

On Using Temporal Networks to Analyze User Preferences Dynamics

Fabíola S. F. Pereira¹, Sandra de Amo¹ and João Gama²

¹ Federal University of Uberlândia, Uberlândia, Brazil
`{fabiola.pereira,deamo}@ufu.br`

² University of Porto, LIAAD INESC TEC, Porto, Portugal
`jgama@fep.up.pt`

Abstract. User preferences are fairly dynamic, since users tend to exploit a wide range of information and modify their tastes accordingly over time. Existing models and formulations are too constrained to capture the complexity of this underlying phenomenon. In this paper, we investigate the interplay between user preferences and social networks over time. We propose to analyze user preferences dynamics with his/her social network modeled as a temporal network. First, we define a temporal preference model for reasoning with preferences. Then, we use evolving centralities from temporal networks to link with preferences dynamics. Our results indicate that modeling Twitter as a temporal network is more appropriated for analyzing user preferences dynamics than using just snapshots of static network.

Keywords: temporal networks, user preferences, evolving centralities

1 Introduction

What drives people’s preferences dynamics? Modeling users’ preferences and needs is one of the most important personalization tasks in the information retrieval domain. User preferences are fairly dynamic, since users tend to exploit a wide range of items and modify their tastes accordingly over time. Moreover, all the time users are facing others’ opinions and being socially influenced. This scenario dispatches several research efforts to investigate the interplay between *user preferences* and *social networks* [1,2].

In this paper, we are interested in user preferences dynamics, i.e., the observation of how a user evolves his/her preferences over time. In our context, user preference is a specific type of opinion, that establishes an order relation between two objects. For example, when a user says: “I prefer sports than religion”, we clearly identify his preference to sports subjects over religion.

Although social networks are fairly dynamic, traditional approaches in online social networks analysis consider nodes and edges as being persistent over time [3]. We go a step beyond and discuss social networks as *temporal networks*, where the times when edges are active are an explicit element of the representation [4]. A classical example of temporal networks is on disease contagion. Usually, the

spreading of diseases occurs through contact between two people. Considering the order of these contacts is important and more meaningful than just analyzing an aggregate static network that ignores when these contacts occurred. In the context of social networks, our topic of interest, temporal networks are being used to represent users interactions [5,6].

We hypothesize that centralities properties of temporal networks can reveal interesting patterns of users' preferences dynamics, especially when handling change detection. The running example below illustrates the preference dynamics problem we are investigating in this paper and its link with temporal networks.

Motivating example. Let us consider the most popular microblogging social network, Twitter. In Twitter, a user A usually follows another B if A is interested in what B posts. A common behavior is, for example, if A is fan of soccer, probably he is following sports channels and personalities like Messi and Neymar. Let us imagine, now, the context of an e-commerce recommendation system able to infer user preferences from Twitter in order to improve recommendations quality. In Figure 1(a) we have a sequence of preferences inferred for user A about different themes (sport, TV, religion and music) that he likes to follow, post, share or read about in Twitter. An edge (a_1, a_2) indicates that a_1 is preferred to a_2 (edges inferred by transitivity are not represented in the Figure). Analyzing A 's preferences, we notice that in the first week of January, his preferences keep stable, just appearing a preference of *TV* over *music* themes and disappearing a preference of *sport* over *music*. However, on day 9, A changed his mind and is preferring music over TV. This can be an indicative of *changes* on his preferences and interests.

Now, in Figure 1(b) we have snapshots of A 's Twitter network. One edge from node A to node B indicates that A is followed by B (the flow of information). It is well know that Twitter network is fairly dynamic, and people start and end relationships (follows) all the time [7]. According to our hypothesis, many aspects of A 's network structure may have influenced on his preferences change. For instance:

- the structural *position* of A in the network can be changed, as well as the global network topology. This means that he could be held a close position to music influential users;
- from second week of January, according to A 's *position*, he is not receiving any information about religion;
- A 's *position* is always close to sport celebrities.

It is an essential point for our hypothetical recommendation system to detect and predict A 's preferences evolution over time and their changes. This could improve the sales of CDs, for example, when recommending to A . We show that the temporal-topological Twitter structure is strongly correlated with user preferences dynamics.

