Temporal Scale of Processes in Dynamic Networks

Rajmonda Sulo Caceres *,Tanya Berger-Wolf † and Robert Grossman ‡
* † *University of Illinois at Chicago*
Email: {rsulo1, tanyabw}@uic.edu
‡*University of Chicago*
Email:grossman@labcomputing.org

***Abstract*—Temporal streams of interactions are commonly aggregated into dynamic networks for temporal analysis. Results of this analysis are greatly affected by the resolution at which the original data are aggregated. The mismatch between the inherent temporal scale of the underlying process and that at which the analysis is performed can obscure important insights and lead to wrong conclusions. To this day, there is no established framework for choosing the appropriate scale for temporal analysis of streams of interactions. Our paper offers the first step towards the formalization of this problem. We show that for a general class of interaction streams it is possible to identify, in a principled way, the inherent temporal scale of the underlying dynamic processes. Moreover, we state important properties of these processes that can be used to develop an algorithm to identify this scale. Additionally, these properties can be used to separate interaction streams for which no level of aggregation is meaningful versus those that have a natural level of aggregation.**

***Keywords*-Dynamic Networks; Temporal Scale; Stationary Processes**

I. Introduction

Whether it is the modeling of on-line social interactions, IP packet traces, email and cell phone communications, or protein interactions in a cell, networks have become an indispensable data abstraction that captures the nature of such complex systems. Networks are graphs with nodes representing entities, such as people, computers, or proteins, and the edges representing interactions between pairs of entities, such as sending an email, meeting a person, or proteins participating in the same regulatory process. All these systems are inherently dynamic and change over time. The dynamic nature of the interactions and the processes that happen over the interactions, such as spread of diseases and dissemination of information, needs to be explicitly addressed in the analysis and the abstract representation of those interactions. The abstraction of choice has been the "dynamic network", a time series of graphs, each representing an aggregation of a small discrete time interval of the stream of interactions. While in many cases the system under observation naturally suggests the size of such a time interval (hours, days, years), it is more often the case that the aggregation is arbitrary and is done for the convenience of the data representation and analysis.

Once a dynamic network is constructed, typically it is analyzed by mining its various structures (such as subgraphs, communities, hierarchies) or measuring the dynamics of graph-theoretic properties (such as degree distributions, path lengths, and centrality). However, it is clear that the choice of the time interval at which the network is discretized and aggregated has great implications on the structures observed, analysis performed, and inference made about the nature of the network and the processes on it.

There is a rich body of literature that recognizes the sensitivity of good analytical tools to the size of the discretization step in various domains [1]–[5]. If the aggregation is done at a fine resolution (small window of aggregation), we end up with a network that has lots of temporal detail, yet interesting events at this level are not fully formed and can not be observed. Explicitly, when the time interval of aggregation is too small, then often very few interactions happen within that interval and the resulting dynamic network is sparse or empty and, thus, does not reflect the temporal correlations inherent in the underlying processes. Moreover, the fact that, locally, some interactions are observed happening in a particular order is an artifact of looking at them at too fine of a temporal resolution. They could have happened in any order as long as they happened within a certain time frame. For example, in a communication network, it may be more important the fact that certain emails are sent sometime in the afternoon, but not that they are sent in a particular order; in a social network it may be more significant that a certain group of individuals has met for an hour at 3 pm rather than the micro-ordering of their arrival at the meeting.

On the other end of the spectrum, too big of an aggregation window generates a temporal network where dynamics evolving in time can not be extracted anymore. For example, the chain of communication in an organization cannot be used to infer the hierarchy at this resolution. There is a natural trade-off between the two scenarios, where the goal is to differentiate between temporal associations of interactions that are essentially noise and those associations that are meaningful and informative. We hypothesize that there is a level (or levels) of time aggregation that represents the natural temporal scale of the network, an inherent rhythm at which interesting dynamics of the network become apparent. It is at this scale (or scales) that sets of interactions are strongly correlated with other sets of interactions, and it is

978-0-7695-4409-0/11 $26.00 © 2011 IEEE
DOI 10.1109/ICDMW.2011.165

at this scale that the overlay of the temporal information on the network becomes truly helpful in understanding the structure of the underlying processes.

Even though the discretization step seems to be a critical step to ensure the success of the analytical tools that we develop, too often this step is ignored or there is no systematic process that justifies the choice of a particular window size [1], [6]. Here, we present a formalization of the problem and explore the effect of the aggregation process of the stream of interactions on the resulting network and the analysis on that network.

