LEARNING CAUSAL INFORMATION FLOW STRUCTURES IN MULTI-LAYER NETWORKS

Başak Güler *Aylin Yener*

The Pennsylvania State University
University Park, PA
basak@psu.edu *yener@ee.psu.edu*

Ananthram Swami

Army Research Laboratory
Adelphi, MD
a.swami@ieee.org

ABSTRACT

We study causal influence structures between the patterns of a multi-layer network. Multi-layer networks are networks in which different types of activities between users represent different types of edges, i.e., layers. We measure the causal influence between network patterns via directed information, and investigate how to learn the influence patterns when users can engage in interactions in multiple contexts. We evaluate the proposed methods using both synthetic and real-world datasets, and demonstrate that directed information measures can be utilized to identify the causal relations between network structures.

1. INTRODUCTION

Modern networked systems, such as social networks, often allow their users to engage in various types of interactions. For instance, the widely used social networking application Twitter allows its users to *follow* other users' online posts, *reply* to them, and *retweet* their posts, i.e., rebroadcast the message. As such, different types of actions define different network structures.

Network patterns, often called network motifs, are graph structures that are observed frequently in a network [1, 2, 3, 4, 5, 6]. They have been widely utilized to describe patterns in biological systems [7, 8, 9]. Network patterns have also found applications in computer science, such as categorizing network topologies [10]. In [11], pattern evolution in the Google+ social networking application has been studied by using frequency measures.

Directed information [12, 13] extends the conventional notion of information that is symmetric taking into account the direction of information flow, and is instrumental in quantifying the fundamental limits of information transmission in communication networks with feedback [13]. The applications where directed information emerges as the key measure range from hypothesis testing to portfolio theory [14] and from estimating neural signals [15] to influence propagation in recommender systems [16]. In [17], directed information has been utilized to define directed information graphs, which is a graphical representation of the causal dependence structures between random processes, based on an extension of Granger causality [18, 19]. Transfer entropy is utilized in [20] for detecting the causal influences between pairs of social media users. This work considers a different application of directed information with the goal of identifying the *causal* dependency structures between network patterns.

By understanding the causal dependencies between network patterns, one can infer and predict the patterns that can occur in a network, given the behavior of the network patterns in the past. In this work, we propose a methodology to analyze the causality relations between network motifs by using directed information. Our goal is to quantify the causal influence of one network pattern on another, across various action types that the users can take in a complex network. Our work is focused on triangular patterns, i.e., network patterns that occur between three users. Initially, we introduce a graph mining algorithm to identify the graph patterns across different network layers and time-instances in an evolving multi-layer network. Then, we propose a methodology to identify the causal influence between network patterns, that is based on quantifying the causal influence between motifs that result from the interactions within the same group of users across multiple time-instances. Our evaluations on a synthetically created network and a real-world dataset from Twitter indicate the effectiveness of the proposed measure, in that asymmetric information measures such as directed information can be leveraged to infer the causal influences between network patterns.

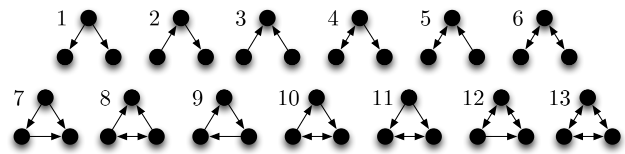


Fig. 1. Non-isomorphic triangular motifs.

2. PATTERN DISCOVERY IN EVOLVING MULTI-LAYER NETWORKS

Our study is focused on identifying causal relationships between non-isomorphic directed triangular motifs with no isolated vertices in Fig. 1 in evolving multi-layer networks. A multi-layer graph is a graph in which vertices can be connected by multiple edges, corresponding to different types of relationships that may take place between them. We focus on networks that evolve over time, depending on the different types of actions the users take. Fig. 2 illustrates an example for a network of 5 people that can communicate with each other through *e-mail, phone*, or *social media*. Each edge in the multi-layer graph identifies who communicated with whom, through one of the three communication types. Each action type, *e-mail, phone*, or *social media*, specifies a distinct network layer. Fig. 2 illustrates the network for two time instances, denoted by $t = 1$ and $t = 2$.