Contributions. The main contributions of this paper can be summarized as follows: (1) proposal of a temporal preference model for representing and reasoning with preferences over time; (2) a preference change detection algorithm

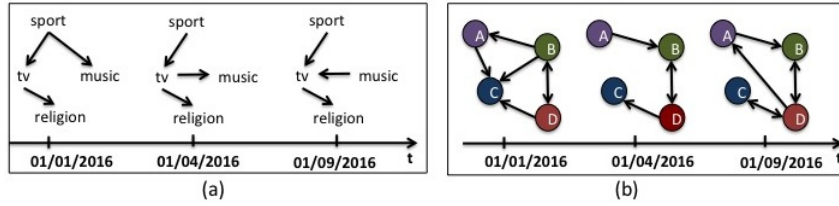


Fig. 1. (a) Sequence of A 's preferences. (b) Snapshots from A 's Twitter network

for detecting changes events in preferences; (3) a centrality change detection algorithm; (4) proposal of a correlation between preferences and centrality measures in temporal networks; (5) a set of experiments validating our proposals, specially against static networks counterpart.

Organization of the paper. First, we discuss some related work. Then, we formalize user preferences dynamics by proposing a temporal preference model able to detect changes on preferences. After that, we present social temporal networks, discussing temporal aspects that were adopted in our analysis. In the Methodology Section, we mainly describe how we elicited preferences from Twitter dataset to validate our proposal. Then, we present the experimental results and conclude the paper.

2 Related Work

We summarize related literature according to two different aspects that we are addressing in this paper: (1) user preferences on social networks and (2) temporal networks.

2.1 User Preferences on Social Networks and Dynamics

Social networks are playing an important role as a type of source knowledge for mining (or inferring) user preferences in the Preference Learning area. The work [1], for example, combines user attributes and social network ties to discover user preferences. In [2], social influence has been used on classifiers to predict users' preferences. In this paper, we go a step beyond, from the point that preference relations between objects are already known and social networks are not used on the learning process, but on the new Preference Dynamics field. The work [8] investigates the music listening histories of Last.fm users focusing on the changes in their preferences based on their choices for different artists at different points in time. The modeling has been done by the survival analysis statistical method. In [9] the focus is on preferences shift (or change) over time. The authors propose a new measure of user preference dynamics (UPD) that captures the shifting rate of preferences. The work [10] is orthogonal to ours because it combines the growth of social links with the generation of user interaction events on those

links. The focus is on social interactions, while ours is on user preferences. The topic of preference change has also been studied on sociology field [11,12].

2.2 Temporal Networks

The literature of temporal networks, focused on social media, is essentially concentrated on understanding patterns of information diffusion [13,14]. Especially, the ideas presented in [5] motivated us to develop this research. The author reviews methods to analyze temporal networks that seek to identify key mediators and how temporal and topological structure of interaction affects spreading processes. In the same way, [15] analyzes temporal centrality metrics. However, our focus is not on information diffusion, but on user preferences dynamics.

Researches on graph metrics for temporal networks essentially address this issue. [16,6] discuss that there are various concepts of *shortest* path for temporal graphs and propose efficient algorithms to compute them. In [17] graph metrics are revisited for temporal networks in order to take into account the effects of time ordering on causality.

Remarking on detecting changes on networks, according to [18], the concept of *change* in a graph sequence falls into one of three main categories according to the structure level of interest: global structure, local structure and the community-level structure. Our focus is on local changes. Global and communities' changes do not hold in our experiments.

3 User Preferences Dynamics

There is no consensus on the definition of preferences dynamics [19]. We adopt the following definition:

Definition 1 (User Preferences Dynamics (UPD)). *UPD refer to the observation of how a user evolves his/her preferences over time.*

First, we propose a temporal preference model able to represent and reasoning with preferences on time. Then, we describe how to detect events of changes on temporal preferences.

3.1 Temporal Preference Model

A *preference* is an *order relation* between two objects. For example, when a user says: “*I prefer sports to politics*”, if we order *sports* and *politics* in a ranking, we can clearly identify that *sports* will be in the top position.

Definition 2 (Temporal Preference Relation \succ_t). *A temporal preference relation (or temporal preference, for short) on a finite set of objects $A = \{a_1, a_2, \dots, a_n\}$ is a strict partial order over A inferred on time t , i.e., a binary relation $R \subseteq A \times A$ satisfying the irreflexivity and transitivity properties on t . Typically, a strict partial order is represented by the symbol \succ . Considering \succ_t as a temporal preference relation, we denote by $a_1 \succ_t a_2$ the fact that a_1 is preferred to a_2 on t .*

Definition 3 (User Temporal Profile Γ_t^u). *User temporal profile Γ_t^u is the transitive closure (TC) of all temporal preferences of user u on t .*

Example 1. Let $A = \{sport, tv, religion, music\}$ be the set of objects in our running domain representing themes of interest of user A . Figure 1(a) illustrates the temporal preferences of A on days 1, 4 and 9 through better-than graphs. Remark that an edge (a_1, a_2) indicates that a_1 is preferred to a_2 and edges inferred by transitivity are not represented. We have: $\Gamma_1^A = \{sport \succ_1 tv, tv \succ_1 religion, sport \succ_1 religion, sport \succ_1 music\}$, $\Gamma_4^A = \{sport \succ_4 tv, tv \succ_4 religion, sport \succ_4 religion, tv \succ_4 music, sport \succ_4 music\}$ and $\Gamma_9^A = \{sport \succ_9 tv, tv \succ_9 religion, sport \succ_9 religion, music \succ_9 tv, music \succ_9 religion\}$.