As we have mentioned, we are interested in the "inherent rhythm" of interactions, something akin to the Fourier analysis for the time series or signal processing. Yet, networks are highly non-linear, non-metric structures. The aggregation of numerical values in a time series does not translate in any known meaningful way to the process of aggregating interactions into a graph. Various graphs theoretic measures and structural properties of these graphs are not simple lossless descriptors of the complex structure of the network and behave unexpectedly different at different temporal scales [7].

In this paper, we will show that for a general class of dynamic networks, we can can indeed use structural measures expressed as linear functions on the edges of the graph to capture important characteristics of its dynamics. More specifically, we analyze the class of dynamic networks generated by an oversampled stationary process that is not complete noise. This probabilistic process is of great interest not only from a theoretical perspective, but practical as well. Meaningful analysis of many processes over physical interactions is done when these processes are stabilized, or have become stationary. Inference about long term trends or typical behavior is not useful if the underlying system is not stable. Also, with the advent of electronic data collection on interactions through communication devices and GPS and proximity sensors, it is more often the case that the data are oversampled at orders of magnitude higher temporal resolution than the temporal scale of the underling processes.

We will in the next sections formally define the notion of a dynamic network as a function of the size of the window of aggregation. As part of the process of formalizing what is interesting in a dynamic network we will define two classes of dynamic networks, one for which there is no optimal window of aggregation, and another for which such an window exists and is identifiable. We will discuss properties of each class and establish some important differences between the two. More specifically, we will show that the behavior of a class of linear functions on edges of the dynamic graph computed at different aggregation levels can be used to distinguish between noisy temporal interactions and structured ones.

The main result of this paper is the following theorem, which we prove in the subsequent sections:

Theorem 1. *Let DG be a dynamic graph which is the result of aggregation over a window ω of a stream of edges, generated by a covariance-stationary process oversampled at a rate of α. Let $F(DG)$ be the time series of a linear function over the edges of DG. Then:*

a) *$F(DG)$ is covariance-stationary when the window of aggregation ω is a multiple of α; and*

b) *there exists ω which is not a multiple of α, s.t. $F(DG)$ is not covariance-stationary.*

II. Related Literature

The problem of identifying the right resolution for analysis of data streams is a very broad problem and covers many research areas such as signal processing [8], [9], discretization of continuous variables [3], time series discretization [4], [5], model granularity [2]. Usually the approach involves a trade-off between loss of information and reduction of noise. While this literature offers a solid foundation on discretization analysis, it does not explicitly address datasets that are represented as networks and, furthermore, it does not address the dynamic nature of these networks. The focus of the work presented here is to develop a framework for identifying the inherent temporal scale for dynamic processes encoded as graph structures that change over time. In addition, it is important to note that the aggregation of temporal edges is an a aggregation of graph structures, not an aggregation of numerical time series. This preserves a lot of the rich structure encoded in the network, but at the same time presents new challenges of how the aggregation affects the structure of the resulting network.

While dynamic networks and their rich temporal structure have gained a lot of interest and motivated a series of informative papers (*e.g.,* [10]–[15] and many more), there is no clear principled framework on how to aggregate temporal graphs in a meaningful way.

Eagle [16], Sun *et al.* [17], and [7] offer empirical ad-hoc approaches to identifying the aggregation level for temporal edges that work for some special cases, but no analysis of generalization or applicability is offered.

This paper is concerned with the theoretical underpinnings of the process of aggregation of a stream of interactions into a dynamic network and the effects of the scale of aggregation on the resulting temporal analysis. We now formally define the objects of study and formalize the problem.

III. Definitions

Here we define the main concepts used throughout the paper. We formalize the notion of a stream of interactions as a probabilistic process. We define two special random interaction streams and networks. We formalize the process of aggregation and the resulting dynamic network as a probabilistic object. We reformulate the general class of linear functions on edges of a graph as a special case of measures on aggregated dynamic networks.

Let V be a set of vertices and E the set of edges defined over $V \times V$. For $\forall e_{ij} \in E, i, j \in V$ and $t \in [1, ..., T]$, the pair (e_{ij}, t) is the time labeled instance of e_{ij}. The *temporal stream of edges* (representing physical interactions) is a pair

$$\langle E, T \rangle = \{(e_{ij}, t)\}, E \subseteq V \times V.$$

Let X_{ijt} be a random variable representing the existence of an edge e_{ij} in the stream $\langle E, T \rangle$ at time t:

$$X_{ijt} = \begin{cases} 1 & \text{if } (e_{ij}, t) \in \langle E, T \rangle \\ 0 & \text{if } (e_{ij}, t) \notin \langle E, T \rangle \end{cases}$$

Definition 1. Dynamic Random Graph ($DynR$**)** *is the probabilistic graph $G(V, \langle E, T \rangle, P)$, defined over the tuple of nodes V, the edge stream on those nodes $(V, \langle E, T \rangle)$ and a probability distribution function P. We use p_{ijt} to denote the probability of an edge in this graph:*

$\forall e_{ij} \in V \times V$ *and* $t \in [1, ..., T], Pr[(e_{ij}, t) \in G)] = p_{ijt}$.

