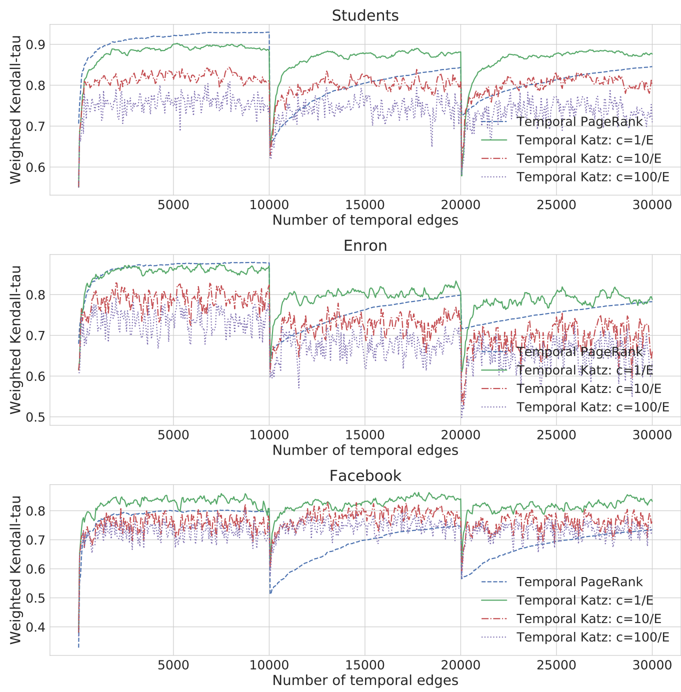


Figure 6: Weighted Kendall tau rank distance of static Katz index and online methods by sampling to simulate concept drift over Students, Enron, and Facebook data. Static Katz index has $\beta = 0.01$. The Weighted Kendall tau curves for temporal Katz centrality with $c = \frac{1}{|E|}$ is green, with $c = \frac{10}{|E|}$ is red, with $c = \frac{100}{|E|}$ is purple, and for temporal PageRank is blue dashed. Noise in temporal Katz centrality is due to the effect of the most recently selected edges.

centrality quickly detects concept drift when we switch between subsamples.

7 SUPERVISED EVALUATION

In this section, we quantitatively analyze the relevance of temporal centrality measures over the UO17 and RG17 Twitter collections. We compare the relevance of temporal Katz centrality to temporal PageRank and other *online* and *static* baseline methods.

To evaluate online metrics, we perform continuous update as the new edges arrive, by considering our data as a time-ordered edge stream. For the static metrics, we consider different graph snapshots. For each centrality measure, we compute the list of the nodes with the highest centrality in *each hour*. We use NDCG [1] for evaluation, with relevance 1 if the owner of the account is a professional player who participated in a game on the given day and 0 otherwise.

7.1 Baseline metrics

We compare temporal Katz centrality to *online* (or time-aware) and *static* (or batch) metrics. Online metrics are updated after the arrival of each edge. By contrast, static metrics are only updated once in each hour. At hour t a static metric is computed on the graph constructed from edges arriving in time window $[t - T, t]$ from the edge stream. For each baseline, we experimentally select the best value of T.

We consider four *static* centrality measures as baseline:

- *PageRank* [19]: We set $\alpha = 0.85$, and 50 iterations.
- *indegree*: We calculate the indegree of each node in time window $[t-T, t]$ by counting each edge once, that is, without multiplicity.
- *negative β-measure* [6]: The normalized version of indegree,

$$\sum_{z \in N_{in}(u)} \frac{1}{\text{outdegree}(z)} \tag{12}$$

where $N_{in}(u)$ denotes the in-neighbors of u.

- *harmonic centrality* [6]: For node u

$$\sum_{z \neq u} \frac{1}{d(z,u)}. \tag{13}$$

Furthermore, we compare temporal Katz centrality with two *online* metrics, temporal PageRank [20] and decayed indegree.

- *temporal PageRank:* We set $\alpha = 0.85$ and $\beta \in \{0.001, 0.01, 0.05, 0.1, 0.5, 0.9\}$ for transition weight.
- *decayed indegree:* The decayed indegree of node u at time t is

$$\sum_{zu \in E(t)} \varphi(t - t_{zu}) \tag{14}$$

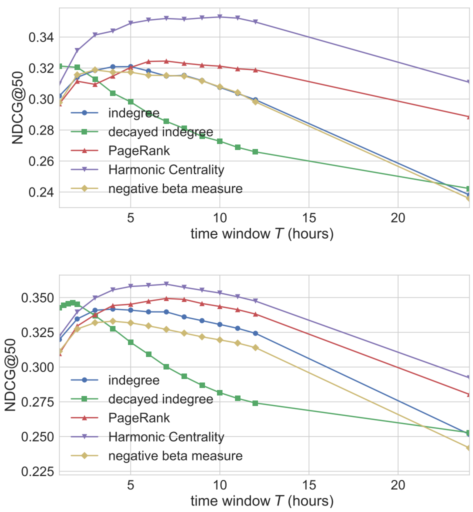
where φ is the time decay function that we set $\varphi(t - t_{zu}) := \exp(-c(t - t_{zu}))$ similarly to temporal Katz centrality.