We consider a network of n nodes, with $\binom{n}{3}$ possible user triplets. The users may take actions from a finite set of actions $\mathcal{L}$ at each time instance over a duration of T time instances. The relationships represented by each action type defines a distinct network layer. We represent the multi-layer graph with $|\mathcal{L}| = L$ layers by $\mathcal{G} = (V, E_t)$ at time $t = 1, \ldots, T$. The vertex set is given by V whereas the edge set is given by $E_t = E_{t1}, \cup, \cdots, \cup E_{tL}$ such that E_{tl} corresponds to the set of edges from layer $l \in \mathcal{L}$. The set of

This research is sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-2-0053.

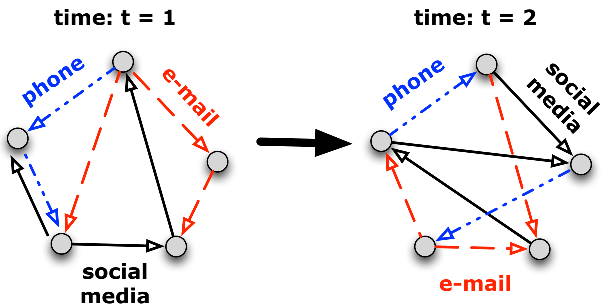


Fig. 2. Multi-layer graph example.

directed triangle motifs is given by $\mathcal{M} = \{1, \dots, 13\}$. In order to identify whether a certain network pattern, i.e., a motif, is observed between a given triplet of users at a specific time instance, we define the following function,

$$e_{it}(v) = 1$$

if pattern $i \in \mathcal{M}$ is observed between user triplet $v \in \{1, \dots, \binom{n}{3}\}$ at time $t \in \{1, \dots, T\}$, whereas

$$e_{it}(v) = 0$$

if pattern $i \in \mathcal{M}$ is not observed between the user triplet v. Given a motif $i \in \mathcal{M}$ in layer $l \in \mathcal{L}$ and a motif $j \in \mathcal{M}$ in layer $k \in \mathcal{L}$, let

$$n_{ij}^{lk}(x^T, y^T) = \sum_{v=1}^{\binom{n}{3}} \mathbf{1}(e_{it}(v) = x_t \wedge e_{jt}(v) = y_t, t = 1, \dots, T) \quad (1)$$

where $x^T = (x_1, \dots, x_T)$ and $y^T = (y_1, \dots, y_T)$ are binary sequences of length T, i.e., $x^T, y^T \in \{0,1\}^T$. $\mathbf{1}(\cdot)$ is an indicator function such that $\mathbf{1}(\phi) = 1$ if $\phi = 1$ and $\mathbf{1}(\phi) = 0$ if $\phi = 0$. For instance, if $x^T = y^T = (1, 1, \dots, 1)$, then $n_{ij}^{lk}(x^T, y^T)$ specifies the number of times pattern i is observed in layer l while pattern j is observed in layer k between the same user triplet, at all time-instants.

We define the probability (the relative frequency) of x, y

$$p_{ij}^{lk}(x^T, y^T) = \frac{n_{ij}^{lk}(x^T, y^T)}{\binom{n}{3}} \quad \text{where} \quad x^T, y^T \in \{0,1\}^T \quad (2)$$

for each $ij \in \mathcal{M}$ and $l, k \in \mathcal{L}$, noting that (2) corresponds to an empirical distribution. It can be observed from (2) that, $p_{ij}^{lk}(x^T, y^T) \geq 0$ for all $x^T, y^T \in \{0,1\}^T$, and

$$\sum_{x^T, y^T \in \{0,1\}^T} p_{ij}^{lk}(x^T, y^T) = 1. \quad (3)$$

We determine the co-occurances of pairs of motifs in (1) across different network layers and multiple time-instances via the pattern mining algorithm in Algorithm 1.