By definition, the probability of an edge in a DynR is $p_{ijt} = P(X_{ijt})$.

Definition 2. Dynamic Uniform Random Graph ($DynUR$**)** *is the graph $G(V, \langle E, T \rangle, p)$ with a constant probability $0 \geq p \leq 1$ for all edges*

$$\forall (e_{ij}, t) \in \langle E, T \rangle \quad Pr[(e_{ij}, t) \in G)] = p.$$

That is, $G(V, \langle E, T \rangle, p)$ is a $DynR$ instance defined over the uniform distribution.

Definition 3. Dynamic Mixture Graph *($DynMix_{M,\{w_l\}}$). Given M probability distributions $\{P_l\}_{l=1}^{M}$ and a set of temporal windows $\{w_l\}_{l=1}^{M}$ such that $W = \sum_{l=1}^{M} w_l$, the* Dynamic Mixture Graph *$DynMix_{M,\{w_l\}}$ is the graph $G(V, \langle E, T \rangle, P_l)$, such that*

$$\forall (e_{ij}, t) \in \langle E, T \rangle, \quad Pr[(e_{ij}, t) \in G)] = p_{lijt}$$

$$\text{where } p_{lijt} = P_l(X_{ijt}) \quad \text{and} \quad l = \bmod_M \left\lfloor \frac{t}{W} \right\rfloor.$$

The Dynamic Mixture Graph is a repeating sequence of Dynamic Graphs. Our paper focuses on two special cases for the probability distribution functions generating the $DynMix_{M,\{w_l\}}$ graph:

Case 1: A sequence of constant probability distribution P_l, so the probability of an edge e_{lij} does not depend on the time index t:

$$Pr[(e_{ij}, t) \in DynMix_{M,\{w_l\}} | (e_{ij})] = p_{lij}.$$

Figure 1(a) gives an illustration of such a Dynamic Mixture Graph when the number of repeating probability distributions M is 3.

Case 2: A sequence of DynURs, where for any given probability distribution P_l, and a given time index t, the probability of all edges e_{lij} at t is the same:

$$Pr((e_{ij}, t) \in DynMix_{M,\{w_l\}} | t) = p_t \quad \forall i, j \in V.$$

Figure 1(b) gives an illustration of such a Dynamic Mixture Graph with M=3. Note that $DynUR$ can be viewed as a special instance of both Case 1 and Case 2 of the

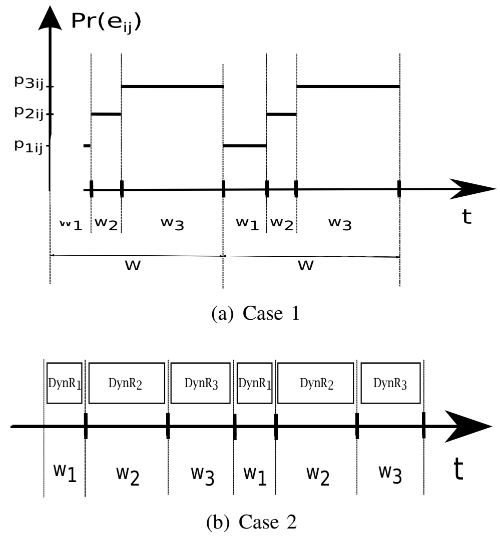


Figure 1. Illustration of DynMix graphs with M=3 for (a), (b)

$DynMix_{M,\{w_l\}}$ graph with $p_{lij} = p_t = p$ for all tuples $\{l, i, j, t\}$.

We now define the operation of aggregation of a temporal stream of edges into a time series of graphs comprising a dynamic network. Given the temporal stream of edges $\langle E, T \rangle$, and a fixed window of aggregation w, we define the aggregation function that takes as an input the temporal stream of edges and outputs a time-series of graphs.