7.2 Results

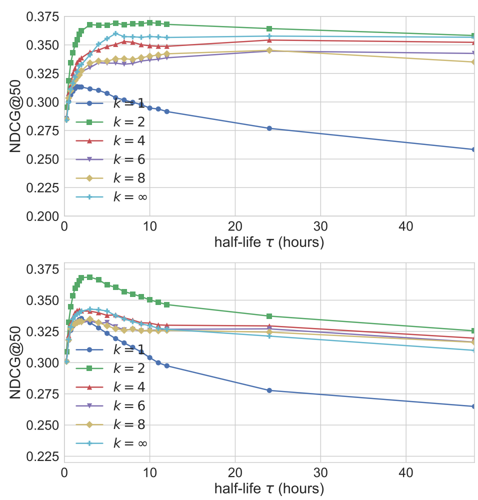
As the final and main analysis of the relevance of centrality measures, we compute hourly lists of top centrality nodes and calculate the NDCG@50 against the ground truth. The hour of the day has a key effect on performance. In the early hours, activity is low, and hence information is scarce to identify the players of the coming day. By contrast, in the late hours after the games are over, we expect that all models easily detect the players of the day based on the tweets of the results. Due to these observations we present two different ways to aggregate hourly NDCG@50 values: (1) For each hour of the day between 0:00 and 24:00, we show averages over the days of the tournament (Figure 8). (2) As a single global value (Figure 7, Table 2), we average NDCG@50 for all days with all hours between 10:00 and 20:00 as daily tennis games start around 10:00. The effect of the hour of the day can be seen in Figure 8, where we plot the average daily performance over UO17.

First, we analyze our baseline models. Each static metric is computed at hour t over the graph defined by edges arriving in time frame $[t - T, t]$. Hence the key parameter of these methods is the length of the time window T. Similarly, online decayed indegree depends on the half-life parameter $\tau := \ln 2/c$. Figure 7, top, shows the overall performance of the static baselines as the function of time frame T, and the quality of decayed indegree as the function of half-life τ. For both data sets, PageRank and harmonic centrality outperform degree-related methods. Furthermore, these path-based methods prefer larger time frames, while degree-based models perform best at smaller values of T.

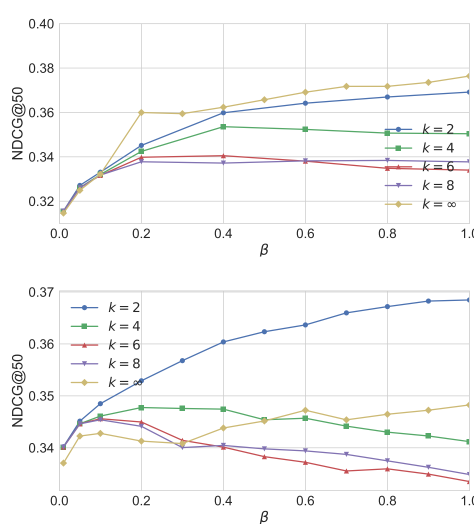
Next we analyze temporal Katz centrality with exponential decay. The key parameters of our method are the parameters of the exponential decay β and $\tau := \ln 2/c$, and truncation k. We parameterize exponential decay with half-life $\tau := \ln 2/c$ instead of c.



Batch methods shown as the function of time window T. Online exponential degree shown as the function of half-life τ.



Temporal Katz centrality with $\beta = 1$ and several truncation values of k shown as the function of half-life τ.



Temporal Katz centrality shown as the function of β, with different values of k and $\tau = 6h$ for UO17, $\tau = 3h$ for RG17.

Figure 7: NDCG@50 performance of metrics with different hyperparameter settings. Top: UO17, Bottom: RG17.