3. IDENTIFYING CAUSAL RELATIONS BETWEEN NETWORK MOTIFS

We explain in this section how directed information can be used to detect the causal dependencies between network patterns. Suppose we are given a pattern $i \in \mathcal{M}$ for layer $l \in \mathcal{L}$, and a pattern $j \in \mathcal{M}$ for layer $k \in \mathcal{L}$. Then, we define the sequence of random variables $X_{il}^T = (X_{il,1}, \dots, X_{il,T})$ and $X_{jk}^T = (X_{jk,1}, \dots, X_{jk,T})$ where each element $X_{il,t}, X_{jk,t}, t = 1, \dots, T$ is defined over a binary alphabet whose joint distribution is specified by (2). We then measure the causal influence of pattern j at layer k on pattern i at layer l by the directed information from X_{jk}^T to X_{il}^T,

Algorithm 1 Pattern Mining for Evolving Multi-Layer Networks

procedure PATTERNDISCOVERY (i, j, k, l, T) ▷ Evaluate the motif counts from (1) for motif i in layer l and motif j in layer k over a duration of T time instances.
- Construct a union graph $H = (V, E)$ with the vertex set V and edge set $E = \cup_{t=1}^T (E_{lt} \cup E_{kt})$.
- $\mathcal{S} = \emptyset$ ▷ Set of processed edges.
- $n_{ij}^{lk}(x^T, y^T) = 0$ for all $i, j \in \mathcal{M}$ and $x^T, y^T \in \{0,1\}^T$ ▷ Counts from (1) for motifs i and j in layers l and k, respectively.
- **for all** (v, v') in E **do**
  - $\mathcal{S} = \mathcal{S} \cup \{(v, v')\}$
  - $\mathcal{U} = \{u : (v,u) \in E \vee (u,v) \in E \vee (v',u) \in E \vee (u,v') \in E,\ u \neq v, v'\}$ ▷ Set of all predecessor/successors of nodes v and v'
  - **if** $\mathcal{U} \neq \emptyset$ **then**
    - **for all** u in $\mathcal{U}$ **do**
      - **if** $(v,u) \notin \mathcal{S} \wedge (u,v) \notin \mathcal{S} \wedge (v',u) \notin \mathcal{S} \wedge (u,v') \notin \mathcal{S} \wedge (v',v) \notin \mathcal{S}$ **then**
        - **for all** $t \in \{1, \dots, T\}$ **do**
          - **for all** m is a permutation of (v, v', u) **do**
            - **if** m is pattern of type i in layer l, time t **then**
              - $M(i,t) = 1$
              - **break**
            - **else**
              - $M(i,t) = 0$
        - **for all** $t \in \{1, \dots, T\}$ **do**
          - **for all** m is a permutation of (v, v', u) **do**
            - **if** m is pattern of type j in layer l, time t **then**
              - $M(j,t) = 1$
              - **break**
            - **else**
              - $M(j,t) = 0$
        - **for all** $(x^T, y^T) \in \{0,1\}^T \times \{0,1\}^T$ **do**
          - **if** $M(i,t) == x_t \wedge M(j,t) == y_t \ \forall t$ **then**
            - $n_{ij}^{lk}(x^T, y^T) = n_{ij}^{lk}(x^T, y^T) + 1$
- **return** $n_{ij}^{lk}(x^T, y^T)$ for all $x^T, y^T \in \{0,1\}^T$

$$I(X_{jk}^T \to X_{il}^T) := \sum_{t=1}^T I(X_{jk}^{t-1}; X_{il,t} | X_{il}^{t-1}) \quad (4)$$

$$= \sum_{t=1}^T \sum_{\substack{x_{jk}^{t-1} \in \{0,1\}^{t-1} \\ x_{il}^t \in \{0,1\}^t}} p(x_{jk}^{t-1}, x_{il}^t) \log \frac{p(x_{il,t} | x_{jk}^{t-1}, x_{il}^{t-1})}{p(x_{il,t} | x_{il}^{t-1})} \quad (5)$$