Definition 4. *For a fixed aggregation window ω, an aggregation function $\mathcal{A}$ on a temporal stream $\langle E, T \rangle$ is defined as:*

$$A : \langle E, T \rangle \times \mathbb{R}^+ \rightarrow \langle V, \langle E, T \rangle \rangle,$$

$$\mathcal{A}(\langle E, T \rangle, \omega) = \langle G_0, G_1 ..., G_k, ...G_{\frac{T}{\omega}-1} \rangle,$$

where each graph $G_k = (V, \langle E_k, T \rangle)$, and $\langle E_k, T \rangle = \{(e_{ij}, t) | e_{ij} \in E, k\omega \leq t < (k+1)\omega\}$.

The aggregation process takes all edges occurring in a stream within a time interval ω and constructs a graph. The dynamic network, then, is a time series of such graphs. Figure 2 shows an aggregation function over the window of aggregation of 2. Note, that edges can occur within each temporal window more than once, but are represented in the corresponding aggregated graph at most once. Linear functions defined over the edges of a graph $G(V, E)$ are of particular interest when analyzing the graph's structural properties. Let $f : E \rightarrow \mathbb{R}^+$ be such a function:

$$f = \sum_{i,j \in V} a_{ij} X_{ij},$$

where X_{ij} is the event of an edge e_{ij} being present in the graph. The density of a graph G is an example of such function f. In this case, $a_{ij} = 1/\binom{n}{2}$ for all edges e_{ij}, where $n = |V|$. Other graph measures on graphs, can be defined similarly and are of great interest when studying graphs that evolve in time.

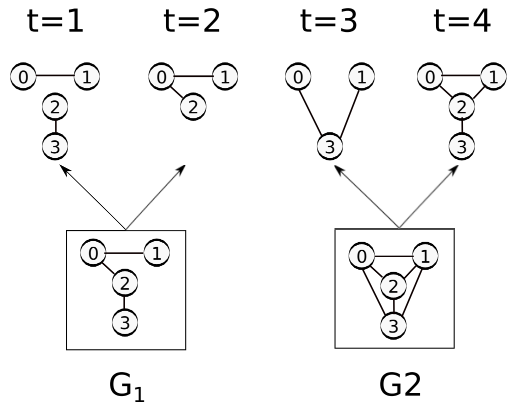


Figure 2. Aggregation of temporal edges with aggregation window $\omega = 2$

IV. Temporal scale of oversampled, noisy stationary processes of edges

Let $P(X_{ijt}) = P_t$ be the general case of a (weak) stationary probability distribution function generating the stream of edges $\langle E, T\rangle = \{(e_{ij}, t)\}, t \in [1, ..., T]$. That is, $\mathbb{E}[p_{ijt}] = \mu_{ij}, Cov(p_{ijt}, p_{ij(t+\tau)}) = \gamma_{ij}$, such that μ_{ij}, γ_{ij} do not depend on time t.

Also, consider an oversampled noisy probabilistic process $Q(X_{ijt'}) = Q_{t'}$ defined over the following time sequence with $t' \in [1, ..., \alpha T]$, and $\alpha > 0$ the oversampling factor. t and t' are related to each other the following way: $\phi : [1, ..., \alpha T] \to [1, ..., T]$, such that, $t = \phi(t') = \lfloor \frac{t'}{\alpha} \rfloor$. Furthermore, $Q_{t'}$ is related to P_t the following way:

$$Q_{t'} = \frac{1}{\alpha} P_{\phi(t')} + \epsilon,$$

with $\epsilon \in N(0, \sigma)$ representing Gaussian noise. Let $\langle E', \alpha T\rangle$ be the stream of edges generated by Q. Note that if $Pr[(e_{ij}, t) \in \langle E, T\rangle] = p_{ijt}$, then $Pr[(e_{ij}, t') \in \langle E', \alpha T\rangle] = q_{ijt'} = \frac{p_{ijt}}{\alpha} + \epsilon$. Therefore,

1) $\mathbb{E}[q_{ijt'}] = \mathbb{E}[\frac{p_{ijt}}{\alpha} + \epsilon] = \frac{\mu_{ij}}{\alpha}$,
2) $Cov(q_{ijt'}, q_{ij(t'+\tau)}) = \mathbb{E}[q_{ijt'} \times q_{ij(t'+\tau)}] - \mathbb{E}[q_{ijt'}]\mathbb{E}[q_{ij(t'+\tau)}] = 0$, because $q_{ijt'}, q_{ij(t'+\tau)}$ are independent variables.