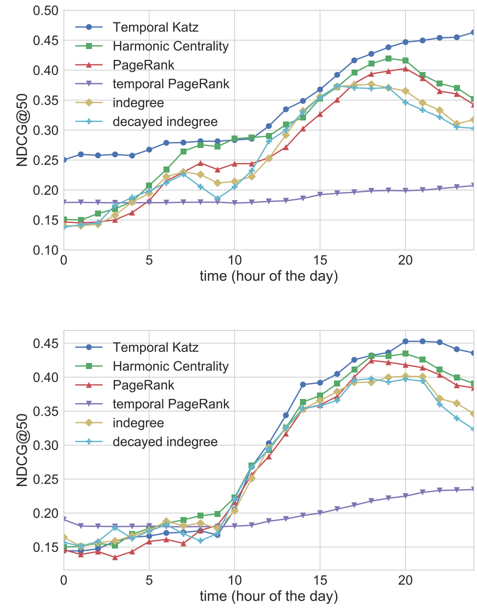


Figure 8: Overall best daily NDCG@50 performance of temporal Katz centrality and the baselines. Top: UO17, Bottom: RG17.

First, we examine the effect of k and half-life τ by setting $\beta = 1$. Figure 7, middle, shows the performance of temporal Katz centrality at various parameter settings for UO17 the RG17. We plot NDCG@50 against parameter τ. Different curves correspond to different k parameters. The effect of k is significant: Models with $k > 1$ strongly outperform models with $k = 1$, a very simple version of temporal Katz centrality similar to online degree. The best performance can be achieved on both data sets by setting $k = 2$ and $\tau \approx 3h$.

In Figure 7, bottom, we analyze the importance of parameter β. For models with larger k (e.g. $k = 8$), the importance of β is to decrease the effect of paths that are too long, with optimal value around $\beta \approx 0.1 - 0.2$. For methods with lower k (e.g. $k = 2$), β is nearly meaningless, and the use of small β in combination with strong exponential decay results in performance deterioration.

As final conclusion, in Figure 8 we compare the hourly performance of each method at their best parameter settings. For temporal Katz centrality we set $\beta = 1, \tau = 3h, k = 2$. In the case of both data sets, temporal Katz centrality can keep up with the performance of harmonic centrality, the strongest baseline model. The quality of temporal PageRank is significantly lower than the quality of other methods. We summarize the best average NDCG@50 scores for temporal Katz centrality and the baselines in Table 2. temporal Katz centrality generally performs better than other baselines. Note that only staic harmonic centrality delivers performance comparable to temporal Katz centrality.

We illustrate various centrality measures by showing the 20 accounts with highest score for the Roland-Garros semifinals. On June 9, more than 70 players participated in several categories (Men's singles, Girl's and Boy's singles, etc.). In Table 3, we show top accounts by various methods. We show the accounts of tennis

NDCG@50	UO17	RG17
indegree	0.321	0.342
decayed indegree	0.321	0.346
negative beta	0.319	0.333
PageRank	0.325	0.349
temporal PageRank	0.187	0.195
harmonic centrality	0.353	0.359
temporal Katz centrality	**0.370**	**0.368**

Table 2: Best average NDCG@50 performances.

1	Simona Halep	@Simona_Halep	0
2	Stanislas Wawrinka	@stanwawrinka	1
3	Andy Murray	@andy_murray	1
4	Rafa Nadal	@RafaelNadal	1
5	Roland-Garros	@rolandgarros	0
6	Ana Ivanovic	@AnaIvanovic	0
7	Timea Bacsinszky	@TimeaOfficial	0
8	Karolina Pliskova	@KaPliskova	0
9	Rohan Bopanna	@rohanbopanna	0
10	Dominic Thiem	@ThiemDomi	1
11	Gaby Dabrowski	@GabyDabrowski	0
12	Gustavo Kuerten	@gugakuerten	0
13	Nicola Kuhn	@NicolaKuhn1	1
14	yonex.com	@yonex_com	0
15	Whitney osuigwe	@whitney_osuigwe	1
16	Caroline Garcia	@CaroGarcia	0
17	NikeCourt	@Nikecourt	0
18	Novak Djokovic	@DjokerNole	0
19	WTA	@WTA	0
20	ATP World Tour	@ATPWorldTour	0

1	Roland-Garros	@rolandgarros	0
2	Stanislas Wawrinka	@stanwawrinka	1
3	Andy Murray	@andy_murray	1
4	Simona Halep	@Simona_Halep	0
5	Rafa Nadal	@RafaelNadal	1
6	Dominic Thiem	@ThiemDomi	1
7	Timea Bacsinszky	@TimeaOfficial	0
8	Rohan Bopanna	@rohanbopanna	0
9	Ana Ivanovic	@AnaIvanovic	0
10	WTA	@WTA	0
11	Gaby Dabrowski	@GabyDabrowski	0
12	Tennis Channel	@TennisChannel	0
13	Rafa Nadal Academy	@rnadalacademy	0
14	Karolina Pliskova	@KaPliskova	0
15	yonex.com	@yonex_com	0
16	Gusti Fernandez	@gustifernandez4	0
17	rolandgarrosFR	@rolandgarros_FR	0
18	Eurosport.es	@Eurosport_ES	0
19	ATP World Tour	@ATPWorldTour	0
20	Caroline Garcia	@CaroGarcia	0