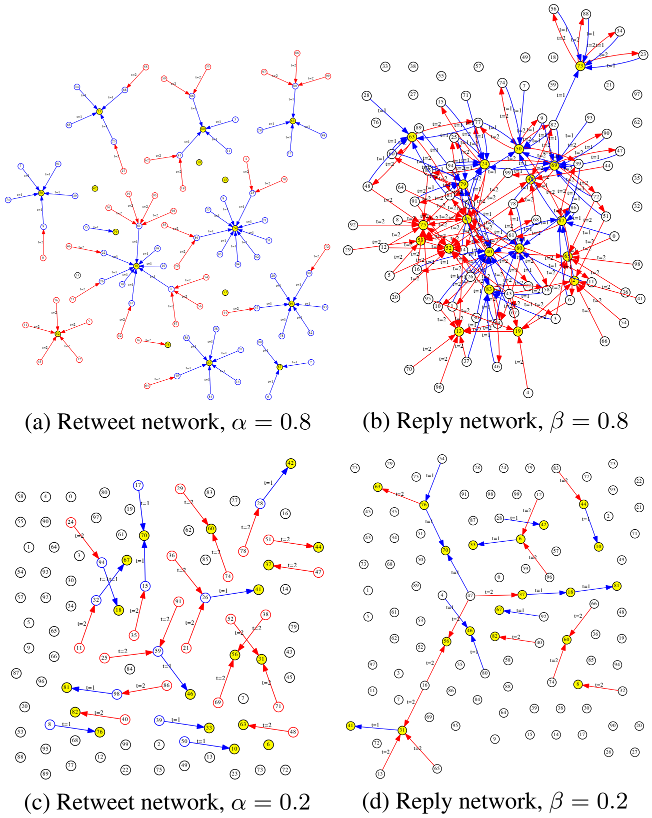
Using (4), we can define a multi-layer directed information graph for identifying the causal independency structures between a group of network motifs as follows. For each network layer $k \in \mathcal{L}$, define a graph $G_k = (\mathcal{M}, \mathcal{E}_k)$ with the vertex set $\mathcal{M}$, and the edge set $\mathcal{E}_k = \mathcal{E}_{k1} \cup \ldots \cup \mathcal{E}_{kL}$ where an edge $(j, i) \in \mathcal{E}_{kl}$ for $l \in \mathcal{L}$ exists if and only if

$$I(X_{jk}^T \to X_{il}^T || X_{\mathcal{M}\setminus\{i\}l}^T, X_{\mathcal{M}\setminus\{j\}k}^T) > 0 \quad (6)$$

where $i, j \in \mathcal{M}$, such that

$$I(X_{jk}^T \to X_{il}^T || X_{\mathcal{M}\setminus\{i\}l}^T, X_{\mathcal{M}\setminus\{j\}k}^T) := \sum_{t=1}^T I(X_{jk}^{t-1}; X_{il,t} | X_{il}^{t-1}, X_{\mathcal{M}\setminus\{i\}l}^{t-1}, X_{\mathcal{M}\setminus\{j\}k}^{t-1}) \quad (7)$$

Here $X_{\mathcal{M}\setminus\{i\}l}^T = \{X_{sl}^T : s \in \mathcal{M}, s \neq i\}$ and similarly $X_{\mathcal{M}\setminus\{j\}k}^T = \{X_{sk}^T : s \in \mathcal{M}, s \neq j\}$. In particular, an edge $(j, i) \in \mathcal{E}_{kl}$ in graph G_k means that patterns of type j in network layer k causally



(a) Retweet network, $\alpha = 0.8$ (b) Reply network, $\beta = 0.8$

(c) Retweet network, $\alpha = 0.2$ (d) Reply network, $\beta = 0.2$

Fig. 3. Retweet and reply networks for $n = 100$ users. Edges created at $t = 1$ and $t = 2$ are colored in blue and red, respectively. Users posting a new message are represented by yellow nodes.

