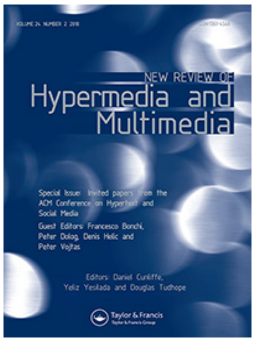


New Review of Hypermedia and Multimedia

ISSN: 1361-4568 (Print) 1740-7842 (Online) Journal homepage: http://www.tandfonline.com/loi/tham20

Implicit negative link detection on online political networks via matrix tri-factorizations

Mert Ozer, Mehmet Yigit Yildirim & Hasan Davulcu

NEW REVIEW OF HYPERMEDIA AND MULTIMEDIA
2018, VOL. 24, NO. 2, 63–87
https://doi.org/10.1080/13614568.2018.1482964

Taylor & Francis
Taylor & Francis Group

Implicit negative link detection on online political networks via matrix tri-factorizations

Mert Ozer, Mehmet Yigit Yildirim and Hasan Davulcu

School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, US

ABSTRACT
Political conversations have become a ubiquitous part of social media. When users interact and engage in discussions, there are usually two mediums available to them; textual conversations and platform-specific interactions such as like, share (Facebook) or retweet (Twitter). Major social media platforms do not facilitate users with negative interaction options. However, many political network analysis tasks rely on not only positive but also negative linkages. Thus, detecting implicit negative links is an important and a challenging task. In this work, we propose an unsupervised framework utilising positive interactions, sentiment cues, and socially balanced triads for detecting implicit negative links. We also present an online variant of it for streaming data tasks. We show the effectiveness of both frameworks with experiments on two annotated datasets of politician Twitter accounts. Our experiments show that the proposed frameworks outperform other well-known and proposed baselines. To illustrate the detected implicit negative links' contribution, we compare the community detection accuracies using unsigned and signed networks. Experimental results using detected negative links show superiority on the three datasets where the camps are known a priori. We also present qualitative evaluations of polarisation patterns between communities which are only possible in the presence of negative links.

ARTICLE HISTORY
Received 27 March 2018
Accepted 28 May 2018

KEYWORDS
Negative link; online political networks; non-negative matrix tri-factorization; online social networks; dynamic graph algorithms

1. Introduction

Beyond any doubt, social media has become a prominent platform for people to express their political stances and opinions for more than a decade (Ausserhofer & Maireder, 2013). It developed into a medium for politicians and political organisations to interact with the public (Pamelee & Bichard, 2011). To name a few, 44th President of the United States, Barack Obama makes an appearance on a Reddit Ask Me Anything, 45th President Donald Trump constantly utilises Twitter for his political messaging, many grassroot organisations mobilise their movements on Twitter and Facebook. Consequently, online social networks more and more are becoming an active field of study for political analysis tasks.

Many researchers have extensively studied the nature of online political networks (Conover, Goncalves, Ratkiewicz, Flammini, & Menczer, 2013; Conover et al., 2011;

CONTACT Mert Ozer mozer@asu.edu School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, US

Johnson & Goldwasser, 2016; Ozer, Kim, & Davulcu, 2016). Most of the existing works utilise platform-specific positive interactions between users such as share and like in Facebook or retweet and like in Twitter to infer insights from and model political activities in such social media platforms. (Conover et al., 2011) present how platform-specific positive interactions in Twitter shows a polarised behaviour in which one side does not retweet or like the other side's contents.

Major online social media platforms, however, do not provide its users options to state negative opinions in the form of a simple click such as "dislike" which might convey opposition or disagreement towards each other. Nonetheless, many political analysis tasks need the information of rivalries, resentments between political actors to get a complete picture of the online political landscape. This very nature of major social media platforms limit the capabilities of researchers studying online political networks effectively. Many researchers usually choose to study the online social networks where explicit negative links are available to them such as Epinions, Slashdot or Wikipedia instead (Dubois, Golbeck, & Srinivasan, 2011; Leskovec, Huttenlocher, & Kleinberg, 2010a; Yang, Smola, Long, Zha, & Chang, 2012). Certainly, these online platforms are not the hotspots where people participate to express their political views through.

Therefore, we focus on inferring the implicit negative links between users of online political networks. We aim to detect the link's negative nature, when any form of an overall disagreement, opposition or hostility is present between two social media users. It is a challenging problem due to two main reasons. First, there is no readily available online political network dataset in which negative links are explicitly present between its users. Therefore, the developed model must be unsupervised. Second, there is no simple predictor of negative links such as "dislike" in major social media platforms where the main body of the online political activity resides. However, opportunities are unequivocally present as well. Recent works in the social media mining research (Liu, Morstatter, Tang, & Zafarani, 2016; Tang, Chang, Aggarwal, & Liu, 2015) show that negative sentiment in the textual interaction between users is a good predictor of the negative link of those two users. Moreover, certain social psychology phenomenons such as social balance or social status theory are proven to be helpful in predicting negative links in certain network configurations (Leskovec, Huttenlocher, & Kleinberg, 2010b).

In this work, we first propose a nonnegative matrix factorization framework SocLS-Fact that combines signals from sentiment lexicon of words, platform-specific positive interactions and social balance theory to detect implicit negative and positive links in online political networks. We do not focus on the accuracy of the positive links since it is already a well studied problem and simple good predictors are already available. Additionally, we extend our SocLS-Fact framework to online settings to allow the integration and analysis of newly acquired data in a computationally efficient manner. Through this extension, it becomes convenient to run SocLS-Fact on a much smaller dataset without compromising effectiveness by utilising previously detected implicit links to calibrate the model.

We discuss two applications where detected implicit negative links can be employed to give a better understanding of the underlying political configuration of the target dataset. The first application is presented to show the added value of the detected implicit negative links in community detection tasks. The second application is proposed to show the informativeness of the detected implicit negative links related to polarisation patterns between political groups.

The main contributions of the paper are,

- Proposing SocLS-Fact an unsupervised model for implicit negative link detection in social media platforms where platform-specific negative interactions or negative links between users are not present.
- Introducing an online extension for SocLS-Fact to dynamically incorporate newly observed data while refraining from retraining the whole dataset.
- Showing the added value of the negative links in community detection tasks for online political networks.
- Providing two human-annotated online political network datasets for further research interest.

The rest of the paper is organised as follows. We discuss related work in the literature in Section 2. Then we propose our offline framework SocLS-Fact in Section 3 for implicit negative link detection task. In the following section we introduce the online variant of SocLS-Fact for online settings. In Section 5, we present experiments showing the effectiveness of both offline and online frameworks. As utilisation cases of detected implicit negative links, we present two applications of our framework in Section 6.

2. Related work

We survey link prediction, sentiment classification and dynamic network modeling methods proposed for similar line of research to ours in social media mining literature.

Link prediction in social media is an extensively studied problem. Its precedings can be traced back to the structuralist social psychology studies (Heider, 1958) that became popular in early 20th century. Link prediction studies standing out as most related to our problem definition are (Kunegis, Preusse, & Schwagereit, 2013; Leskovec et al., 2010a; Tang et al., 2015; Yang et al., 2012). Leskovec et al. (2010a) propose a framework that predicts the sign of user links in online networks. They train classifiers using certain triad configuration and graph features to learn from existing data in which both explicit positive and negative links are present. Yang et al. (2012) make use of explicit negative links through items that users comment to rather than using direct negative links between users. Signed bipartite graph of users and items is used to infer connectivity patterns among users. In their prediction model, they accommodate the principles of balance and status from social psychology theory.

However, these methods are not capable of being trained for major social media platforms (i.e. Twitter, Facebook) due to the nonexistence of explicit negative links or platform-specific negative interaction capabilities of users in those platforms. To address this limitation, (Kunegis et al., 2013) present an approach to predict negative links when only positive links are available explicitly. They further investigate the added value of negative links when they are predictable to a certain extent by using only properties of the positive links and not using any additional information such as textual content. However, they experiment only with Slashdot and Epinions datasets in which negative links or interactions between users are explicitly available. How generalisable their approach for other major social media platforms such as Facebook or Twitter, in which no platform-specific negative interaction is available, is not discussed. In Tang et al. (2015), Tang et al. introduce a supervised classification scheme to predict the negative links among missing links assuming that in many social media platforms, negative links are indirect and implicit. They use negative sentiment polarity of textual interactions between user pairs to synthetically generate the negative labelled links. This method also relies on experiments conducted only on Slashdot and Epinions datasets. On the other hand, our framework stands out as it is proposed for the social media platforms that do not provide any platform-specific negative interaction capabilities to their users.

Second line of research related to our work is sentiment classification in social media. Hu, Tang, Tang, & Liu (2013) propose a supervised sentiment classification model which takes advantage of connected text messages having similar sentiment labels. Hu, Tang, Gao, & Liu (2013) further investigate whether emotional signals such as emoticons can be incorporated in order to infer the sentiment classes of the tweets in Twitter. To credit the informative value of the overall sentiment of the textual interactions between users for predicting the polarity of the user link, Hassan, Abu-Jbara, & Radev (2012) propose a supervised classification framework. It considers all textual interactions of the user pairs' and learn relevant sentiment features from human annotated prior user link polarities. However, it does not use any platform-specific interaction types which are vastly available on many social media platforms. West, Paskov, Leskovec, & Potts (2014) develop a model that combinatorially optimises the agreement between the sentiment class of user pairs' textual interaction and the polarity label of the explicit user link. They make use of Wikipedia, and U.S. Congress dataset, in which explicit negative links or platform specific negative interactions are available. Our work differentiates itself, from aforementioned others in the literature by using platform-specific positive interactions, a sentiment lexicon of words and socially balanced triads to detect the implicit negative links between users.