3.2 Detecting Changes on Temporal Preferences

A key property of temporal preferences is the irreflexivity. We say that a temporal profile Γ_t^u is *inconsistent* when there is a preference $a_1 \succ_t a_1 \in \Gamma_t^u$. It would mean that “*I prefer X better than X!*”, which does not hold for a strict partial order.

Our proposal for detecting preference change is based on the *consistency* of user temporal profiles. The idea is to compute the union of user profiles collected over time, infer temporal preferences by transitivity considering all timestamps and verify if there is any inconsistency on the resulted set of preferences. If yes, we detect an event of preference change. These concepts are formalized in what follows.

Definition 4 (Temporal Profile Union Ω_t^u). *Two temporal preferences of the type $a_1 \succ_{t-1} a_2$ and $a_2 \succ_t a_3$, can unite to infer a third temporal preference $a_1 \succ_{t'} a_3$, once considering transitivity of both, temporal preference relation and timestamp order. A temporal profile union Ω_t^u is the transitive closure (TC) of all irreflexive relations given by $\Gamma_{t-1}^u \cup \Gamma_t^u$.*

Definition 5 (Preference Change δ_t^u). *If there is a temporal preference inconsistency in Ω_t^u a preference change has been detected on time t for user u . In other words, a preference change δ_t^u is defined as:*

$$\delta_t^u = \begin{cases} 1, & \text{if there is a temporal preference inconsistency in } \Omega_t^u \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Remarking on Example 1, let us consider $W = \{1, 4, 9\}$ the set of intervals time stamps. The temporal profile union $\Omega_9^A = \{\dots, tv \succ_4 music, music \succ_9 tv, tv \succ_9 tv, \dots\}$ contains the inconsistency $tv \succ_9 tv$. So, a preference change has been detected on time 9 ($\delta_9^A = 1$). Intuitively, we have that on day 1, for example, A prefers to read/post/share on his social network news about *sport*, but between *tv* and *religion* he is in the mood for *tv*. On the following days, A 's preferences practically do not change, just appearing a preference of *tv* over

music. However, on day 9, *A*'s presented a preference change, as *music* became preferred over *tv*.

The size of the intervals time stamps in W determines if we are tracking short-term or long-term preference events. As example of real events, we can cite new product releases and special personal occasions such as birthdays [20].

In order to formalize the detection of changes in our temporal preference model, we propose the *PrefChangeDetection* algorithm. The intuition of this algorithm is to analyze *better-than graphs* of a user during the observation intervals in W from social networks. If the resulting graph has at least one cycle (meaning an inconsistency) we have detected a preference change. The Algorithm 1 formalizes this idea.

Algorithm 1 PrefChangeDetection

Input: User u , set of intervals W , a vector Γ^u of size $|W|$ containing u 's temporal profiles for each $t \in W$ extracted from the social network $G = (V, E)$
Output: A vector δ^u of size $|W|$ containing u 's preference changes for each $t \in W$

```

1:  $BG_{prev}^u \leftarrow \emptyset$ 
2: for all  $t \in W$  do
3:   build better-than graph  $BG_t^u$  from  $\Gamma^u[t]$ 
4:    $BG^u \leftarrow BG_{prev}^u \cup BG_t^u$ 
5:   if  $BG^u$  is not acyclic then
6:      $\delta^u[t] = 1$ 
7:   else
8:      $\delta^u[t] = 0$ 
9:    $BG_{prev}^u \leftarrow BG_t^u$ 
10: return  $\delta^u$ 

```

Remarking on the complexity analysis, the time to build a better-than graph (line 3) varies according to the preference mining algorithm used. In Section 5 we present the algorithm we use in this paper for mining preferences from social networks. In worst case, its complexity is $O(|V|)$, where $|V|$ is the number of nodes. The time to detect if a directed graph is acyclic (line 5) is $O(|A| + |\Omega_t^u|)$ where A is the set of objects in the domain (the nodes) and Ω_t^u is the temporal profile union containing the preference orders (the edges). Hence, *PrefChangeDetection*, in the worst case, has complexity of $O(|W||V|(|A| + |\Omega_t^u|))$.

4 Social Networks Evolution

We introduce the background of temporal networks, leveraging terms like temporal networks, temporal graphs, static networks and evolving graphs. Next, we propose an algorithm for detecting changes on centralities metrics.