influence patterns of type i in layer l, when causally conditioned on all the remaining patterns on both graphs. Since each network layer corresponds to a specific type of action that may take place between the users, the multi-layered structure of the influence graph identifies what types of actions have the greatest impact on the network patterns created by a specific type of action. Due to the computational complexity involved in identifying multiple patterns jointly in a multi-layer setting to estimate the causally conditioned directed information values in (6), we define a simpler influence graph by utilizing the pairwise directed information measures between different network patterns. To do so, we replace the edge test in (6) with the pairwise directed information measure from (4), such that an edge $(j, i) \in \mathcal{E}_{kl}$ for $l \in \mathcal{L}$ and $j, i \in \mathcal{M}$ exists if and only if

$$I(X_{jk} \to X_{il}) > 0. \quad (8)$$

An early work on using pairwise information measures to create a graphical representation of causal influences is [20], where transfer entropy is utilized to detect causal influences between pairs of social media users.

4. NUMERICAL RESULTS

4.1. Synthetic Network

We first provide numerical evaluations on a synthetic network of n nodes. We consider networks with two layers, i.e., $L = 2$, corresponding to two types of actions common in social network applications. The first action type corresponds to *broadcasting* a message to a group of friends, such as posting a tweet or retweeting another user's post on Twitter. The second action corresponds to replying to a message from another user. In the former case every user who follows the posts of the sender gets notified by the application, whereas in the latter, only the target of the reply message gets notified. In the following, we consider two message propagation models corresponding to these two action types, inspired by the retweeting and replying behavior on Twitter.

We define a directed graph $G = (V, E_G)$ with the vertex set $V = \{1, \dots, n\}$. We consider for each $v, v' \in V$, an edge $(v, v') \in E_G$ exists with probability ρ. An edge $(v, v') \in E_G$ means that user v follows the messages posted by v'. We let $\mathcal{N}(v) = \{v' : (v', v) \in E_G\}$ denote the set of followers of node $v \in V$. We focus on propagation of a single topic and assume that $\lfloor \lambda n \rfloor$ randomly selected users post a new message about this topic at each time instant $t \in \{1, \dots, T\}$, where $0 \leq \lambda \leq 1$. Users who have already posted a message do not post a new message again.

We then model the retweet behavior as follows. At time $t + 1$, each message posted by one of the users at time t is retweeted by a random α fraction of their followers. In particular, if a user $v \in V$ shares a message at time t, either by posting a new message or retweeting another user's message, then, $\lfloor \alpha |\mathcal{N}(v)| \rfloor$ of her friends retweet the same message at time $t + 1$. In that sense, $0 \leq \alpha \leq 1$ specifies the propagation speed of the retweet behavior. Users who have already posted or retweeted a message do not retweet again.

We next model the reply behavior. Each new message posted at time t gets replied by a randomly selected β fraction of the neighbors at time $t + 1$. That is, if $v \in V$ posts a new message at time t, then $\lfloor \beta |\mathcal{N}(v)| \rfloor$ of her followers reply to her at time $t + 1$. At time $t + 2$, the same user replies back to a β fraction of the followers who have replied to her in the previous round, i.e., at time $t + 1$. This iteration continues until there are no users to reply to. In that sense, by varying $0 \leq \beta \leq 1$, one can control the density of the reply behavior.