The last line of research related to our work is dynamic network modeling methods. Aktunc, Ozer, Toroslu, & Davulcu (2015) propose an extension of well-known modularity based smart local moving algorithm for dynamic networks. The main goal of the modeling is community detection. The work concerns with only explicitly defined links on networks. Mankad & Michailidis (2013) propose a non-negative matrix factorization approach for modeling dynamic networks. They also utilise the concept of temporal smoothing as our online framework does. However they do not take any other form of interaction in networks other than explicit links. On a similar line of research, Yu, Aggarwal, & Wang (2017) propose modeling dynamics of networks by using temporal matrix factorization. They investigate the modeling options of temporal unfolding of networks by factorising different snapshots of networks into one constant and one time-varying matrix. Similar to our modeling, they also use a decay function to weight importance of previous snapshots in temporal order. For predicting links in dynamic networks, Zhu, Guo, Yin, Ver Steeg, & Galstyan (2016) propose using a temporal latent space. They assume two users who are located closer in a temporal latent space is likely to form link in the next snapshot. The concept of temporal smoothing also plays an important role in their dynamic network modeling. Our work stands out from aforementioned four works by focusing on implicit links rather than explicit ones as majority of social media platforms do not allow explicit negative links. Moreover, these previous efforts do not involve incorporating textual interactions, sentiment signals or social balance theory.

Table 1. Notation.

Symbol	Size	Explanation
m		Number of interacting user pairs
n		Number of sentiment words
I_k	$k \times k$	Identity matrix of size k
$\mathbf{X}$	$[m \times n]$	Matrix of occurrences of sentiment words in textualinteractions ofuser pairs
$\mathbf{S}_u$	$[m \times 2]$	User link polarity
$\mathbf{S}_{u0}$	$[m \times 2]$	Initial user link polarity
$\mathbf{D}_u$	$[m \times m]$	Binary diagonal matrix of user pairs with positiveinteraction
$\mathbf{S}_w$	$[n \times 2]$	Sentiment word polarity
$\mathbf{S}_{w0}$	$[n \times 2]$	Initial sentiment lexicon
$\mathbf{M}$	$[m \times m]$	Social balance matrix

3. Offline framework

In this section, we first present the notation used throughout the paper, formally define the problem and then propose the SocLS-Fact optimisation solution. Finally, we provide the details of how to build the prior knowledge that the SocLS-Fact requires.

Before going into the details of the framework, the notation that is used throughout the paper can be seen in Table 1. m is the number of interacting user pairs, and n is the number of unique sentiment words. An example with 3 interacting user pairs and 8 unique sentiment words can be seen in Figure 1(a,b). All textual interaction happening between two users are represented as rows of $\mathbf{X}$. $\mathbf{X}$ encodes how many times each sentiment word occurs in textual interactions of two users. In Figure reffig:pipeline(b), when user a and b interacts they use 2nd, 3rd, 5th and 6th words while user b and c interacts they use 1st, 3rd and 8th and so on. Initial user link polarities are embedded in matrix $\mathbf{S}_{u0}$. Initial sentiment lexicon is embedded in $\mathbf{S}_{w0}$. Positive and negative polarities are represented as two latent dimensions in matrix $\mathbf{S}_{u0}$, and $\mathbf{S}_{w0}$. Which user links should have the same polarity following the social balance theory is governed by matrix $\mathbf{M}$. Further details of how matrices $\mathbf{S}_{u0}$, $\mathbf{S}_{w0}$, $\mathbf{M}$ are derived is given later in this section.

As we discuss earlier, sentiment of words used in user interactions are proven to be good predictors of the polarity of user links. Moreover, built-in positive interactions (i.e. retweet, like, share) are good predictors of positive user links by their nature. As

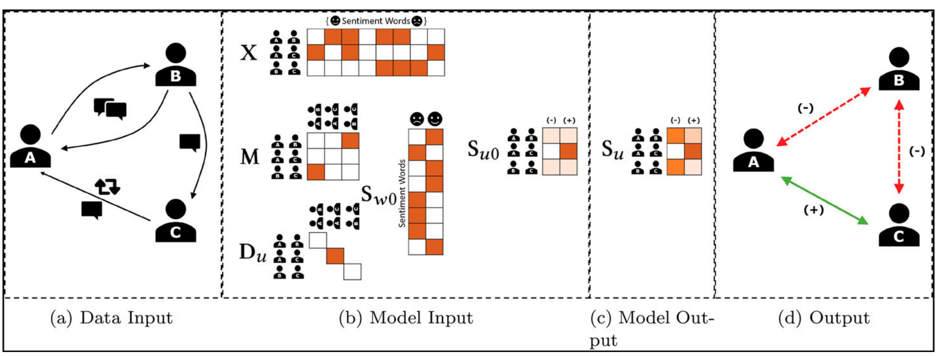


Figure 1. (Colour online) Modeling of social media data and interpretation of output. Darker hues of orange imply larger values, while white implies the 0 in the corresponding entry. (a) Data Input, (b) Model Input, (c) Model Output, (d) Output.

referred in Section 1, how user links form triangles with each other is also a decisive factor of their polarities since they tend to follow social balance theory. To factorise all textual interactions between users into two latent dimensions as positive and negative and enjoy aforementioned three predictors of polarity of user links at the same time, we propose the following optimisation problem;

$$\min_{\mathbf{S_u},\mathbf{H},\mathbf{S_w}} \quad \|\mathbf{X}-\mathbf{S}_u\mathbf{H}\mathbf{S}_w^T\|_F^2 \tag{0}$$

$$+\,\alpha\|\mathbf{S}_w-\mathbf{S}_{w0}\|_F^2 \tag{1}$$

$$+\,\beta\,\mathrm{Tr}((\mathbf{S}_u-\mathbf{S}_{u0})^T\mathbf{D}_u(\mathbf{S}_u-\mathbf{S}_{u0})) \tag{2}$$

$$\begin{aligned} &+\,\gamma\|\mathbf{M}-\mathbf{S}_u\mathbf{S}_u^T\|_F^2 \\ \text{subjectto} \quad &\mathbf{S}_u>0,\mathbf{S}_w>0,\mathbf{H}>0 \end{aligned} \tag{3}$$

Optimisation formulation consists of four terms. (0)th term factorises user pair textual interactions into three matrices. $\mathbf{S}_u\in\mathbb{R}_+^{m\times 2}$ is the lower-rank projection of matrix $\mathbf{X}$. First column of $\mathbf{S}_u$ is the latent negative and second column is the latent positive dimension. $\mathbf{S}_w$ is the lower-rank projection of columns of matrix $\mathbf{X}$. Each column of $\mathbf{X}$ represents a sentiment word. Projection matrix $\mathbf{S}_w$ corresponds to distributed polarity representation of each sentiment word. As in $\mathbf{S}_u$, first column of $\mathbf{S}_w$ is the latent negative and the second column is the latent positive dimension.

(1)st term in the optimisation formulation penalises the meaning change of the sentiment words compared to their initial lexicon meaning. Parameter α governs the relaxation on the penalty.

(2)nd term governs how much the polarity prediction of links diverges from their initial inferred labels. Initial labels are inferred as positive if there is any platform-specific positive interaction between users that the link connecting to. Diagonal matrix D_u helps to penalise divergences of links which have platform-specific positive interactions only.

(3)rd term in the optimisation formulation penalises the triangles in the user network that do not follow social balance theory. $\mathbf{M}$ encodes the information of pair of links that should have the same polarity if they form a triangle with another positive link.

3.1. Constructing $\mathbf{S}_{w0}$

An off-the-shelf sentiment word lexicon is utilised[1] to populate the initial sentiment polarities of words. A word is represented as [1, 0] if it has negative sentiment. It is represented as [0, 1] if it has a positive sentiment. In Figure 1(b), initial sentiment lexicon is embedded in $\mathbf{S}_{w0}$ such that 1st, 3rd, 4th and 8th words as positive sentiment words and 2nd, 5th, 6th and 7th words as negative sentiment words.

3.2. Constructing $\mathbf{S}_{u0}$ and $\mathbf{D}_u$

Each row of the initial user link polarity matrix $\mathbf{S}_{u0}$ encodes the information of the prior inference of the polarity of user link. First column of the polarity matrix $\mathbf{S}_{u0}$ is the latent

[1]http://www.cs.uic.edu/~liub/FBS/opinion-lexicon-English.rar

negative dimension, while the second column is the latent positive dimension. For the links that connect user pairs having previous platform-specific positive interaction, we infer the initial polarity of them as positive and embed it as $[0, 1]$ in the corresponding row of $\mathbf{S}_{u0}$ and as 1 in the corresponding diagonal entry of $\mathbf{D}_u$. For the links that connect user pairs having no previous platform-specific positive interaction, we do not infer any initial polarity and represent them as $[0.5, 0, 5]$ in $\mathbf{S}_{u0}$ and as 0 in the corresponding diagonal entry of $\mathbf{D}_u$. To illustrate in Figure 1(b), the positive interaction between user A and C is represented as $[0, 1]$ in the second row of $\mathbf{S}_{u0}$ and as 1 in the second diagonal entry of $\mathbf{D}_u$.

3.3. Incorporating social balance theory

The theory of social balance of signed links in triads is extensively studied since its introduction by Heider et al. in (Heider, 1958) as structural balance of signed triads. It suggests that for a signed triad to be balanced, it has to have an odd number of positive links (i.e. one or three positive links), otherwise it is not balanced. The balanced configurations among all possible configurations are presented with dashed frames in Figure 2. The definition of structural balance is analogous to common phrase of "enemy of my enemy is my friend" and "friend of my friend is my friend" in social settings.

To encode the social balance theory, we utilise the prior knowledge of positive links inferred from platform-specific positive interactions. Our intuition is that if two users have any prior platform-specific positive interaction, the polarity of their interaction with any other third user should be similar. They can connect to third user either with both negative or positive links (i.e. Triad-1 and Triad-2 in Figure 2). The cases which they connect to a third user with different polarities are not socially balanced configurations (i.e. Triad-3 in Figure 2).