4.1 Temporal Networks vs Static Networks

In this paper, we explore two different representation of social networks: as a static graph structure and as a temporal graph [4,6]. The static graph structure is a traditional approach where the temporal aspects are aggregated and the evolution is analyzed just as a set of graphs snapshots over time [6]. On the other hand, in temporal graphs (or temporal networks) the information of when interactions between nodes happen is taken into account. Let us formalize these concepts.

Definition 6 (Static Networks). A network $G_s = (V, E)$ is static (or aggregate) if there is not any time reference in the edges.

Definition 7 (Temporal Networks). Temporal networks or temporal graphs $G_t = (V, E)$ are graphs with temporal edges, i.e., each edge contains the information of when it has been created and when it has been deleted [6].

Example 2. Consider the temporal and the aggregate graphs in Figure 2. They represent a social network, where the nodes are users and the edges are interactions (for example, tweets) between two users. Suppose that node A has a high impact information to spread in the network. If we analyze the network from the aggregate graph perspective, the information will reach node F . This is not true for the temporal network, as A just interacts on time t_3 with B and after that, it is not possible to reach F from B .

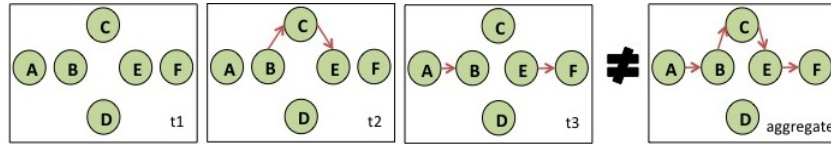


Fig. 2. Temporal network vs. static (aggregate) network

The centrality metrics analysis is inherent to what network representation we are using. Remarking on Example 2, there is a path between nodes A and F in aggregate graph, but not in temporal graphs. This implies in different values of centralities for these nodes. The problem of evolving centralities in temporal networks is addressed in [6]. In Section 6 we show that the *betweenness* and *closeness* centralities have different behaviors according to the network representation and, consequently, they correlate with user preferences in different ways.

4.2 Detecting Changes on Centrality Metrics

In order to detect changes on temporal metrics, and consequently on graph structure, we defined a baseline approach founded on change-point in rankings

[21]. The idea is to maintain a ranking R of all the nodes in the graph according to their metrics values for each interval time stamp t_i inside the observation set of intervals W . Based on the variations of metrics values and ranking positions from t to $t + 1$, we detect changes.

Definition 8 (Ranking Position). Let $G = (V, E)$ be a graph and $W = \{t, t + 1, \dots\}$ a set of intervals. We define C_t^u as the centrality value of u on time t , for $u \in V$ and $t \in W$. Let us consider a ranking R_t where the centrality values of all nodes in V are ranked in descending order on time t . We define pos_t^u as the position of node u in R_t , i.e., $C_t^u > C_t^v$ iff $pos_t^u > pos_t^v$, for $u, v \in V$.

Definition 9 (Temporal Metric Change λ_t^u). We define Λ_t^u as the acceleration of node u in centrality ranking position from time $t - 1$ to time t :

$$\Lambda_t^u = \frac{|pos_t^u - pos_{t-1}^u|}{\max(pos_t^u, pos_{t-1}^u)} \quad (2)$$

A temporal metric change λ_t^u is detected when Λ_t^u is greater than a threshold θ :

$$\lambda_t^u = \begin{cases} 1, & \Lambda_t^u > \theta \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The algorithm *CentralityChangeDetection* (Algorithm 2) implements change detection in centrality metrics based on above definitions. The complexity time is given by the computation of centralities (line 3). In the worst case, the betweenness centrality involves calculating the shortest paths between all pairs of vertices on a graph, which takes $O(|V|^3)$. Hence, the total complexity is $O(|W||V|^3)$.

Algorithm 2 CentralityChangeDetection

Input: User u , set of intervals W , social network $G = (V, E)$, threshold θ

Output: A vector λ^u of size $|W|$ containing temporal metrics changes for each $t \in W$

```

1:  $R_{t-1} \leftarrow \emptyset$ 
2: for all  $t \in W$  do
3:   calculate centralities from  $G$  and build ranking  $R_t$ 
4:   if  $R_{t-1} = \emptyset$  then
5:      $\Lambda_t^u = 0$ 
6:   else
7:      $\Lambda_t^u \leftarrow \frac{|pos_t^u - pos_{t-1}^u|}{\max(pos_t^u, pos_{t-1}^u)}$ 
8:   if  $\Lambda_t^u > \theta$  then
9:      $\lambda^u[t] = 1$ 
10:  else
11:     $\lambda^u[t] = 0$ 
12:   $R_{t-1} \leftarrow R_t$ 
13: return  $\lambda^u$ 

```

5 Methodology

In our methodology we chose as temporal social network the Twitter follower/followee network and extract preferences based on the structure of this network.