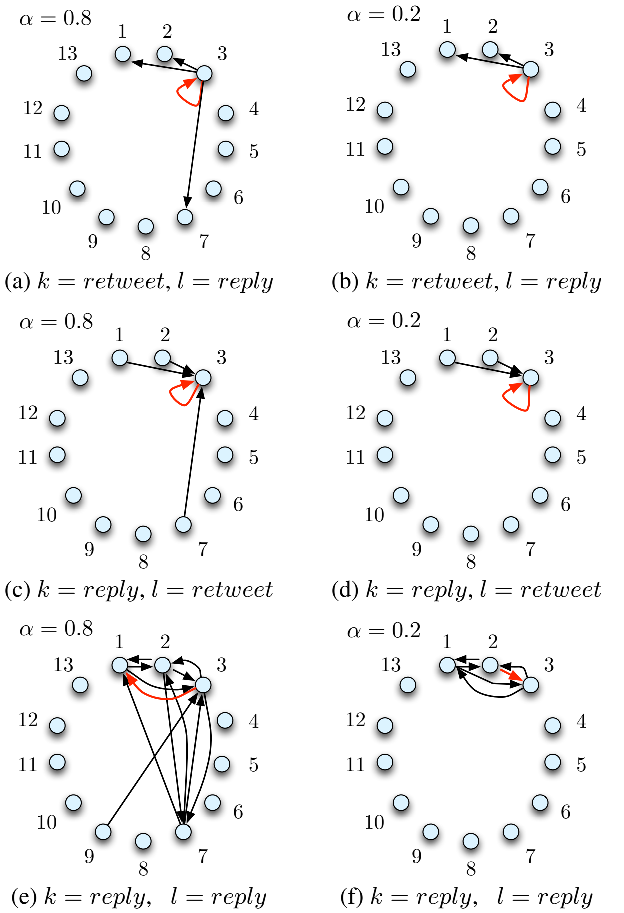
The retweet and reply processes are illustrated in Fig. 3 for a network of $n = 100$ users, $\rho = 0.1$, and $\lambda = 0.1$, for $\alpha = \beta = 0.8$ and $\alpha = \beta = 0.2$, respectively. An edge from node v to node v' with a label $t = t'$ in Fig. 3 implies that user v has retweeted/replied to the post of user v' at time t'.

In our analysis, we simulate a network of $n = 1000$ users and determine the network patterns for two cases, a slow propagation scheme where $\alpha = \beta = 0.2$, and a fast propagation scheme where $\alpha = \beta = 0.8$, by letting $T = 2$, $\lambda = 0.01$, and $\rho = 0.1$. We construct a multi-layer network with two layers corresponding to the retweet and reply behaviour, i.e., $|\mathcal{L}| = 2$, such that a directed edge (v, v') in the retweet layer implies that user $v \in V$ has retweeted a post from user $v' \in V$, whereas an edge (v, v') in the reply layer implies that user v has replied to user v'. We then identify the directed triangular patterns in Fig.1 via Algorithm 1.

Fig. 4 the provides graphical representations of causal influence structures between graph motifs that we have obtained using the edge test in (8). In Fig. 4 (a)-(f), each graph node corresponds to a distinct motif in $\mathcal{M}$, and we draw an edge from motif $j \in \mathcal{M}$ to motif $i \in \mathcal{M}$ whenever $I(X^T_{jk} \to X^T_{il}) > 0$ for the given pair of network layers $k, l \in \mathcal{L}$. In that sense, each edge indicates that motif j from layer k causally influences motif i in layer l. Accordingly, Fig. 4 (a)-(b) correspond to the case when $k = retweet$, $l = reply$, whereas Fig. 4 (c)-(d) correspond to the case when $k = reply$, $l = retweet$, and Fig. 4 (e)-(f) to the case when $k = reply$, $l = reply$. In each graph, the edge (i^*, j^*) from node i^* to node j^* with

$$(i^*, j^*) = \arg \max_{i,j \in \mathcal{M}} I(X^T_{jk} \to X^T_{il}) \quad (9)$$

is colored in red, representing the motif pair with the highest causal influence level. We observe from Fig. 4 (a)-(b) that motif 3 in the



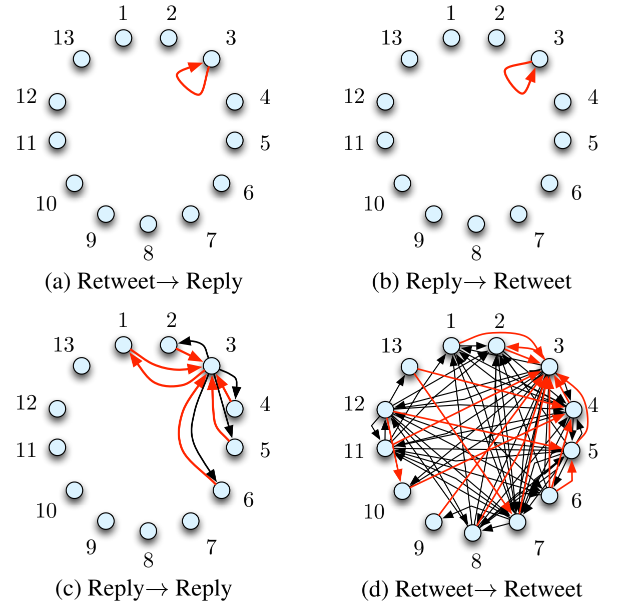
(a) $k = retweet, l = reply$ (b) $k = retweet, l = reply$