1	Roland-Garros	@rolandgarros	0
2	Rafa Nadal	@RafaelNadal	1
3	Andy Murray	@andy_murray	1
4	Stanislas Wawrinka	@stanwawrinka	1
5	Simona Halep	@Simona_Halep	0
6	Dominic Thiem	@ThiemDomi	1
7	Rohan Bopanna	@rohanbopanna	0
8	Timea Bacsinszky	@TimeaOfficial	0
9	Ana Ivanovic	@AnaIvanovic	0
10	Tennis Channel	@TennisChannel	0
11	yonex.com	@yonex_com	0
12	WTA	@WTA	0
13	Caroline Garcia	@CaroGarcia	0
14	Rafa Nadal Academy	@rnadalacademy	0
15	Gaby Dabrowski	@GabyDabrowski	0
16	ATP World Tour	@ATPWorldTour	0
17	Whitney osuigwe	@whitney_osuigwe	1
18	Nicola Kuhn	@NicolaKuhn1	1
19	NikeCourt	@Nikecourt	0
20	Anabel Medina	@anabelmedina	0

1	Roland-Garros	@rolandgarros	0
2	Andy Murray	@andy_murray	1
3	Stanislas Wawrinka	@stanwawrinka	1
4	Rafa Nadal	@RafaelNadal	1
5	Dominic Thiem	@ThiemDomi	1
6	Timea Bacsinszky	@TimeaOfficial	0
7	Simona Halep	@Simona_Halep	0
8	Rohan Bopanna	@rohanbopanna	0
9	Ana Ivanovic	@AnaIvanovic	0
10	Tennis Channel	@TennisChannel	0
11	Gaby Dabrowski	@GabyDabrowski	0
12	Gusti Fernandez	@gustifernandez4	0
13	Rafa Nadal Academy	@rnadalacademy	0
14	WTA	@WTA	0
15	yonex.com	@yonex_com	0
16	Eurosport.es	@Eurosport_ES	0
17	Caroline Garcia	@CaroGarcia	0
18	Eurosport UK	@Eurosport_UK	0
19	Stéphanie Loire	@Stephloire	0
20	Emilie Lopez	@emilielopez	0

Table 3: Top list for RG17 semi final day (June 9) at 12:00. Top left: temporal Katz centrality $\beta = 1.0$; Top right: $\beta = 0.2$. Bottom left: harmonic centrality; Bottom right: decayed indegree. Relevant daily players are highlighted orange. Accounts of players who did not play on this day are highlighted yellow.

players playing participating in the June 9 semifinals in orange and of those who did not play in yellow, for example, women semifinalists of the previous day, Simona Halep, Timea Bacsinszky, Caroline Garcia and Gabriela Dabrowski. All methods listed 4–6 daily players among the most central 20 accounts. All methods assigned high centrality to Men semi-finalists Nadal, Murray, Wawrinka and Thiem. Furthermore, temporal Katz centrality with $\beta = 1.0$ and harmonic centrality could recover two additional young daily players, Osuigwe and Kuhn. Retired tennis legends Ivanovic and Kuerten are not relevant in our experiment as they did not participate.

Notice that decayed indegree and temporal Katz centrality with $\beta = 0.2$ rank sports media accounts (Tennis Channel, WTA, ATP World Tour, Eurosport) higher compared to harmonic centrality and temporal Katz centrality with $\beta = 1.0$. We did not attempt to curate the relevance to media sources, as the number of such Twitter accounts is abundant. Finally, sponsors 'yonex.com' and 'NikeCourt', as well as the official Twitter account of the event '@rolandgarros' also rank high. Most of these accounts are active every day, with little observable change in time, which justifies why we do not consider them relevant for the temporal evaluation.

8 CONCLUSION

In this paper, we designed an online updateable, dynamic graph centrality measure based on the Katz index. We presented multiple unsupervised experiments to show that our method can adapt to changes in the distribution of the edge stream. We compiled a supervised evaluation for the mention graphs of Twitter tennis tournament collections along with temporal importance ground truth information. To the best of our knowledge, these are the first Twitter collections enhanced with dynamic node importance labels. Our data and codes are publicly available.