The matrix $\mathbf{M} \in \{0, 1\}^{m \times m}$ encodes the link pairs that are needed to have the same polarity to follow social balance theory by having 1 in the related row and column of $\mathbf{M}$ and 0 for the rest. In Figure 1(a), link between user A and B should have the same polarity with link between user B and C. It is because they are forming a triad with link between user A and C which has prior platform-specific positive interaction. In Figure 1(b), it is encoded as 1 in the $\mathbf{M}(1, 3)$ and $\mathbf{M}(3, 1)$. Eventually, minimising the squared frobenious norm of the difference between $\mathbf{M}$ and $\mathbf{S}_u\mathbf{S}_u^T$ forces triads to have odd number of positive links in the whole network.

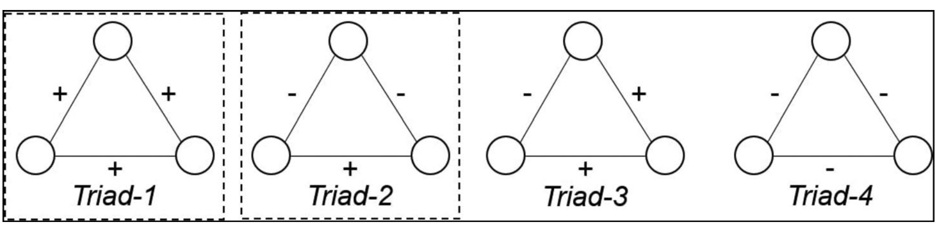


Figure 2. All possible configurations of undirected signed links in a triad. Balanced ones are framed with dashed rectangles.

3.4. Algorithm

The objective function proposed in Section 3 is not convex for all variables of $\mathbf{S}_u$,$\mathbf{S}_w$,$\mathbf{H}$. We introduce an alternating optimisation solution for our problem similar to Li, Zhang, & Sindhwani (2009). We update each variable $\mathbf{S}_u$,$\mathbf{S}_w$,$\mathbf{H}$ iteratively while fixing others to find a local minimum in the solution space. The update rules for each variable is given as;

$$\mathbf{S}_u \leftarrow \mathbf{S}_u \odot \sqrt{\frac{\mathbf{X}\mathbf{S}_w\mathbf{H}^T + \gamma(\mathbf{M}+\mathbf{M}^T)\mathbf{S}_u + \beta\mathbf{D}_u\mathbf{S}_{u0}}{\mathbf{S}_u\mathbf{H}\mathbf{S}_w^T\mathbf{S}_w\mathbf{H}^T + \gamma\mathbf{S}_u\mathbf{S}_u^T\mathbf{S}_u + \beta\mathbf{D}_u\mathbf{S}_u}} \tag{3.1}$$

$$\mathbf{H} \leftarrow \mathbf{H} \odot \sqrt{\frac{\mathbf{S}_u^T\mathbf{X}\mathbf{S}_w}{\mathbf{S}_u^T\mathbf{S}_u\mathbf{H}\mathbf{S}_w^T\mathbf{S}_w}} \tag{3.2}$$

$$\mathbf{S}_w \leftarrow \mathbf{S}_w \odot \sqrt{\frac{\mathbf{X}^T\mathbf{S}_u\mathbf{H} + \alpha\mathbf{S}_{w0}}{\mathbf{S}_w\mathbf{H}^T\mathbf{S}_u^T\mathbf{S}_u\mathbf{H} + \alpha\mathbf{S}_w}} \tag{3.3}$$

Derivation of the update rules is presented in Appendices of a previous work where SocLSFact was initially introduced Ozer, Yildirim, & Davulcu (2017). The proposed algorithm employs an iterative scheme of the above rules until convergence or certain number of iterations are met. Each step of the algorithm is shown in Algorithm 1.

Finally, the polarity of the latent dimension with higher numerical value in the i^{th} row of $\mathbf{S}_u$ is assigned as the polarity output of the link i. To illustrate in Figure 1(c,d), it can be seen that the value in the first column is greater than the second column for the first and the third rows of $\mathbf{S}_u$. Therefore, the polarity of the link between user A and B and the link between user B and C are inferred as negative. Since the value in the second column is greater than the first column for the second row of $\mathbf{S}_u$ the polarity of the link between user A and C is inferred as positive. Further details of the proposed offline framework can be found in Ozer et al. (2017).

4. Online framework

Given the dynamic nature of online political networks, it is necessary to handle streaming data in an online fashion. It usually is computationally expensive to re-run offline methods from scratch each time a new piece of arrives. One naive solution is running the offline method only on the new data. It is a faster solution, yet, it ignores the rich historical information.

Algorithm 1: Proposed Algorithm for Offline Framework's Optimisation Problem

Input: $\mathbf{X}$,$\mathbf{S}_{u0}$,$\mathbf{S}_{w0}$, $\mathbf{M}$.
Output: $\mathbf{S}_u$,$\mathbf{S}_w$.
1 Initialise $\mathbf{S}_u \leftarrow \mathbf{S}_{u0}$,$\mathbf{H} \leftarrow I_2$, $\mathbf{S}_w \leftarrow \mathbf{S}_{w0}$.
2 **while** *not convergent* **do**
3 Update $\mathbf{S}_u$ using Equation 3.1.
4 Update $\mathbf{H}$ using Equation 3.2.
5 Update $\mathbf{S}_w$ using Equation 3.3.

To alleviate the aforementioned problems, we introduce an online framework. It follows similar principles as the offline framework besides the modelling of temporal dimension. It uses sentiment words, prior positive interactions, and socially balanced triads to infer implicit negative links. To take previous snapshots' detected implicit links into account, we propose using a temporal smoothing term. This smoothing term penalises abrupt changes of the signs of the links in consecutive snapshots. Thus, we propose solving the following optimisation problem for online settings;

$$\min_{\mathbf{S_u}^{(t)},\mathbf{H}^{(t)},\mathbf{S}_w^{(t)}} \quad \|\mathbf{X}^{(t)} - \mathbf{S}_u^{(t)}\mathbf{H}^{(t)}\mathbf{S}_w^{(t)^T}\|_F^2 \tag{0}$$

$$+\, \alpha\|\mathbf{S}_w^{(t)} - \mathbf{S}_{w0}\|_F^2 \tag{1}$$

$$+\, \beta\,\mathrm{Tr}((\mathbf{S}_u^{(t)} - \mathbf{S}_{u0}^{(t)})^T\mathbf{D}_u^{(t)}(\mathbf{S}_u^{(t)} - \mathbf{S}_{u0}^{(t)})) \tag{2}$$

$$+\, \gamma\|\mathbf{M}^{(t)} - \mathbf{S}_u^{(t)}\mathbf{S}_u^{(t)^T}\|_F^2 \tag{3}$$

$$\begin{aligned} &+\, \tau\sum_{i=1}^{t} \mathrm{e}^{-(t-i)}\|\mathbf{S}_u^{(t)} - \mathbf{S}_u^{(i)}\|_F^2 \\ \text{subject to} \quad &\mathbf{S}_u^{(t)} > 0, \mathbf{S}_w^{(t)} > 0, \mathbf{H}^{(t)} > 0 \end{aligned} \tag{4}$$

In above formulation, t stands for the current time snapshot and any matrix superscripted by parameter t (e.g. $\mathbf{X}^{(t)}$) spans the data of tth snapshot from time $(t-1)$ to t. First four terms (0,1,2,3) in the objective function are inherited from the offline framework. $(4)^{th}$ term controls the divergence of current snapshot's signs of links from previous time snapshots'. An inverse exponential decay function ($\mathrm{e}^{-(t-i)}$) is employed to weight previous snapshots' importance in temporal order. One can simply plug another decay function based on their application's constraints when necessary. Parameter τ controls the importance of temporal smoothing.

4.1. Algorithm

To optimise the online framework objective function, we follow similar iterative multiplicative update rules as in the offline framework. While updating $\mathbf{S}_u^{(t)}$, we treat emerging links and continuing links exclusively, since they are subject to different temporal smoothing constraints. For emerging links as there is no precedent of them in previous snapshot, we employ the update rule of the offline framework. We denote rows of $\mathbf{S}_u$ corresponding to emerging links as $\mathbf{S}_{ue}$,

$$\mathbf{S}_{ue}^{(t)} \leftarrow \mathbf{S}_{ue}^{(t)} \odot \sqrt{\frac{\mathbf{X}_e^{(t)}\mathbf{S}_w^{(t)}\mathbf{H}^{(t)^T} + \gamma(\mathbf{M}_e^{(t)} + \mathbf{M}_e^{(t)^T})\mathbf{S}_{ue}^{(t)} + \beta\mathbf{D}_{ue}^{(t)}\mathbf{S}_{ue0}^{(t)}}{\mathbf{S}_{ue}^{(t)}\mathbf{H}^{(t)}\mathbf{S}_w^{(t)^T}\mathbf{S}_w^{(t)}\mathbf{H}^{(t)^T} + \gamma\mathbf{S}_{ue}^{(t)}\mathbf{S}_{ue}^{(t)^T}\mathbf{S}_{ue}^{(t)} + \beta\mathbf{D}_{ue}^{(t)}\mathbf{S}_{ue}^{(t)}}} \tag{4.1}$$