(c) $k = reply, l = retweet$ (d) $k = reply, l = retweet$

(e) $k = reply, \quad l = reply$ (f) $k = reply, \quad l = reply$

Fig. 4. Causal influences between motifs, where an edge from node j to node i implies that motif j in layer k causally influences motif i in layer l. A red edge denotes the motif pair with the highest influence.

retweet network causally influences motif 3 and motif 1 in the reply network. The intuition behind this is as follows. Motif 3 in the retweet graph implies that a message that was posted by a person at $t = 0$ is retweeted by her friends at time $t = 1$. This can be observed from Fig. 1 by noting that the top node has two incoming edges from the bottom nodes, indicating that the bottom left and bottom right nodes have retweeted the post of the top node. In addition to retweeting the post from $t = 0$, one can expect that those friends may also have replied to it at time $t = 1$. The user will then reply back to her friends from the previous round, leading to the pattern illustrated by motif 1 in the reply graph at time $t = 2$. Hence, observing motif 3 in the retweet graph implies that it is likely that motif 1 will be observed between these users in the next time slot. Hence, motif 3 in the retweet layer causally influences motif 1 in the reply layer. In addition, we also observe that if motif 3 is observed in the retweet graph at time $t = 1$, then it is unlikely that it will also be observed in the reply graph at time $t = 2$, since the reply pattern at time $t = 2$ will be motif 1. Hence, motif 3 in the retweet layer causally influences motif 3 in the reply layer. We also point out that the reason why the same reasoning cannot be made for the remaining motifs in the reply graph, such as motif 13, is that their frequency is very low whether or not motif 3 is observed in the retweet graph, hence their absence is not due to the occurrence of motif 3 in the retweet layer. Similarly, we observe from Fig. 4 (c)-(d) that motif 3 is causally influenced by motifs 1 and 3. We also observe from Fig. 4 (e) that motif 3 has the highest causal influence on motif 1



(a) Retweet→ Reply (b) Reply→ Retweet

(c) Reply→ Reply (d) Retweet→ Retweet

Fig. 5. Causal influences between the motifs in the Twitter network. For each pattern $i \in \mathcal{M}$, we show the patterns that receive positive influence from i, with the highest one indicated by a red edge.

in the reply network, which results from the reciprocative nature of the replying behavior that can be observed from Fig. (3). That is, at one time instance friends reply to a message, and in the next time-instance the owner of the post replies to the friends, which explains the causal influence from motif 3 to motif 1 and from motif 1 to motif 3. We also observe that fewer causal influences are identified as the network becomes more sparse, i.e., when $\alpha = \beta = 0.2$ than the case for $\alpha = \beta = 0.8$.

4.2. Real-world Network

In order to study the causal network pattern dependencies in a real-world network, we consider the Higgs Twitter dataset from [21] that is a collection of messages related to the discovery of the Higgs boson during the period between July 1, 2012 and July 7, 2012. The dataset consists of $n = 456,631$ nodes and $14,855,875$ edges in the follower network. The network also specifies retweet and reply actions with respective message timings. In our analysis, we initially divide the timeline into two time-slots with respect to the median of the tweet timings. Then, we identify the motif patterns and influence structures using Algorithm 1 and (8), respectively. The results are provided in Fig. 5, from which we observe that the results from the real-world dataset are similar to the synthetic propagation models from Section 4.1. This also implies that judicious selection of synthetic propagation schemes can be utilized to understand the behavior of real-world network patterns.