5.1 Dataset

The Twitter dataset from [6] was used to validate our proposal. In order to correlate user preferences and the structure of social networks, from evolution viewpoint, we need a dataset (1) containing the information of *when* relationships start and end in the network and (2) some semantic information about the nodes from which it is possible to extract preferences.

The temporal information of Twitter has the following meaning: each node is a user and an edge $(u, v, t_{init}, t_{end})$ indicates that v starts following u at t_{init} and unfollows u at $t_{end}+1$ (v follows u during $[t_{init}, t_{end}]$). As we are dealing with a real dynamic social network, a user can follow and unfollow another user all the time. This is the most interesting aspect that we are investigating: how following relationships on Twitter can allow us to understand user preferences dynamics? The dataset contains 144975 users and 1222118 temporal edges, observed from 08/28/2015 to 12/15/2015. In [6] there are more insights about time-changing characteristics of data.

5.2 Preference Mining

The dataset we used was crawled from Twitter based on a “seed celebrity” policy: choose s celebrities users as seeds and for each seed, select his/her followers. These seeds play an important role for the extraction of users preferences from data. We labeled the 27 seeds based on 9 themes that they represent as celebrities. The themes are *politics*, *sport*, *religion*, *news*, *music*, *humor*, *TV*, *fashion* and *health*. For example, Neymar is a sport celebrity, Gisele Bündchen from fashion and the pope is a religion representative.

The 9 themes adopted are the domain of preferences. The intuition in this preference mining process is: if user u follows a lot of religions personalities and does not follow anyone from fashion field, then u has more interest in religion than in fashion. Thus, we mined preferences of the type: $religion \succ_u^t fashion$.

We consider the preference strength w as the number of seeds of the same theme that a user u follows. For example, if u follows 3 *religion* seeds and 1 *news* seed, we have $religion \succ_u^t news$. If u does not follow any *health* seed, we have that the remaining themes are preferred over health. These situations are illustrated in Figure 3. This method is solely language-agnostic and based on the semantics inferred by the structure of the network [22].

There are many drawbacks on mining preferences in this way. We can infer that a user prefer a topic even if the posts are badly written. Or, that soccer players always post about soccer, which is not necessary true. As this paper is a pioneer quantitative analysis of preferences dynamics in Twitter, we chose this preference mining approach as baseline. In [23] a technique to extract preferences

from tweets has been proposed, but we did not use it because our dataset do not contain users tweets. We just have the network topology.

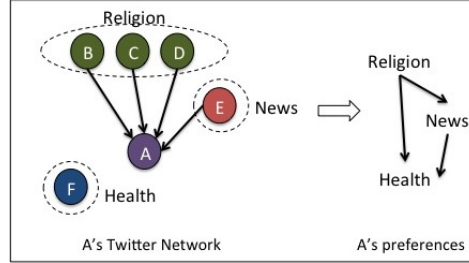


Fig. 3. Mining preferences from Twitter

5.3 Discussion

There are many directions to explore from the concepts presented in this paper: (i) networks are related to preferences to what extent? (ii) Any social network can be used to analyze UPD? (iii) What are the best centrality metrics? (iv) What is the best network modeling to analyze UPD: static or temporal? In this paper, we perform experiments to validate the direction (iv).

We analyze how effective are temporal networks to track user preferences dynamics against static networks in Twitter dataset. The analysis is founded on the correlation strength of preferences changes and centralities metrics changes over time. According to our methodology, the preferences are elicited from the same graph structure that we perform centralities measures analysis. Hence, preference change and centrality change are naturally correlated. Our focus here is to show that there is a significant difference between correlations strength obtained with temporal networks representation in relation to static networks counterpart. So, the methodology does not imply a bias in the experiments.

6 Experimental Results

We show that temporal networks are better representations than static networks for the analysis of user preferences dynamics.

6.1 Experimental Environment

Intervals. The solutions we are proposing for the problem of preferences and centralities events detection are highly sensitive to the granularity of observation window W . We define four intervals to perform the experiments, described in Table 1. Remark that the intervals are the elements of W . If we are interested in

tracking short-term events, then short intervals like Daily and Weekly fit better. For instance, preferences over the domains of news or restaurants have a high changing rate. On the other hand, long intervals are more appropriate when the events are not frequent, for example preferences about movies and politics.