Let $DG_{Q,\omega}$ be the dynamic graph defined over $\langle E', \alpha T\rangle$ at window of aggregation $\omega : DG_{Q,\omega} = \mathcal{A}(\langle E', \alpha T\rangle, \omega)$:

$$DG_{Q,\omega} = \langle G_0, G_1, ..., G_k, ...G_{\alpha T/\omega - 1}\rangle$$

Let F be the resulting time-series that we get by applying function f (as defined on Section 3) on $DG_{Q,\alpha}$:

$$F = \langle f(G_0), f(G_1), ..., f(G_k), ...f(G_{\alpha T/\omega-1})\rangle$$

The following theorem states that stationarity of F depends on the value of the windows of aggregation.

Theorem 1. *Let DG be a dynamic graph which is the result of aggregation over a window ω of a stream of edges, generated by a covariance-stationary process oversampled at a rate of α. Let $F(DG)$ be the time series of a linear function over the edges of DG. Then:*

a) *$F(DG)$ is covariance-stationary when the window of aggregation ω is a multiple of α; and*
b) *there exists ω which is not a multiple of α, s.t. $F(DG)$ is not covariance-stationary.*

Before giving the proof the Theorem 1, we will first consider a special case of a stationary process, the probabilistic periodic process that generates the $DynMix$ graph.

A. $DynMix_{M,\{w_l\}}$ graph

We consider what happens when the aggregation function $\mathcal{A}$ is applied to a temporal stream of edges, each of them representing an edge in the $DynMix_{M,\{w_l\}}$ graph. As mentioned in Section III, M represents the number of probability distribution functions P_l generating the edge stream, and w_l represents the length of the temporal window during which edges were generated by P_l. We will first consider Case 1 of the $DynMix_{M,\{w_l\}}$. Recall that in Case 1, probability of an edge e_{lij} does not depend on time index t, for any given probability distribution P_l: $Pr[(e_{ij}, t) \in DynMix_{M,\{w_l\}} | (e_{ij})] = p_{lij}$.

The result of aggregating $DynMix_{M,\{w_l\}}$ is time series of graphs which we call $DynMix_\omega = \mathcal{A}(DynMIX_{M,\{w_l\}}, \omega)$. We will show that in the case of $DynMix_\omega$, stationarity of a linear function on edges is achieved only when the aggregation is done at the period level $W = \sum_{l=1}^{M} w_l$.

Lemma 1.

a) *The time series $F(DynMix_\omega)$ is covariance-stationary when the window of aggregation ω is a multiple of W, $\omega = nW$, where $W = \sum_{l=1}^{M} w_l$, and $n \in \mathbb{Z}$.*
b) *If $\omega \neq nW$, $\exists \omega$ s.t. $F(DynMix_\omega)$ is not covariance-stationary.*

Proof: a) Let $n = 1, \omega = W$. Let G_k be a graph in $DynMix_W$. Then, $f(G_k) = \sum_{i,j \in V} a_{ij} X_{ij}^k$, where X_{ij}^k is the indicator variable for the event $(e_{ij}, t) \in \langle E_k, T\rangle$ for any $t \in [kW, (k+1)W)$. Therefore, $X_{ij}^k = 1$, if (e_{ij}, t) is generated from P_1, or P_2, or ..., P_M. By this observation, the probability of an edge e_{ij} being in G_k is $Pr[X_{ij}^k = 1] = \sum_{l=1}^{M} p_{lij}$. Then the expectation of function $f(G_k)$ is:

$$\mathbb{E}[f(G_k)] = \mathbb{E}[\sum_{i,j \in V} a_{ij} X_{ij}^k]$$

$$= \sum_{i,j \in V} a_{ij} \mathbb{E}[X_{ij}^k] = \sum_{i,j \in V} a_{ij} \sum_{l=1}^{M} p_{lij} = \mu,$$

where μ doesn't depend on k.

Note that the proof trivially generalizes to $\omega = nW$ for arbitrary values of $n \in \mathbb{Z}$. By the periodicity of $DynMix$, $X_{ij}^k = X_{ij}^{k+\tau}$, and furthermore $f(G_k) = f(G_{k+\tau}) = \sum_{i,j \in V} a_{ij} X_{ij}^{k+\tau}$. Therefore, $Cov(f(G_k), f(G_{k+\tau})) = Var(f(G_k))$

b) Now consider the case when $DynMix_\omega$ is aggregated at windows $\omega \neq nW$. We will show, by giving an explicit example, that there exists a window of aggregation ω at which the time series $F(DynMix_\omega)$ is not stationary. For simplicity, assume $w_l = w$ for each P_l. Let the window of aggregation $\omega = w$.