REFERENCES

[1] Azzah Al-Maskari, Mark Sanderson, and Paul Clough. 2007. The relationship between IR effectiveness measures and user satisfaction. In *Proc SIGIR*. ACM, 773–774.

[2] Ahmad Alsayed and Desmond J Higham. 2015. Betweenness in time dependent networks. *Chaos, Solitons & Fractals* 72 (2015), 35–48.

[3] Bahman Bahmani, Abdur Chowdhury, and Ashish Goel. 2010. Fast incremental and personalized pagerank. *Proc. VLDB* 4, 3 (2010), 173–184.

[4] Bahman Bahmani, Ravi Kumar, Mohammad Mahdian, and Eli Upfal. 2012. Pagerank on an evolving graph. In *Proc. SIGKDD*. ACM, 24–32.

[5] Albert Bifet, Geoff Holmes, Richard Kirkby, and Bernhard Pfahringer. 2010. Moa: Massive online analysis. *JMLR* 11, May (2010), 1601–1604.

[6] Paolo Boldi and Sebastiano Vigna. 2014. Axioms for centrality. *Internet Mathematics* 10, 3-4 (2014), 222–262.

[7] Dan Braha and Yaneer Bar-Yam. 2006. From centrality to temporary fame: Dynamic centrality in complex networks. *Complexity* 12, 2 (2006), 59–63.

[8] Dániel Fogaras, Balázs Rácz, Károly Csalogány, and Tamás Sarlós. 2005. Towards Scaling Fully Personalized PageRank: Algorithms, Lower Bounds, and Experiments. *Internet Mathematics* 2, 3 (2005), 333–358.

[9] Peter Grindrod and Desmond J Higham. 2014. A dynamical systems view of network centrality. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 470.

[10] Scott A Hill and Dan Braha. 2010. Dynamic model of time-dependent complex networks. *Physical Review E* 82, 4 (2010), 046105.

[11] Weishu Hu, Haitao Zou, and Zhiguo Gong. 2015. temporal PageRank on social networks. In *Int. Conf. Web Information Systems Engineering*. Springer, 262–276.

[12] Leo Katz. 1953. A new status index derived from sociometric analysis. *Psychometrika* 18, 1 (1953), 39–43.

[13] Hyoungshick Kim and Ross Anderson. 2012. Temporal node centrality in complex networks. *Physical Review E* 85, 2 (2012), 026107.

[14] Kyung Soo Kim and Yong Suk Choi. 2015. Incremental iteration method for fast PageRank computation. In *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*. ACM, 80.

[15] Kristina Lerman, Rumi Ghosh, and Jeon Hyung Kang. 2010. Centrality metric for dynamic networks. In *Proc. MLG 2010*. ACM, 70–77.

[16] Bundit Manaskasemsak, Pramote Teerasetmanakul, Kankamol Tongtip, Athasit Surarerks, and Arnon Rungsawang. 2013. Computing personalized PageRank based on temporal-biased proximity. In *Information Technology Convergence*. Springer, 375–385.

[17] Michael Mitzenmacher. 2004. A brief history of generative models for power law and lognormal distributions. *Internet mathematics* 1, 2 (2004), 226–251.

[18] Naoto Ohsaka, Takanori Maehara, and Ken-ichi Kawarabayashi. 2015. Efficient PageRank tracking in evolving networks. In *Proc. SIGKDD*. ACM, 875–884.

[19] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1998. *The PageRank Citation Ranking: Bringing Order to the Web.* Technical Report 1999-66. Stanford University. citeseer.nj.nec.com/page98pagerank.html

[20] Polina Rozenshtein and Aristides Gionis. 2016. temporal PageRank. In *Proc. ECML*. Springer, 674–689.

[21] Atish Das Sarma, Sreenivas Gollapudi, and Rina Panigrahy. 2011. Estimating pagerank on graph streams. *Journal of the ACM (JACM)* 58, 3 (2011), 13.

[22] John Tang, Mirco Musolesi, Cecilia Mascolo, Vito Latora, and Vincenzo Nicosia. 2010. Analysing information flows and key mediators through temporal centrality metrics. In *Proceedings of the 3rd Workshop on Social Network Systems*. ACM, 3.

[23] Dane Taylor, Sean A Myers, Aaron Clauset, Mason A Porter, and Peter J Mucha. 2017. Eigenvector-based centrality measures for temporal networks. *Multiscale Modeling & Simulation* 15, 1 (2017), 537–574.

[24] Sebastiano Vigna. 2015. A weighted correlation index for rankings with ties. In *Proc. WWW*. 1166–1176.