Period	# of intervals	Values
Daily (DA)	110	$DA_1 = [08/28, 08/28], \dots, DA_{110} = [12/15, 12/15]$
Weekly (WE)	15	$WE_1 = [09/01, 09/07], WE_2 = [09/08, 09/14], \dots, WE_{15} = [12/08, 12/14]$
Fortnightly (FO)	7	$FO_1 = [09/01, 09/15], FO_2 = [09/16, 09/30], \dots, FO_7 = [11/30, 12/14]$
Monthly (MO)	3	$MO_1 = [09/01, 09/30], MO_2 = [10/01, 10/31], MO_3 = [11/01, 11/30]$

Table 1. Intervals inside observation window $W = [08/28/2015, 12/15/2015]$

Users. The dataset contains 144975 users. For computing the events along time intervals, the values λ_t^{AVG} and δ_t^{AVG} correspond to the average across all users for each time interval t .

Centrality Metrics. *Betweenness* and *closeness* metrics are used in our experiments. Betweenness considers how important nodes are in connecting other nodes. In closeness centrality the intuition is that the more central nodes are, the more quickly they can reach other nodes [24]. Thus, these metrics are related with nodes that play *influence* and *spreading* roles in the network, respectively. These are potential features for understanding users preferences dynamics. For centrality events detection we vary the threshold $\theta = \{0.2, 0.4, 0.6\}$.

Social Network Representation. We compare the behavior of *temporal graphs* and *static graphs* structures in relation to preferences and centralities changes over time.

6.2 Analyzing Change Events and Correlations

Q1: What are the changes behaviors of users (preferences) and nodes (centralities)? Are these variables really dynamics?

We verify the change rate for both preferences and centralities over different time intervals. The results are illustrated in Figure 4. The analysis of daily intervals is not interesting here as there is no difference between temporal and static networks for daily intervals (the temporal network has granularity of one day). The parameter θ has been fixed as 0.4 in this analysis, corresponding to the intermediate value of our range. In next analysis we show that it does not affect the correlations behaviors.

On average, the rate of users that change their preferences are 30.39%, 32.17% and 40.11% for week, fortnight and month intervals, respectively. For temporal

betweenness centrality, the averages of changes are 43.56%, 46.44% and 46.39%; and 36.42%, 35.07%, 45.01% for static betweenness. Temporal closeness changes averages are 38.85%, 38.78% and 41.23%; and 24.16%, 26.28% and 33.78% for static closeness. Generally, as the interval size increases, the change rate increases as well. This occurs due to the trade off between domain of preferences and social network. We are investigating the domain of users' preferences to post/share/read in Twitter. According to our analysis, this is very dynamic and even short-term data (week intervals) have relevancy.

The most important observation is that the curves of preferences and temporal centralities have similar behaviors for all scenarios, different from static centralities. This observation indicates that if we consider temporal modeling of Twitter network to track evolving betweenness and closeness nodes centralities, we have a better notion of users respective preferences dynamics than if considering static modeling.

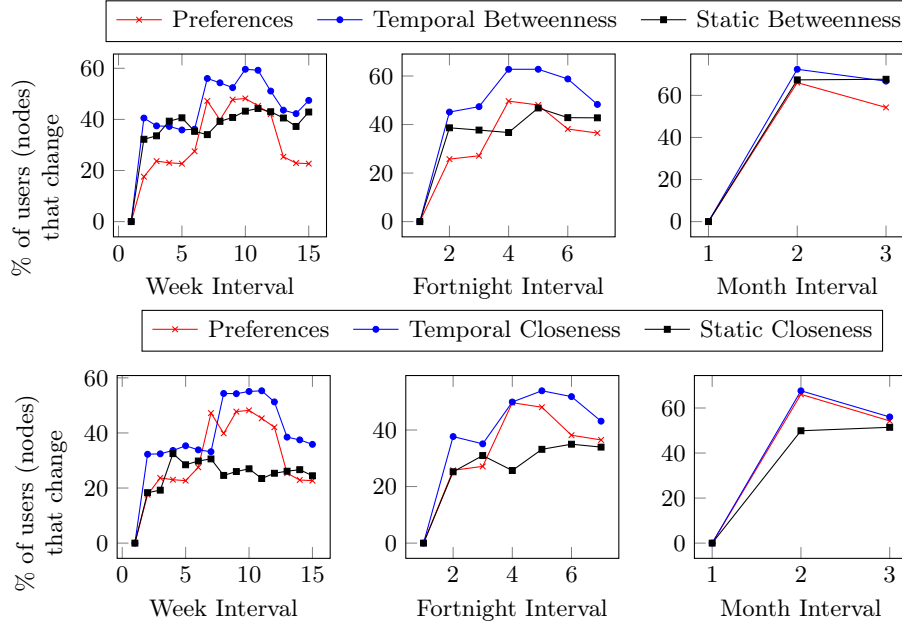


Fig. 4. Change rates for betweenness (up) and closeness (down) centralities

Q2: Are temporal networks more appropriate to analyze user preferences dynamics than static networks?