$$\mathbb{E}[f(G_k)] = \mathbb{E}[\sum_{i,j \in V} a_{ij} X_{ij}^k]$$
$$= \sum_{i,j \in V} a_{ij} \mathbb{E}[X_{ij}^k] = \sum_{i,j \in V} a_{ij} p_{lij} = \mu \quad (1)$$

Since the value of p_{lij} depends on the value of k, μ is not constant with respect to k. Therefore, the time series $F(DynMix_w)$ is not stationary. Similar results can follow for Case 2 of probability distribution functions $\{P_l\}$. Recall that in Case 2, each P_l is a function that is constant over the edges and only depends on time index t. Following similar arguments as in the proofs of Lemma 4, it can be shown that time series F is stationary for any window of aggregation $\omega = nM, n \in \mathbb{Z}$, and that for values aggregation $\omega \neq nM$, there exist some ω when the corresponding time series F is not stationary. ■

We showed that for a periodic process representing the edge probabilities in a temporal stream, a linear function on the corresponding dynamic graph becomes stationary at specific windows of aggregation. These windows of aggregation correspond to the period (or any multiple of the period) of the underlying edge probability process. A periodic process is just one example of a stationary process. We can generalize the result in Lemma 1 for the class of oversampled noisy stationary process and show that the time series of linear functions on the discretized stream of edges generated by this process becomes stationary only at particular windows of aggregation (Theorem 1). The following is the proof of Theorem 1:

Proof: **a)** Let $\omega = \alpha, (n = 1)$

(1) $$\mathbb{E}[f(G_k)] = \sum_k f(G_k) Pr[f(G_k)],\ k \in [0, ..., T-1]$$
$$= \sum_k f(G_k) Pr[\sum_{t'=k\alpha}^{(k+1)\alpha} \sum_{i,j \in V} a_{ijt'} X_{ijt'}],$$
by the definition of f, G_k,
$$= \sum_k f(G_k) \sum_{t'=k\alpha}^{(k+1)\alpha} \sum_{i,j \in V} a_{ijt'} Pr[X_{ijt'}]$$
$$= \sum_k f(G_k) \sum_{t'=k\alpha}^{(k+1)\alpha} \sum_{i,j \in V} a_{ijt'} Q(X_{ijt'})$$
$$= \sum_k f(G_k) \alpha \sum_{i,j \in V} a_{ijt} (\frac{p_{ijt}}{\alpha} + \epsilon),$$
where $t = \lfloor \frac{t'}{\alpha} \rfloor$

Therefore,
$$\mathbb{E}[f(G_k)] = \sum_k f(G_k) \sum_{i,j \in V} a_{ijk\alpha}(p_{ijk\alpha} + \alpha\epsilon),$$
where $k\alpha$ = t
$= \mu_{Q,\alpha}$ a constant with respect to t'. ■

The proof generalizes trivially to $\omega = n\alpha$, for arbitrary values of $n \in \mathbb{Z}$

(2) $Cov(f(G_k), f(G_{k'+\tau})) = 0, \tau > 0$, follows by similar arguments used in the proof of Lemma 1.

(b) $\omega \neq n\alpha$

Consider the simple case of an underlying periodic process generating the temporal stream of edges. Let the period be α. Then, by Lemma 1, there exists a window of aggregation $\omega < \alpha$ such that F is not covariance-stationary.

V. Characteristics of $DynUR_\omega$ graphs

Consider what happens when the aggregation function $\mathcal{A}$ is applied to a temporal stream of edges, each of them representing an edge in the DynUR graph. The result is a time series of graphs which we call $DynUR_\omega = \mathcal{A}(DynUR, \omega)$. The following Lemma shows that $DynUR_\omega$ is a time series of instances of the same Erdős-Rényi graph.

Lemma 2. *Every $G_k \in DynUR_\omega$ is a $G(|V|, q)$ Erdős-Rényi graph, where $q = 1 - (1-p)^\omega$.*

Proof: By definition, $G_k \in DynUR$ is an Erdős-Rényi graph if each edge $e_{ijt} \in G_k$ exists with equal probability independently of other edges. The independence condition is trivially satisfied by the definition of the DynUR graph. We now show that the $Pr[(e_{ij}, t) \in G_k]$ is also the same $\forall (e_{ij}, t) \in G_k$.

An edge (e_{ij}, t) exists in G_k, if it exists in at least one of the w time values representing the time window for G_k: $t \in [k\omega, (k+1)\omega)$. By definition at any time t, $Pr[(e_{ij}, t) \in DynUR] = P(X_{ijt}) = p$. Therefore, $Pr[(e_{ij}, t) \in G_k] = 1 - (1-p)^\omega$. ■

Note, that for $\omega = 1$, $DynUR_1$ represents the original stream of temporal edges with $t \in [1, ..., T]$ and each $G_k = G_t$ is a $G(|V|, p)$ Erdős-Rényi graph.