For continuing links, we incorporate the temporal smoothing term, so, the update rule for continuing links become,

$$\mathbf{S}_{uc}^{(t)} \leftarrow \mathbf{S}_{uc}^{(t)} \odot \sqrt{\frac{\mathbf{X}_c^{(t)}\mathbf{S}_w^{(t)}{\mathbf{H}^{(t)}}^T + \gamma(\mathbf{M}_c^{(t)} + {\mathbf{M}_c^{(t)}}^T)\mathbf{S}_{uc}^{(t)} + \beta \mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc0}^{(t)} + \tau\sum_{i=1}^{t} \mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(i)}}{\mathbf{S}_{uc}^{(t)}\mathbf{H}^{(t)}{\mathbf{S}_w^{(t)}}^T\mathbf{S}_w^{(t)}{\mathbf{H}^{(t)}}^T + \gamma\mathbf{S}_{uc}^{(t)}{\mathbf{S}_{uc}^{(t)}}^T\mathbf{S}_{uc}^{(t)} + \beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)} + \tau t\mathbf{S}_{uc}^{(t)}}} \tag{4.2}$$

From the perspective of matrices $\mathbf{H}^{(t)}$ and $\mathbf{S}_w^{(t)}$, there is no temporal smoothing involved. So, same update rules can be employed as in the offline framework in a snapshot-based fashion;

$$\mathbf{H}^{(t)} \leftarrow \mathbf{H}^{(t)} \odot \sqrt{\frac{{\mathbf{S}_u^{(t)}}^T\mathbf{X}^{(t)}\mathbf{S}_w^{(t)}}{{\mathbf{S}_u^{(t)}}^T\mathbf{S}_u^{(t)}\mathbf{H}^{(t)}{\mathbf{S}_w^{(t)}}^T\mathbf{S}_w^{(t)}}} \tag{4.3}$$

$$\mathbf{S}_w^{(t)} \leftarrow \mathbf{S}_w^{(t)} \odot \sqrt{\frac{{\mathbf{X}^{(t)}}^T\mathbf{S}_u^{(t)}\mathbf{H}^{(t)} + \alpha\mathbf{S}_{w0}^{(t)}}{\mathbf{S}_w^{(t)}{\mathbf{H}^{(t)}}^T{\mathbf{S}_u^{(t)}}^T\mathbf{S}_u^{(t)}\mathbf{H}^{(t)} + \alpha\mathbf{S}_w^{(t)}}} \tag{4.4}$$

Derivation of the update rule of $\mathbf{S}_{uc}^{(t)}$ is given in Appendix 1. Derivation of the update rules of $\mathbf{S}_{ec}^{(t)}$, $\mathbf{H}^{(t)}$, and $\mathbf{S}_w^{(t)}$ follow the exact offline framework calculations in Ozer et al. (2017). Given the update rules, algorithm for finding the optimal $\mathbf{S}_{ue}^{(t)}$ and $\mathbf{S}_{uc}^{(t)}$ becomes straightforward, and presented in Algorithm 2

As in the offline framework, the polarity of a link is assigned based on the values in the corresponding row of the link in $\mathbf{S}_u^{(t)}$. If the first value is larger, the link is inferred as negative, and otherwise, as positive.

Complexity of the algorithm can be formulated as follows. Update rule for $\mathbf{S}_{ue}^{(t)}$ (Eq. (4.1)), $\mathbf{H}^{(t)}$ (Eq. (4.3)), and $\mathbf{S}_w^{(t)}$ (Eq. (4.4)) are same as in the offline framework; Equation (3.1), 3.2 and 3.3. They are $\mathcal{O}(mn + m^2 + m + n^2m)$, $\mathcal{O}(mn + m + n)$ and $\mathcal{O}(mn + m^2n)$, respectively. Complexity of updating $\mathbf{S}_{uc}^{(t)}$ is $\mathcal{O}(mn + m^2 + m + n^2m + tm)$. Therefore, whole complexity of the online framework becomes $\mathcal{O}(i(m^2n + n^2m + m^2 + mn + tm + n))$ where i is the number of iterations of applying multiplicative update rules. The added time complexity the online framework introduces is due to the term tm. The source code for the whole running pipeline for both offline and online frameworks can be reached at www.public.asu.edu/~mozer/NRHMcode.tar.gz.

Algorithm 2: Proposed Algorithm for the Online Framework's Optimisation Problem

Input: $\mathbf{X}^{(t)}$, $\mathbf{S}_{u0}^{(t)}$, $\mathbf{S}_{w0}$, $\mathbf{M}^{(t)}$, $\mathbf{S}_u^{(i)}$, i=1,2…,t−1
Output: $\mathbf{S}_u^{(t)}$, $\mathbf{S}_w^{(t)}$.
Initialise $\mathbf{S}_u^{(t)} \leftarrow \mathbf{S}_{u0}^{(t)}$, $\mathbf{H}^{(t)} \leftarrow I_2$, $\mathbf{S}_w^{(t)} \leftarrow \mathbf{S}_{w0}$.
while *not convergent* **do**
    Update emerging links $\mathbf{S}_{ue}^{(t)}$ using Equation 4.1.
    Update continuing links $\mathbf{S}_{uc}^{(t)}$ using Equation 4.2.
    Update $\mathbf{H}^{(t)}$ using Equation 4.3.
    Update $\mathbf{S}_w^{(t)}$ using Equation 4.4.

5. Experiments

In this section, we present experiments to evaluate the performance of our offline and online frameworks. In the first experiment, we investigate the effectiveness of the offline framework and in the second experiment, we compare online framework's performance with variants of offline framework in implicit negative link detection task.

5.1. Dataset

We crawl tweets by members of the 56th and 57th Parliament of United Kingdom using GET user_timeline function of Twitter API. Each parliament member usually self-describes when the account is associated with their parliament identity in their user profile. All of the accounts in the dataset are verified Twitter accounts.

- **56th Parliament Dataset** covers 1074 user pairs sampled from 400 members of the 56th Parliament of United Kingdom on Twitter. Polarity of each user link is annotated using three human annotators.
- **57th Parliament Dataset** covers 1349 user pairs sampled from 561 members of the 57th Parliament of United Kingdom on Twitter. Polarity of each user pair is annotated by yearly snapshots. It spans three snapshots, namely, "2016 → 2017", "2017 → 2018", and "2018 →". "2018 →" snapshot spans the first two months of 2018. The task of annotation involves three human annotators. Details of the annotation are explained in Section 5.1.1.

Users who do not participate in any textual user interaction are removed from the dataset. For implicit negative link detection task, it is essential to obtain labels for the links between users to (1) test the effectiveness of our algorithm, (2) have a grasp on the effect of the parameters. Thus, we hired three graduate students for our annotation task. An overview of the annotated datasets can be seen in Table 2. Tweet ids, user ids and annotated user links of both datasets used in our experiments can be retrieved fromwww.public.asu.edu/ mozer/NRHMdata.tar.gz.

To evaluate the performance of the offline algorithm, we experiment with the 56th Parliament dataset and an aggregated single-view of the 57th Parliament dataset over three snapshots. To aggregate human annotated labels of the 57th Parliament dataset, we use the latest available label in three snapshots for each link. For online algorithm's experiments, we use the 57th Parliament dataset as it is.

Table 2. Dataset statistics.

	56th Parliament	57th Parliament		
		2016 → 2017	2017 → 2018	2018 →
Textual interactions	4217	1297	3947	1459
Interacting user pairs	1074	460	1099	602
+/− links	948/126	433/27	977/122	526/76
(+, +, +) triads	732	150	1257	294
(+, +, −) triads	61	0	72	15
(+, −, −) triads	68	12	126	30
(−, −, −) triads	11	0	3	0
Sentiment Tokens	1,225	543	1,064	615

5.1.1. Annotation task

For 56th Parliament dataset, we aggregated all the textual interactions (i.e. tweets identified as mentions and reply to's) of user pairs. For 57th Parliament we aggregated interactions into three snapshots ("2016 → 2017", "2017 → 2018", "2018 →"). We filtered the data to include textual interactions which contains a single user mention to avoid the confusion as it is ambiguous which user is addressed in the multiple mentions case.

We requested 3 graduate students who had knowledge of UK politics to rate the polarity of the interactions between two politician accounts. For a pair of users, we have provided all textual interactions, political party affiliations, and retweet counts between the users to help annotators assess the polarity of the link better. After retrieving all the answers from three annotators, we assigned the polarity labels using majority voting.

We analysed the labelers inter-rater agreement using Cohen's Kappa (Landis & Koch, 1977) and Fleiss' Kappa (Fleiss, 1971) to ensure annotation quality. Two-way inter-rater agreement is nearly perfect according to Landis & Koch (1977) with Cohen's Kappa scores calculated as 0.810, 0.898 and 0.911. Fleiss' kappa is reported as 0.731.

Finally, we remove the neutral user links as they are not covered by our problem formulation.

5.2. Offline framework performance

Our first experiment aims to demonstrate the implicit negative link detection performance of SocLS-Fact in offline settings. To assess the performance of our method, we explain and compare with two existing state-of-the-art matrix factorization approaches along with three other baseline predictors we define as follows:

- **Random**: Motivated by Liben-Nowell (2003), this method predicts signs of user links randomly.
- **Only Sentiment**: This predictor infers the polarity of user pairs' links using only textual interaction. Sum of the inverse distance weighted sentiment values $(+1, -1)$ of words in textual interactions is given as the polarity of the link between user pairs.
- **Only Link**: This predictor infers user pairs' links as positive if there is any historical platform-specific positive interaction between them and negative otherwise.
- **NMTF[Ding, Li, Peng, & Park (2006)]**: This predictor is a simple non-negative matrix tri-factorization method without any regularizers of sentiment lexicon, link prior or social balance.
- **SSMFLK[Li et al. (2009)]**: Proposed as sentiment classification method, it is a semi-supervised matrix factorization framework utilising prior sentiment lexicon knowledge. This method is similar to SocLS-Fact method, however, it does not encode platform-specific positive interaction between users or social balance theory.
- **LS-Fact**: This predictor is a variant of the proposed method but it does not embed social balance theory. It is introduced as a baseline to show the effect of social balance regularizer.