5. CONCLUSIONS

We have studied causal dependency structures between network patterns in a multi-layer network, where each layer corresponds to a different relationship type. We have proposed a method for identifying causal motif relationships using directed information, and provided numerical evaluations using both a synthetic network and a real-world dataset. Future directions include lower complexity pattern analysis and directed information estimation. Our techniques can also be extended to beyond triangular motifs to study causal dependencies between subgraphs of different sizes.

6. REFERENCES

[1] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[2] F. Schreiber and H. Schwöbbermeyer, "Frequency concepts and pattern detection for the analysis of motifs in networks," in *Transactions on Computational Systems Biology III*, pp. 89–104. Springer, 2005.

[3] M. Mayo, A. F. Abdelzaher, E. J. Perkins, and P. Ghosh, "Motif participation by genes in E. coli transcriptional networks," *Frontiers in physiology*, vol. 3, pp. 357, 2012.

[4] S. Chu and J. Cheng, "Triangle listing in massive networks," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 4, pp. 17, 2012.

[5] P. Wang, J. Lui, B. Ribeiro, D. Towsley, J. Zhao, and X. Guan, "Efficiently estimating motif statistics of large networks," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 9, no. 2, pp. 8, 2014.

[6] R. N. Lichtenwalter and N. V. Chawla, "Vertex collocation profiles: subgraph counting for link analysis and prediction," in *International Conference on World Wide Web (WWW)*. ACM, 2012, pp. 1019–1028.

[7] U. Alon, "Biological networks: the tinkerer as an engineer," *Science*, vol. 301, no. 5641, pp. 1866–1867, 2003.

[8] R. Sharan and T. Ideker, "Modeling cellular machinery through biological network comparison," *Nature Biotechnology*, vol. 24, no. 4, pp. 427–433, 2006.

[9] S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, "Network motifs in the transcriptional regulation network of Escherichia coli," *Nature Genetics*, vol. 31, no. 1, pp. 64–68, 2002.

[10] K. Juszczyszyn, P. Kazienko, and K. Musiał, "Local topology of social network based on motif analysis," in *Knowledge-based Intelligent Information and Engineering Systems*. Springer, 2008, pp. 97–105.

[11] D. Schiöberg, F. Schneidery, S. Schmid, S. Uhlig, and A. Feldmann, "Evolution of directed triangle motifs in the Google+ OSN," *arXiv:1502.04321*, 2015.

[12] H. Marko, "The bidirectional communication theory–A generalization of information theory," *IEEE Transactions on Communications*, vol. 21, no. 12, pp. 1345–1351, 1973.

[13] J. Massey, "Causality, feedback and directed information," in *International Symposium on Information Theory and Its Applications*, 1990, pp. 303–305.

[14] H. H. Permuter, Y.-H. Kim, and T. Weissman, "Interpretations of directed information in portfolio theory, data compression, and hypothesis testing," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3248–3259, 2011.

[15] C. J. Quinn, T. P Coleman, N. Kiyavash, and N. G. Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *Journal of computational neuroscience*, vol. 30, no. 1, pp. 17–44, 2011.

[16] B. Guler, K. Tutuncuoglu, and A. Yener, "Maximizing recommender's influence in a social network: An information theoretic perspective," in *IEEE Information Theory Workshop*, 2015, pp. 262–266.

[17] C. J. Quinn, N. Kiyavash, and T. P. Coleman, "Directed information graphs," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6887–6909, 2015.

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

[19] M. Eichler, "Graphical modelling of multivariate time series," *Probability Theory and Related Fields*, vol. 153, no. 1-2, pp. 233–268, 2012.

[20] G. Ver Steeg and A. Galstyan, "Information transfer in social media," in *International Conference on World Wide Web (WWW)*, 2012, pp. 509–518.

[21] M. De Domenico, A. Lima, P. Mougel, and M. Musolesi, "The anatomy of a scientific rumor," *Scientific Reports*, vol. 3, 2013.