We use Pearson Correlation Coefficient (PCC) metric to evaluate which network model better represents preferences dynamics. As previously appointed, we mine *preferences* from the same graph structure that we perform *centralities* measures analysis. It is expected that both variables are correlated. As a matter

of fact, in this analysis, we explore the *correlation strength* difference between temporal and static networks representations in our context.

The results are illustrated in Figure 5. $PCC(\delta_t^{AVG}, \lambda_t^{AVG})$ has been calculated considering betweenness and closeness centralities. Each scenario has three periods – Day, Week and Fortnight. The month intervals are not illustrated due the small size of the series (only three values). For each centrality we vary the parameter θ . This parameter indicates that the closer to 1, more significant are the centralities changes that are being considered.

Both centralities metrics correlate significantly (as compared to the corresponding critical values – in all scenarios critical values are lower than 0.1). Two random variables (with no correlation) would have a 95% probability of PCC greater than a critical value or lower. As expected, we observe a high correlation between the change events in user preferences and in centrality metrics.

We highlight the difference between temporal and static values. Considering the analyzed scenarios, on average, temporal betweenness has a correlation strength 40 times higher than static betweenness. For closeness, the correlation strength is 59 times higher. This corroborates our investigation that changes in temporal metrics indicate changes on users preferences. The correlation difference between temporal and static metrics is an evidence that dynamics of preferences fit better in temporal networks representation.

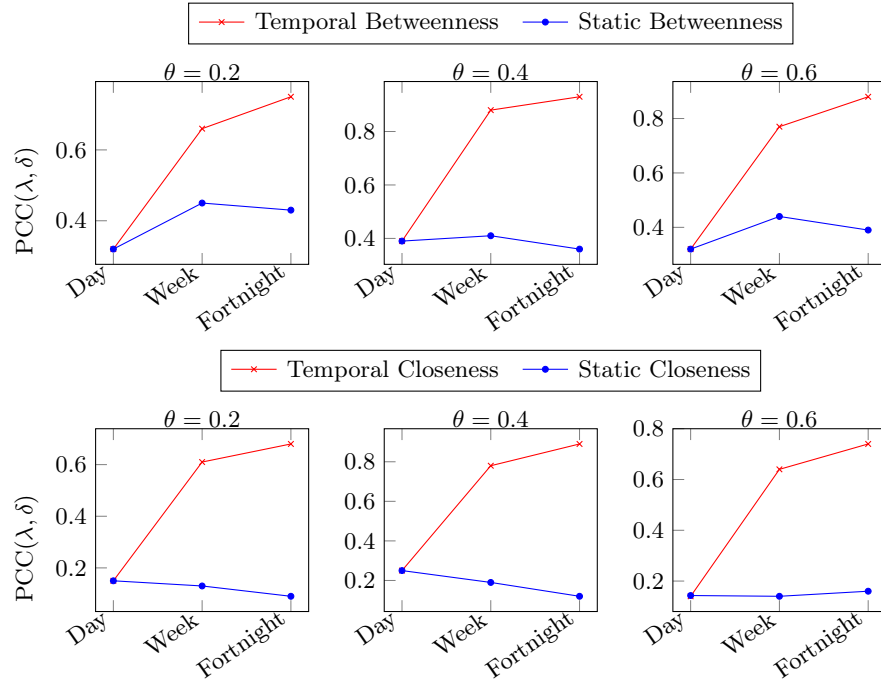


Fig. 5. PCC between centralities and preferences

Q3: Why temporal networks are good social network representation in the analysis of user preferences dynamics?

The results obtained so far can be explained by the phenomena of information propagation and inherent consequences of homophily and influence. The main difference between temporal and static networks, as discussed in Section 4, is that temporal networks take into account the contact sequence between nodes to compute paths [6] and this has an impact on different centralities measures. The related work [14] discuss about the relation of preferences and information propagation in social networks. The aspects described on motivating example (Section 1) could illustrate that preferences are directed by information flow in the social network. Finally, temporal networks represent information flow more realistically.

7 Conclusion

In this paper we have investigated the interplay between *user preferences dynamics* and *evolving social networks*. We have introduced a new temporal preference model able to describe dynamics of user preferences through user profiles and change detection. We have presented a social network analysis correlating centralities metrics evolution with user preferences dynamics over Twitter data. Our findings have shown that there is a high correlation between changes on temporal centrality metrics – *betweenness* and *closeness* – and changes on user preferences, against static centrality metrics counterpart.