Methods using regularizer coefficients (i.e. SSMFLK, LS-Fact, SocLS-Fact) are experimented with all powers of 10 from -6 to 2 and the best performance is reported.

Evaluation metrics

We use three gold-standard metrics, namely; accuracy, precision, and *F*-measure to evaluate our method. Scores are reported in terms of our method's detection performance on the negative links. We do not report recall explicitly as we emphasise quality over quantity; retrieving meaningful negative links is the most important task in this work as suggested for many tasks in Wang, Pedreschi, Song, Giannotti, & Barabasi (2011). The change in recall can be indirectly observed through *F*-measure. Although we present the accuracy for reader convenience solely focusing on accuracy may be misleading considering the imbalanced nature of our dataset. Hence, we focus mainly on precision and *F*-measure throughout the discussion of our results.

Results

An overview of the implicit negative link detection performance of the proposed and baseline methods can be found in Table 3. As can be clearly observed through the table, performance increase is consistent among all three metrics: precision, *F*-measure and accuracy. Important findings are reported below:

- Encoding the sentiment information using SSMFLK improves the performance over the random classifier.
- An interesting finding can be observed when "only sentiment" predictor is used. It yields better results than SSMFLK due to its deterministic nature; whereas SSMFLK may be highly affected by the random starting conditions.
- Only link predictor gives much better results than using just the sentiment information. A steep increase in all three metrics is evident that prior platform specific positive interaction is a very strong signal that the link between users is not negative.
- Co-optimising the link information with sentiment information in LS-Fact framework results in superior performance compared to both only link and only sentiment predictors.
- Finally, our framework, SocLS-Fact obtains the best results by incorporating the social balance theory into the framework. SocLS-Fact performs slightly better than LS-Fact thanks to the user link triads following social balance theory in formation. *F*-measure performance contribution of socially balanced triangles is higher for 57th Parliament dataset than 56th, as higher ratio of socially balanced triangles can be observed in Table 2.

Table 3. Offline implicit negative link detection performance on the 56th and 57th Parliament datasets.

	56th parliament dataset			57th parliament dataset		
	Prec.	*F*-meas.	Acc.	Prec.	*F*-meas.	Acc.
Random	0.1450	0.2344	0.5317	0.1707	0.2709	0.5664
SSMFLK[a]	0.3143	0.4490	0.7737	0.3708	0.4599	0.8426
Only Sentiment	0.4010	0.4892	0.8464	0.3364	0.4207	0.8333
Only Link	0.6032	0.6726	0.9062	0.5312	0.6733	0.9021
NMTF[b]	0.6741	0.6973	0.9264	**0.8243**	0.5622	0.9271
LS-Fact	0.6976	0.7059	0.9302	0.7091	0.7548	0.9434
SocLS-Fact	**0.7236**	**0.7149**	**0.9339**	0.7742	**0.8**	**0.9553**

[a]Li et al. (2009).
[b]Ding et al. (2006).

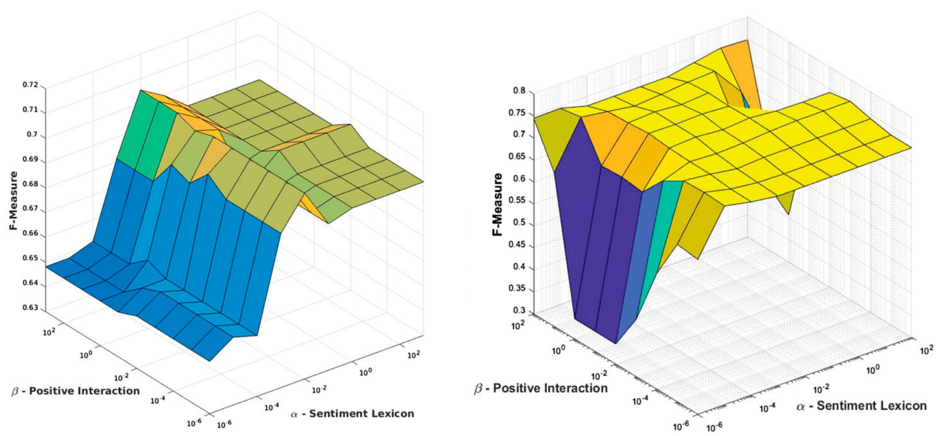
Parameter analysis

It is essential that our framework performs effectively under different parameter settings. So, we experiment with various values of α, β, and γ then report the performance in terms of F-measure scores. Best performance was obtained using the parameters $\alpha = 10^{-2}$, $\beta = 100$, and $\gamma = 10^{-1}$ for 56th Parliament Dataset, $\alpha = 10$, $\beta = 10^{-5}$, and $\gamma = 10^{-5}$.

Figure 3 demonstrates the effect of sentiment lexicon parameter α and prior platform-specific positive interaction parameter β when the social balance regularizer γ is fixed at 0. α and β are tweaked as powers of 10 between −6 to 2. Parameters out of this range gives very low F-measure scores thus excluded.

- SocLS-Fact is robust to changes of α and β when α is in the range of 10^{-5} and 1 as F-measure does not differ more than 0.07 for both datasets.
- Lower values of α yield the lowest F-measure scores. Performance sharply increases when α is incremented from 10^{-6} towards 10^{-2}.
- Change of β creates rather stable results for any given α in 56th Parliament dataset and αs between 10^{-5} and 1 in 57th Parliament dataset.

Figure 4 shows how social balance regularizer γ affects the performance when the other parameters are fixed at optimal values, 10^{-2} and 100 for 56th Parliament and 10 and 10^{-5} for 57th Parliament dataset, respectively. γ is supplied incrementally as powers of 10 between −5 to 1. As the chart shows, SocLS-Fact is robust also to changes of γ performing in a F-measure margin of 0.025 for 56th Parliament dataset. The margin for 57th Parliament dataset is 0.1. Both chart shows that with the optimal setting of γ, social balance theory can contribute to achieve a superior performance in implicit negative link detection task. The optimal γ parameters are 10^{-1} for the 56th Parliament dataset and 10^{-5} for the 57th Parliament dataset.



(a) (b)

Figure 3. (Colour online) Effect of regularizer coefficients. (a) 56th Parliament Dataset Experiments, (b) 57th Parliament Dataset Experiments.

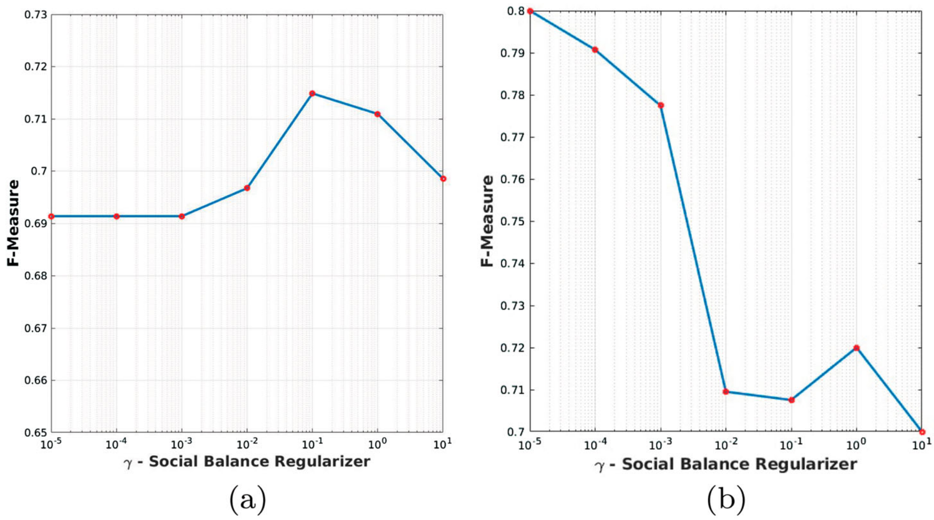


Figure 4. (Colour online) Effect of Social Balance Regularizer Under Optimal Positive Prior and Sentiment Lexicon Regularizers. (a) 56th Parliament Dataset, (b) 57th Parliament Dataset.

5.3. Online framework performance

In this section, we discuss the performance of the online framework by presenting comparisons with variants of the offline framework. As mentioned in the previous sections, conventional ways of dealing with streaming data using an offline methodology usually involves either computing everything from scratch, or ignoring the historical data. Both extremes have their disadvantages. To show the trade-offs between these two approaches, we propose experimenting with the following two baselines and our online method;

- **SocLSFact** detects signs of the links only based on the current snapshot data.
- **SocLSFact [A]** detects signs of the links based on aggregation of current and all previous snapshots.
- **SocLSFact (Online)** detects signs of the links based on signs of the detected implicit links in the previous snapshots and factorise only the current snapshot data.

In this experimental setup, we utilise 57th Parliament dataset which is labelled in three snapshots. Results are reported based on the last snapshot (2018 →) data. SocLSFact works only on the last (2018 →) snapshot data to detect implicit links in it. SocLSFact [A] aggregates all three snapshots into single view and detect implicit links, accordingly. SocLSFact (Online) model uses SocLSFact outputs of two previous snapshots (2016 → 2017 and 2017 → 2018) for temporal smoothing and factorises only the last snapshot (2018 →) data to detect implicit links. Parameters (α, β, γ, τ) are explored with all powers of 10 from −6 to 2.

Results

An overview of the implicit negative link detection in online settings can be seen in Table 4 and in Figure 5(a). In terms of maximum performance, SocLSFact (Online) model

Table 4. Online implicit negative link detection maximum performances on the last snapshot of the 57th Parliament dataset.