A. Temporal Order Invariance

Consider the following probabilistic permutation process: a pair of time labeled edges $\langle (e_{i_1 j_1}, t_1), (e_{i_2 j_2}, t_2) \rangle$ is chosen independently at random from the set of all pairs in $\langle E, T \rangle$, with uniform probability $1/\binom{|\langle E,T \rangle|}{2}$. Given such a pair of edges, a temporal permutation function π is defined as follows:

Definition 5. *Let $\langle (e_{i_1 j_1}, t_1), (e_{i_2 j_2}, t_2) \rangle$ be a pair of edges chosen i.i.d. from $\langle E, T \rangle$, a temporal permutation function*

$\pi : \langle E, T\rangle \rightarrow \langle E, T\rangle$ *is defined as:*

$$\pi(\langle \ldots, (e_{i_1 j_1}, t_1), \ldots, (e_{i_2, j_2}, t_2), \ldots\rangle) = \langle \ldots, (e_{i_2 j_2}, t_1), \ldots, (e_{i_1 j_1}, t_2), \ldots\rangle \quad (2)$$

Lemma 3. *Aggregation of $DynUR_\omega$ is invariant under π :*

$$\mathcal{A}(\langle E, T\rangle, \omega) = \mathcal{A}(\pi(\langle E, T\rangle), \omega)$$

Proof: Let $\langle (e_{i_1 j_1}, t_1), (e_{i_2 j_2}, t_2)\rangle$ be the pair of edges chosen i.i.d. from $\langle E, T\rangle$, by the permutation function π. We consider the two cases:

Case 1: $k\omega \leq t_1, t_2 < (k+1)\omega, 0 \leq k = \frac{T}{\omega} - 1$. By the definition of the aggregation function $\mathcal{A}$, both edges $e_{i_1 j_1}, e_{i_2 j_2}$ belong to the same graph G_k, in the time-series $DynUR_\omega$. Therefore, even after π permutes the time labels of $e_{i_1 j_1}, e_{i_2 j_2}$, the aggregation function $\mathcal{A}$ will place them in the same graph G_k. Hence the resulting time series of G_k graphs will be identical before and after the permutation.

Case 2: $t_2 - t_1 > \omega$ We now compute the $Pr[(e_{ij}, t) \in G_k]$ after permutation. Lets consider the following mutually exclusive events:

E1: Edge e_{ij} was selected by π, and it got swapped out from graph G_k.

E2: Edge e_{ij} was not in graph G_k and it was not selected by π.

Let r=$Pr[e_{ij}$ selected by $\pi] = Pr[e_{ij}$ selected by $\pi | e_{ij}$ exists$] = \frac{p}{\binom{\langle E,T\rangle}{2}}$. If edge e_{ij} was removed from G_k after it was selected by π, that means, edge e_{ij} occurred in exactly one timestep during the time interval corresponding to G_k. Therefore, $Pr[\text{E1}] = rwp(1-p)^{w-1}$.

Let us now compute the probability of E2. Probability of edge e_{ij} not occurring in G_K is $1 - q$, where q represents the probability of edge e_{ij} being present in G_k by Lemma 2. Probability of edge e_{ij} not being selected by π is $1 - r$. Therefore, $Pr[\text{E2}] = (1-q)(1-r)$.

The event of edge e_{ij} being present in G_k after the permutation is the complementary event of the union of E1 and E2. Therefore, $Pr[(e_{ij}, t) \in G_k]$ after permutation is :

$$\begin{aligned} Pr[(e_{ij}, t) \in G_k] &= 1 - (Pr[\text{E1}] + Pr[\text{E2}]) \\ &= 1 - rwp(1-p)^{w-1} - (1-q)(1-r) \\ &= 1 - rwp(1-p)^{w-1} - (1-p)^w(1-r) \\ &= 1 - rwp(1-p)^{w-1} - (1-p)^w + (1-p)^w r. \quad (3) \end{aligned}$$

■

The result from Lemma 3 shows that even though the permutation process changes the probability of an edge existing in each partition, the type of graph representing each partition is still an Erdős-Rényi graph, but more importantly, it is the same type of graph across all the partitions. Also, this result is true regardless of the value of the aggregation window, a unique characteristic of the $DynUR$ graph.