Acknowledgments This work was supported by the research project “TEC4Growth - Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact / NORTE-01-0145-FEDER-000020”, financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF) and by European Commission through the project MAESTRA (Grant number ICT-2013-612944). Fabiola Pereira is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalization - COMPETE 2020 Programme within project POCI-01-0145-FEDER-006961, and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) as part of project UID/EEA/50014/2013. This work was also supported by the Brazilian Research Agencies CAPES and CNPq.

References

1. Li, J., Ritter, A., Jurafsky, D.: Inferring user preferences by probabilistic logical reasoning over social networks. arXiv preprint arXiv:1411.2679 (2014)
2. Abbasi, M.A., Tang, J., Liu, H.: Scalable learning of users’ preferences using networked data. In: Proceedings of the 25th ACM conference on Hypertext and social media, ACM (2014) 4–12
3. Costa, L.d.F., Rodrigues, F.A., Traverso, G., Villas Boas, P.R.: Characterization of complex networks: A survey of measurements. *Advances in Physics* **56**(1) (2007) 167–242

4. Holme, P., Saramäki, J.: Temporal networks. *Physics reports* **519**(3) (2012) 97–125
5. Holme, P.: Analyzing temporal networks in social media. *Proceedings of the IEEE* **102**(12) (2014) 1922–1933
6. Pereira, F.S.F., Amo, S., Gama, J.: Evolving centralities in temporal graphs: a twitter network analysis. In: *Mobile Data Management (MDM), 2016 17th IEEE International Conference on.* (2016)
7. Arias, M., Arratia, A., Xuriguera, R.: Forecasting with twitter data. *ACM Transactions on Intelligent Systems and Technology (TIST)* **5**(1) (2013) 8
8. Kapoor, K., Srivastava, N., Srivastava, J., Schrater, P.: Measuring spontaneous devaluations in user preferences. In: *19th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM* (2013) 1061–1069
9. Rafailidis, D., Nanopoulos, A.: Modeling the dynamics of user preferences in coupled tensor factorization. In: *Proceedings of the 8th ACM Conference on Recommender systems, ACM* (2014) 321–324
10. Yang, Z., Xue, J., Wilson, C., Zhao, B.Y., Dai, Y.: Process-driven analysis of dynamics in online social interactions. In: *Proceedings of the 2015 ACM on Conference on Online Social Networks, ACM* (2015) 139–149
11. Liu, F.: Preference change and information processing. Technical report, ILLC, University of Amsterdam (2006)
12. Liu, F.: Preference change a quantitative approach. *Studies in Logic* **2**(3) (2009) 12–27
13. Tang, J., Musolesi, M., Mascolo, C., Latora, V.: Temporal distance metrics for social network analysis. In: *Proceedings of the 2nd ACM workshop on Online social networks, ACM* (2009) 31–36
14. Guille, A., Hacid, H.: A predictive model for the temporal dynamics of information diffusion in online social networks. In: *Proceedings of the 21st international conference companion on World Wide Web, ACM* (2012) 1145–1152
15. Tang, J., Musolesi, M., Mascolo, C., Latora, V., Nicosia, V.: Analysing information flows and key mediators through temporal centrality metrics. In: *3rd Workshop on Social Network Systems, ACM* (2010) 3
16. Wu, H., Cheng, J., Huang, S., Ke, Y., Lu, Y., Xu, Y.: Path problems in temporal graphs. *Proceedings of the VLDB Endowment* **7**(9) (2014) 721–732
17. Nicosia, V., Tang, J., Mascolo, C., Musolesi, M., Russo, G., Latora, V.: Graph metrics for temporal networks. In: *Temporal Networks. Springer* (2013) 15–40
18. Koujaku, S., Kudo, M., Takigawa, I., Imai, H.: Community change detection in dynamic networks in noisy environment. In: *24th International Conference on World Wide Web Companion.* (2015) 793–798
19. Liu, F.: Reasoning about Preference Dynamics. *Spring* (2011)
20. Xiang, L., Yuan, Q., Zhao, S., Chen, L., Zhang, X., Yang, Q., Sun, J.: Temporal recommendation on graphs via long-and short-term preference fusion. In: *16th ACM SIGKDD Int. Conf. Knowledge discovery and data mining.* (2010) 723–732
21. Wei, W., Carley, K.M.: Measuring temporal patterns in dynamic social networks. *ACM Transactions on Knowledge Discovery from Data (TKDD)* **10**(1) (2015) 9
22. Klochko, M.A., Ordeshook, P.C.: *Endogenous Time Preferences in Social Networks.* Edward Elgar Publishing
23. Pereira, F.S.F.: Mining comparative sentences from social media text. *Workshop on Interactions between Data Mining and Natural Language Processing 2015 co-located with European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases* (2015)
24. Zafarani, R., Abbasi, M.A., Liu, H.: *Social Media Mining: An Introduction.* Cambridge University Press, New York, NY, USA (2014)