	2018→		
	Prec.	*F*-Meas.	Acc.
SocLSFact	0.6588	0.7134	0.9233
SocLSFact [A]	0.7742	0.8	0.9553
SocLSFact (Online)	**0.8406**	**0.8227**	**0.9574**

performs better than both SocLSFact and SocLSFact [A] baseline models. In other words, while modelling SocLSFact for online settings, temporal smoothing among other two options increases the performance in all three metrics. SocLSFact [A], which aggregates previous snapshots' data as they are, is still the better choice than running factorization only on the last snapshot. SocLSFact (Online) achieves 15% higher *F*-measure, 28% higher precision, and 4% higher accuracy than SocLSFact and achieves 3% higher *F*-measure, 9% higher precision, and 0.2% higher accuracy than SocLSFact [A] in implicit negative link detection.

To better evaluate the trade-off between run-time and effectiveness of these three methods, we run each method 100 times with parameters α,β,γ, and τ set to 0.01, arbitrarily. We report their run-times in Figure 5(b). The online framework runs approximately %3 faster on average than SocLSFact [A]. Furthermore, it shows 1% higher *F*-measure performance on average than SocLSFact [A] method and 11% higher than SocLSFact on average. Much shorter run-time of SocLSFact method should be noted. However, it is not significant as it factorises a much smaller size of data, and shorter run-time is expected.

Temporal smoothing parameter analysis

In this section, we discuss the effect of the temporal smoothing parameter τ in the online framework. We introduce the parameter τ to weight the importance of previous snapshots' detected implicit links. We expect it to behave as a temporal regularizer in the case of data sparsity and any other type of abrupt changes. To evaluate the behaviour of online

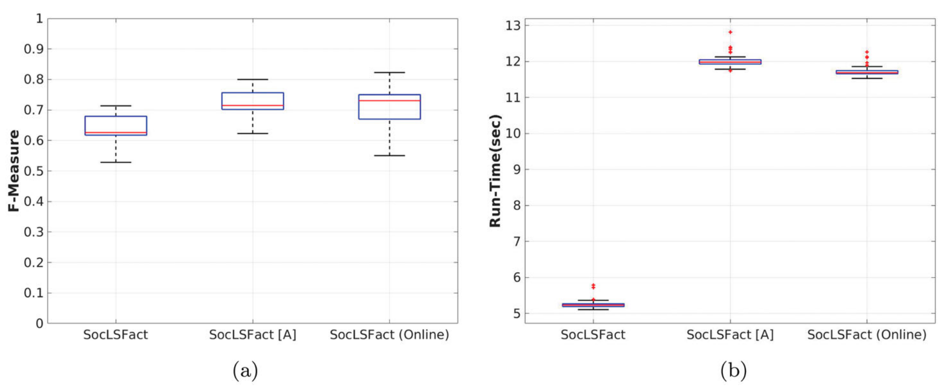


Figure 5. (Colour online) Offline & Online Algorithms' performance comparison for 57th Parliament dataset. Online SocLSFact achieves competetitive performances while having shorter run-times. (a) Performance Comparison, (b) Run-Time Comparison.

framework under different τ's, we present the F-measure performances of different parameter settings in Figure 6. When the value of τ gets larger, variation in the performance due to the positive prior and sentiment lexicon regularizer parameters decrease. Tweaking τ does not improve the F-measure performance under optimal α and β choices, significantly.

6. Applications

In this section, we present two applications for online political networks which would not be possible or as effective without detecting implicit negative links, first. As the first application, we demonstrate the added value of implicit negative links in community detection task. Second, we qualitatively analyse the key role of implicit negative links in disclosing group polarisation dynamics.

6.1. Dataset

Before going into the details of the applications, first, we introduce the datasets we utilise in the application settings. We crawl three datasets from politician accounts of Twitter from United Kingdom, United States, and Canada using GET user_timeline function of Twitter API. They consist of either parliament member accounts of their country or prominent political figure accounts. Each politician account in the dataset either self declares their political party membership in their user profile or has the abbreviation of the political party in their user name as suffix or prefix. Baseline communities are constructed according to each account's self-identification of political party memberships.

- **United Kingdom Dataset** covers Twitter accounts of 421 prominent members of 56th United Kingdom Parliament from 5 major political parties, namely, Conservative Party

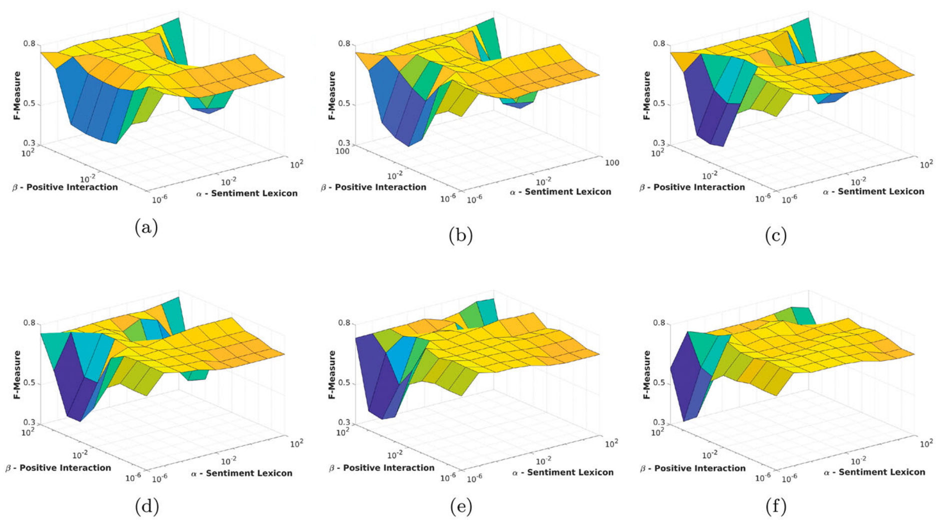


Figure 6. (Colour online) Effect of temporal smoothing parameter τ. Deviation in F-measure decreases with increasing τs. (a) $\tau = 0.0001$, (b) $\tau = 0.001$, (c) $\tau = 0.01$, (d) $\tau = 0.1$, (e) $\tau = 1$, (f) $\tau = 10$.

(Cons), Labour Party (Lab), Scottish National Party(SNP), Liberal Democrats (LibDem), and United Kingdom Independence Party (UKIP).
- **United States Dataset** covers 596 prominent political figures' Twitter accounts from Republican and Democrat Party.
- **Canada Dataset** covers Twitter accounts of 136 members of 41st Parliament of Canada from 5 major political parties, namely, Liberal Party of Canada, Green Party of Canada, Conservative Party of Canada, New Democratic Party, and Bloc Quebecois (BLOC).

Further statistics about the datasets can be found in Table 5, and tweet ids and user ids can be downloaded from www.public.asu.edu/~mozer/NRHMdata.tar.gz.

6.2. Community detection

To evaluate the added value of negative links we test the contribution of negative links in detecting the underlying political communities in the dataset. To that end, we employ a simple spectral clustering algorithm for signed networks. We feed both unsigned links of the given dataset and predicted signed links by our framework SocLS-Fact separately. We employ United Kingdom, Canada and United States datasets to evaluate the performance of our method. Parameters for SocLS-Fact are set to be the ones which minimises the residual error of the objective function.

6.2.1. Spectral clustering on signed networks

As proposed by Kunegis et al. (2010), we define the laplacian matrix $\overline{L}$ of an adjacency matrix A of signed network as;

$$\overline{L} = \overline{D} - A \tag{6.1}$$

where

$$\overline{D}_{ii} = \sum_{j \sim i} |A_{ij}| \tag{6.2}$$

The rest of the clustering framework follows the standard spectral clustering as given in Algorithm 3.

6.3. Evaluation metrics

To evaluate the contribution of predicted negative links in community detection tasks, we make use of two well known clustering quality metrics, namely; purity and normalised mutual information(NMI).

Table 5. Dataset statistics.

	United Kingdom	United States	Canada
Textual interactions	18,903	31,276	5001
Users	400	596	136
Interacting user pairs	3367	6114	1291
Sentiment tokens	1685	1987	1078
# of communities	5	2	5

Algorithm 3: Spectral Clustering Algorithm for Signed and Unsigned Networks

Input: $\bar{L}$ (signed) or L (unsigned).
Output: Clusters $C_1, C_2, \ldots, C_k$.
1 Find the smallest k eigenvalues of $\bar{L}$ (or L).
2 Form matrix U as $[v_1, v_2, \ldots, v_k]$ with corresponding k eigenvectors as columns.
3 Cluster the rows of U into $C_1, C_2, \ldots, C_k$ by applying k-means.

6.4. Community detection results

Table 6 shows the community detection results for United Kingdom, United States and Canada datasets. Inclusion of the predicted negative links of our framework consistently contributes to the performance of community detection tasks.

For experiments having matching k's with number of ground-truth communities of datasets, following observations are made. Significant improvement in all three metrics can be observed in the results of United Kingdom and Canada datasets. United States dataset reveals even more intriguing results: purity increases by %25, and NMI by % 241. This finding suggests that addition of negative links does not only boost the performance but can be of very critical importance for community detection.

Another observation we make is the higher contribution of the predicted negative links in community detection tasks when the number of clusters k given to spectral clustering algorithm is equal to the ground-truth community count of the datasets. The ground-truth community counts for United Kingdom is 5, Canada is 5 and United States is 2 as described in 5.1. Most increase by percentage in all three metrics is achieved when k=5 in United Kingdom and Canada, and $k = 2$ in United States dataset. This further suggests the informativeness of the predicted negative links in implying the exact number of underlying communities.

6.5. Group polarisation

To show another powerful use-case of our framework SocLS-Fact, we set up an experiment that quantifies the group polarisation patterns over time among UK politicians who interact with each other in Twitter. We demonstrate how our method and predicted negative links can be used to represent political dynamics such as emerging and diminishing rivalries or coalitions among political party members. We visualise and qualitatively analyse the detected polarities of links among groups and their change over time.

Table 6. Contribution of detected implicit negative links in community detection tasks with different k's.

k		United Kingdom		Canada		United States	
		Purity	NMI	Purity	NMI	Purity	NMI
2	Unsigned Links	0.4818	0.3829	**0.8013**	**0.5485**	0.7445	0.1863
	SocLS-Fact Links	**0.4844**	**0.4052**	0.7947	0.5057	**0.9294**	**0.6364**
3	Unsigned Links	0.8333	0.6770	**0.9338**	**0.7481**	0.8622	0.3962
	SocLS-Fact Links	**0.8411**	**0.6854**	**0.9338**	0.7473	**0.8807**	**0.4709**
4	Unsigned Links	**0.9167**	0.7838	0.9338	0.7026	0.8605	0.3770
	SocLS-Fact Links	**0.9167**	**0.7859**	**0.9470**	**0.7424**	**0.8773**	**0.4268**
5	Unsigned Links	0.9167	0.7794	0.9272	0.6803	0.8706	0.3935
	SocLS-Fact Links	**0.9427**	**0.8041**	**0.9536**	**0.7456**	**0.8790**	**0.4304**

We sample United Kingdom dataset and create three datasets spanning different time intervals to represent political climate change on social media. First dataset covers the whole timespan which we treat as the overall political climate among members. This dataset constitutes our baseline for detecting divergences from conventional behaviours of political party members in the sampled representative data. Second dataset spans all tweets in 2015. General election held on 5 May 2015 is considered to be the major political event of the year. We refer to the second dataset as general election dataset for future references. Third dataset spans the time interval of first 6 months of the year 2016. Brexit unequivocally being the major political event of that time interval, we refer to the third dataset as Brexit sample for future references.

After sampling these three datasets, we run offline SocLS-Fact algorithm and detect the polarity of each user link. Links that connect users are aggragated with users' affiliated political parties. Aggregation yields the polarisation scores among and within political parties. Positive scores are mapped to hues of greens while negative scores are mapped to reds. Darker color means higher polarity. White color stands for the non-existence or very few links between groups, thus omitted. The overview of the resulting polarity among and within groups for each of the three datasets is presented in Figure 7.

General election dataset

Major event of the 2015 which this dataset covers is the United Kingdom general election 2015 as implied by the popular hashtags presented in Table 7. It took place on May, 5 2015. Conservative Party and Labour Party was the prominent candidates of winning the election. Government before the election was a coalition between Conservative Party and Liberal Democrat Party. Further background information about United Kingdom political parties can be obtained from (Moran, 2011).

Brexit dataset

The biggest political event of the first 6 months of the year 2016 that Brexit Dataset covers, is clearly the European Union (EU) Referandum [Hobolt, 2016] that took place on 23 June 2016. UKIP and some politicians from Conservative party supported leaving the EU. On the opposite side of leave campaign, SNP, Labour Party, Liberal Democrats and part of the

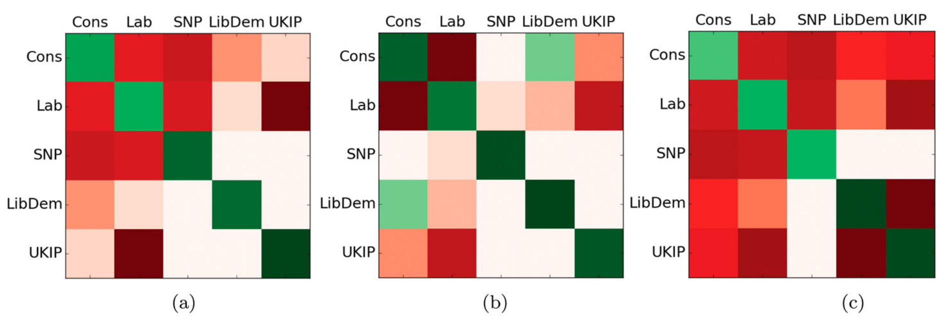


Figure 7. (Colour online) United Kingdom link prediction results for political parties for different time frames. Predicted polarities of user links are aggregated according to users' political party affiliation. Red color implies negative links while green color implies positive links. The darker the color is the higher the polarity is between two parties. (a) Overall Political Climate, (b) General Election Dataset, (c) Brexit Dataset.

Table 7. Popular hashtags in the textual interactions of two samples from United Kingdom dataset.

Sampled datasets	Popular hashtags
General Election	#GE2015, #labourdoorstep, #GE15, #VoteSNP, #Labour, #VoteLabour, #bedroomtax, #NHS, #PMQs, #voteSNP
Brexit	#StrongerIn, #Brexit, #EUref, #VoteLeave, #labourdoorstep, #Remain, #LabourInForBritain, #BackZac2016, #BothVotesSNP, #EU

Consertavite Party were for staying in the EU. UKIP was a prominent political actor in the campaign. As implied by the popular hashtags used in the textual interactions between users, the dataset also covers London mayoral election (i.e. #BackZac2016) and Scottish Parliament Election (#BothVotesSNP). The election in Scotland resulted as a victory for SNP.

6.5.1. Tracking the divergence of political parties from overall behaviour

In this section, we elaborate on how much polarisation between groups deviate from their overall representation in the full dataset. Findings can be summarised as;

- Comparing Figure 7(a,b) shows the increasing positive link ratio in inner-party links. De Nooy & Kleinnijenhuis (2013) suggest that if two politicians belong to the same political party, they are more likely to support each other in an election season as the partisanship increases. The behaviour can be justified with the existence of the general election.

6.5.2. Tracking the temporal dynamics of polarisation among political parties

To evaluate the performance of the tracking the temporal dynamics of polarisation between groups, we qualitatively analyse the polarity shifts from 2015 to 2016 between groups.

- Inner group positive link ratio of Conservative Party members decrease from 2015 (Figure 7(b)) to 2016 (Figure 7(c)) which can be explained by the members of the party diverging apart by having different point of views for EU Referandum.
- The rivalry between Conservative Party and Labour Party members dissolves slightly in 2016, because they were the two most prominent competiters in the general election and considerable amount of two parties' members campaigned for the same voting stance on Brexit election.
- The coalition in 2015 between Conservative Party and Liberal Democrats shifts to rivalry in 2016. It may be due to the coalition government that still existed in 2015 but were not formed again after the election.
- Rivalry increases between UKIP and other parties in Brexit dataset compared to General Election dataset. It can be explained by the EU Referandum in which UKIP was a leading figure.

7. Conclusion

In this paper, we study rivalries, antagonisms, or disagreements in the form of implicit negative links in online political networks and their use-cases. First, we propose an implicit negative link detection framework that performs well on online political networks in which no platform-specific negative interactions or explicit negative links between users are present. Our framework presents robust performance under various parameter settings. It makes the framework applicable to the datasets where there is no ground-truth labels to train a supervised negative link detection model. Then, we propose an online variant of the offline SocLS-Fact framework to allow efficient computation for streaming data. We investigate the effectiveness of online implicit negative link detection and argue that online framework can achieve similar performance in shorter run-times.

We introduce a use-case of implicit negative links on online political networks, namely, community detection. By utilising three different political Twitter datasets, we show that communities can be unfolded more effectively when implicit negative links are detected first.

We believe the existence of implicit negative links in online political networks is evident and effective detection of it can facilitate further applications. Our work stands out as being the first effort in this direction and by providing researchers a rare dataset of annotated implicit negative links.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by the Office of Naval Research Global under Grant N00014-16-1-2015; and Office of Naval Research under Grant N00014-15-1-2722.

ORCID

Mert Ozer http://orcid.org/0000-0003-2687-6783

References

Aktunc, R., Ozer, M., Toroslu, I. H., & Davulcu, H. (2015). A dynamic modularity based community detection algorithm for large-scale networks: DSLM. *Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining 2015 (ASONAM '15)*, Paris, France (pp. 1177–1183).

Ausserhofer, J., & Maireder, A. (2013). National politics on twitter. *Information, Communication and Society*, *16*(3), 291–314.

Conover, M. D., Goncalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F (2013). Predicting the political alignment of twitter users. *2011 IEEE third international conference on privacy, security, risk and trust and 2011 ieee third international conference on social computing*, Boston, MA, USA (pp. 192–199).

Conover, M. D., Goncalves, B., Ratkiewicz, J., Francisco, M., Flammini, A., & Menczer, F. (2011). Political polarization on twitter. *Proceedings of the 5th international conference on Weblogs and social media*, Barcelona, Spain (pp. 89–96).

De Nooy, W., & Kleinnijenhuis, J. (2013). Polarization in the media during an election campaign: A dynamic network model predicting support and attack among political actors. *Political Communication*, *30*(1), 117–138.

Ding, C., Li, T., Peng, W., & Park, H. (2006). Orthogonal nonnegative matrix t-factorizations for clustering. *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '06)*, Philadelphia, PA, USA (pp. 126–135).

DuBois, T., Golbeck, J., & Srinivasan, A. (2011). Predicting trust and distrust in social networks. *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, Barcelona, Spain (pp. 418–424).

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382.

Hassan, A., Abu-Jbara, A., & Radev, D. (2012). Extracting signed social networks from text. *Workshop proceedings of TextGraphs-7 on graph-based methods for natural language processing* (pp. 6–14).

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Psychology Press.

Hobolt, S. B. (2016). The Brexit vote: A divided nation, a divided continent. *Journal of European Public Policy*, *23*(9), 1259–1277.

Hu, X., Tang, J., Gao, H., & Liu, H. (2013). Unsupervised sentiment analysis with emotional signals. *Proceedings of the 22nd international conference on world wide web*, Rio de Janeiro, Brazil (pp. 607–618).

Hu, X., Tang, L., Tang, J., & Liu, H. (2013). Exploiting social relations for sentiment analysis in microblogging. *Proceedings of the sixth ACM international conference on web search and data mining*, Rome, Italy (pp. 537–546).

Johnson, K., & Goldwasser, D. (2016). All I know about politics is what I read in Twitter: Weakly supervised models for extracting politicians' stances from Twitter. *Proceedings of COLING 2016*, Osaka, Japan (pp. 2966–2977).

Kunegis, J., Preusse, J., & Schwagereit, F. (2013). What is the added value of negative links in online social networks? *Proceedings of the 22nd international conference on world wide web*, Rio de Janeiro, Brazil (pp. 727–736).

Kunegis, J., Schmidt, S., Lommatzsch, A., Lerner, J., De Luca, E. W., & Albayrak, S. (2010). Spectral analysis of signed graphs for clustering, prediction and visualization. *Proceedings of the 2010 SIAM international conference on data mining*, Columbus, OH, USA (pp. 559–570).

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010a). Predicting positive and negative links in online social networks. *Proceedings of the 19th international conference on world wide web*, Raleigh, NC, USA (pp. 641–650).

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010b). Signed networks in social media. *Proceedings of the SIGCHI conference on human factors in computing systems*, Atlanta, GA, USA (pp. 1361–1370).

Li, T., Zhang, Y., & Sindhwani, V. (2009). A non-negative matrix tri-factorization approach to sentiment classification with lexical prior knowledge. *Proceedings of the joint conference of the 47th annual meeting of the ACL and the 4th international joint conference on natural language processing of the AFNLP*, Suntec, Singapore (Vol. 1, pp. 244–252).

Liben-Nowell, D., & Kleinberg, J. (2003). The link prediction problem for social networks. *Proceedings of the twelfth international conference on information and knowledge management*, New Orleans, LA, USA (pp. 556–559).

Liu, H., Morstatter, F., Tang, J., & Zafarani, R. (2016). The good, the bad, and the ugly: uncovering novel research opportunities in social media mining. *International Journal of Data Science and Analytics*, *1*(3-4), 137–143.

Mankad, S., & Michailidis, G. (2013). Structural and functional discovery in dynamic networks with non-negative matrix factorization. *Physical Review E*, *88*(4), 042812.

Moran, M. (2011). *Politics and Governance in the UK*. New York, NY: Palgrave Macmillan.

Ozer, M., Kim, N., & Davulcu, H. (2016). Community detection in political Twitter networks using Nonnegative Matrix Factorization methods. *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, San Francisco, CA, USA (pp. 81–88).

Ozer, M., Yildirim, M. Y., & Davulcu, H. (2017). Negative link prediction and its application in online political networks. *28th ACM conference on hypertext and social media*, Prague, Czech Republic (pp. 125–134).

Pamelee, J. H., & Bichard, S. L. (2011). *Politics and the Twitter revolution: How tweets influence the relationship between political leaders and the public*. Lanham, MD: Lexington Books.

Tang, J., Chang, S., Aggarwal, C., & Liu, H. (2015). Negative link prediction in social media. *Proceedings of the eighth ACM international conference on web search and data mining*, Shanghai, China (pp. 87–96).

Wang, D., Pedreschi, D., Song, C., Giannotti, F., & Barabasi, A. (2011). Human mobility, social ties, and link prediction. *Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining*, San Diego, CA, USA (pp. 1100–1108).

West, R., Paskov, H., Leskovec, J., & Potts, C. (2014). Exploiting social network structure for person-to-person sentiment analysis. *Transactions of the Association for Computational Linguistics*, *0*, 297–310.

Yang, S., Smola, A. J., Long, B., Zha, H., & Chang, Y. (2012). Friend or Frenemy?: Predicting signed ties in social networks. *Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval*, Portland, OR, USA (pp. 555–564).

Yu, W., Aggarwal, C., & Wang, W. (2017). Temporally factorized network modeling for evolutionary network analysis. *Proceedings of the tenth ACM international conference on web search and data mining (WSDM '17)* (pp. 455–464).

Zhu, L., Guo, D., Yin, J., Ver Steeg, G., & Galstyan, A. (2016). Scalable temporal latent space inference for link prediction in dynamic social networks. *IEEE Transactions on Knowledge and Data Engineering*, *28*(10), 2765–2777.

Appendix. Derivation of $\mathbf{S}_{uc}^{(t)}$'s update rule

Objective function with respect to $\mathbf{S}_{uc}^{(t)}$ of the rewritten optimisation formulation of online framework is;

$$\begin{aligned}\min_{\mathbf{S}_{uc}^{(t)}} \quad & -2\mathrm{Tr}(\mathbf{X}^{(t)}\mathbf{S}_w^{(t)}\mathbf{H}^{(t)^T}\mathbf{S}_{uc}^{(t)^T}) + \mathrm{Tr}(\mathbf{S}_{uc}^{(t)}\mathbf{H}^{(t)}\mathbf{S}_{\mathbf{w}}{}^{(t)^T}\mathbf{S}_w^{(t)}\mathbf{H}^{(t)}\mathbf{S}_{uc}^{(t)^T}) \\ & + \beta\,\mathrm{Tr}(\mathbf{S}_{uc}^{(t)^T}\mathbf{D}_u^{(t)}\mathbf{S}_{uc}^{(t)}) - 2\beta\,\mathrm{Tr}(\mathbf{S}_{uc}^{(t)^T}\mathbf{D}_u^{(t)}\mathbf{S}_{u0}^{(t)}) - \gamma\,\mathrm{Tr}(\mathbf{M}^{(t)}\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}) \\ & - \gamma\,\mathrm{Tr}(\mathbf{M}^{(t)^T}\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}) + \gamma\,\mathrm{Tr}(\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}) \\ & + \tau\sum_{i=1}^{t}(\mathrm{e}^{-(t-i)}(-2\mathrm{Tr}(\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(i)^T}) + \mathrm{Tr}(\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)})) - \mathrm{Tr}(\Gamma\mathbf{S}_{uc}^{(t)^T})\end{aligned}$$

where Γ is the Lagrange multiplier for the constraint of $\mathbf{S}_u \geq 0$. The derivative of the objective function with respect to $\mathbf{S}_u$ is;

$$\begin{aligned}\frac{\partial L_{\mathbf{S}_{uc}^{(t)}}}{\partial \mathbf{S}_{uc}^{(t)}} = & -2\mathbf{X}_c^{(t)}\mathbf{S}_w^{(t)}\mathbf{H}^T + 2\mathbf{S}_{uc}^{(t)}\mathbf{H}^{(t)}\mathbf{S}_{\mathbf{w}}{}^T\mathbf{S}_w^{(t)}\mathbf{H}^{(t)} + 2\beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)} - 2\beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc0}^{(t)} \\ & + \gamma(\mathbf{M}_c^{(t)} + \mathbf{M}_c^{(t)^T})\mathbf{S}_{uc}^{(t)} - 2\gamma\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}\mathbf{S}_{uc}^{(t)} - 2\tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(i)} \\ & + 2\tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(t)} - \Gamma\end{aligned}$$

By setting the derivative to 0, we get;

$$\begin{aligned}\Gamma = & -2\mathbf{X}_c^{(t)}\mathbf{S}_w^{(t)}\mathbf{H}^{(t)^T} + 2\mathbf{S}_{uc}^{(t)}\mathbf{H}^{(t)}\mathbf{S}_{\mathbf{w}}{}^{(t)^T}\mathbf{S}_w\mathbf{H}^{(t)} + 2\beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)} - 2\beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc0}^{(t)} \\ & + \gamma(\mathbf{M}_c^{(t)} + \mathbf{M}_c^{(t)^T})\mathbf{S}_{uc} - 2\gamma\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}\mathbf{S}_{uc} - 2\tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(i)} \\ & + 2\tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(t)}\end{aligned}$$

Having Karush Kuhn Tucker (KKT) complementary condition of the nonnegativity of $\mathbf{S}_{uc}^{(t)}$ as $\Gamma_{ij}(\mathbf{S}_{uc}^{(t)})_{ij} = 0$ gives;

$$\begin{aligned}&\left(\mathbf{S}_{uc}^{(t)}\mathbf{H}^{(t)}\mathbf{S}_{\mathbf{w}}{}^T\mathbf{S}_w\mathbf{H}^{(t)} + \beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)} + \gamma(\mathbf{M}_c^{(t)} + \mathbf{M}_c^{(t)^T}) + \tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(t)}\right)_{ij}(\mathbf{S}_{uc}^{(t)})_{ij} \\ & - \left(\mathbf{X}_c^{(t)}\mathbf{S}_w\mathbf{H}^{(t)^T} + \beta\mathbf{D}_{uc}^{(t)}\mathbf{S}_{uc0}^{(t)} + \gamma\mathbf{S}_{uc}^{(t)}\mathbf{S}_{uc}^{(t)^T}\mathbf{S}_{uc}^{(t)} + \tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(i)}\right)_{ij}(\mathbf{S}_{uc}^{(t)})_{ij} = 0\end{aligned}$$

which leads to the update rule of $\mathbf{S}_{uc}^{(t)}$;

$$\mathbf{S}_{uc}^{(t)} \leftarrow \mathbf{S}_{uc}^{(t)} \odot \sqrt{\frac{\mathbf{X}_c\mathbf{S}_w\mathbf{H}^T + \gamma(\mathbf{M}_c + \mathbf{M}_c^T)\mathbf{S}_{uc} + \beta\mathbf{D}_{uc}\mathbf{S}_{uc0} + \tau\sum_{i=1}^{t}\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(i)}}{\mathbf{S}_{uc}\mathbf{H}\mathbf{S}_w^T\mathbf{S}_w\mathbf{H}^T + \gamma\mathbf{S}_{uc}\mathbf{S}_{uc}^T\mathbf{S}_{uc} + \beta\mathbf{D}_{uc}\mathbf{S}_{uc} + \tau t\,\mathrm{e}^{-(t-i)}\mathbf{S}_{uc}^{(t)}}}$$