B. Stationarity of functions on $DynUR_\omega$

Let f be a linear function on edges of a graph as described in Section 3. Let F be the resulting time-series that we get by applying function f to a Dynamic Uniform Random graph: $DG_\omega = DynUR_\omega$. The following lemma shows that F is a covariance-stationary time series for any value of aggregation window ω:

Lemma 4. *$F(DynUR_\omega)$ is covariance-stationary. That is, for some constants μ, γ, and τ :*

(1) $\mathbb{E}_k[f(G_k)] = \mu$,
(2) $Cov(f(G_k), f(G_{k+\tau})) = \gamma_\tau, \forall 0 < k < T, \tau > 0$

Proof: Let G_k be a graph in $DynUR_\omega$. Let X_{ij}^k be the indicator variable for the event $(e_{ij}, t) \in \langle E_k, T\rangle$ for any $t \in [k\omega, (k+1)\omega)$. Then, $f(G_k) = \sum_{i,j \in V} a_{ij} X_{ij}^k$. Recall that while edge e_{ij} can occur more than once in interval $[k\omega, (k+1)\omega)$, it shows up at most once in the aggregated graph G_k.

$$(1)\ \mathbb{E}[f(G_k)] = \mathbb{E}[\sum_{i,j\in V} a_{ij} X_{ij}^k] = \sum_{i,j \in V} a_{ij} \mathbb{E}[X_{ij}^k] = \sum_{i,j\in V} a_{ij} Pr(X_{ij}^k) = \sum_{i,j\in V} a_{ij} q = \mu$$

where μ is a constant with respect to the time index t.

(2) Let $X_{ij}^{k+\tau} = 1$ if $(e_{ij}, t) \in \langle E_{k+\tau}, T\rangle$ for any $t \in [(k+\tau)\omega, (k+\tau+1)\omega)$. By the definition of $DynUR$, $X_{ij}^k, X_{ij}^{k+\tau}$ are independent variables. $f(G_k) = \sum_{i,j\in V} a_{ij} X_{ij}^k$ and $f(G_{k+\tau}) = \sum_{i,j \in V} a_{ij} X_{ij}^{k+\tau}$ are independent variables as well, since they are defined as linear combinations of independent variables. Therefore, $Cov(f(G_k), f(G_{k+\tau})) = 0$

It is important to note that for the class of functions f, the property of stationarity is true at any value of aggregation of the $DynUR$ graph. Results of Lemma 2, 3, 4, all point to an inherent characteristic of the $DynUR$, the uniformity across windows of aggregations. Whether we analyze $DynUR$ at the graph level as a series of Erdős-Rényi graphs, or at the function level by looking at the class of functions f, the behavior of $DynUR$ is invariant with respect to the window of aggregation. This illustrates the fact that the $DynUR$ graph has no optimal temporal scale and in this context, it is the representative of the null hypothesis. ■

VI. Empirical Results

In this section, we investigate empirically how the window of aggregation at which the dynamic network is constructed affects the behavior of linear functions on this dynamic network. The goal is to identify the difference in behavior of these functions when they measure noisy processes versus structured ones. We use density and the average degree as examples of popular network measures that can be naturally expressed as linear functions on the edges of the network. We then analyze the time series of density and average degree as functions of the window of aggregations.

While there are several sophisticated methods to determine the stationarity of a time series, there is no standard, agreed process to do this. Often times, existing methods require artful tweaking of parameters. Since the goal in this work is not to determine stationarity per se, but to establish the differences between noisy and structured stable processes, we use the variance of the time series as a simple proxy for stationarity. The following section gives a brief description of the datasets used for analysis.

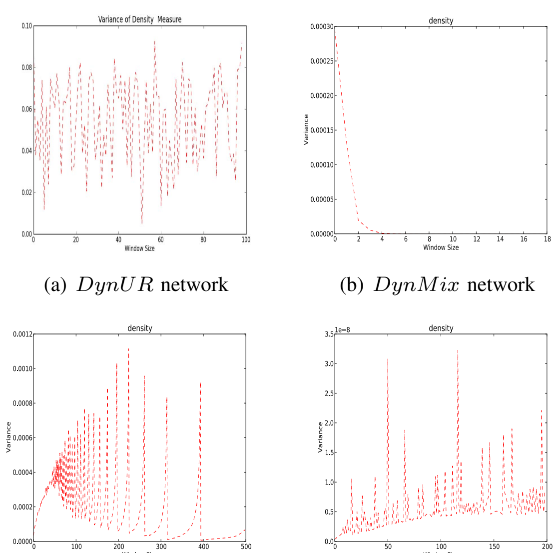
A. Datasets

DynUR is a simulated network where each edge is present at any time t with some fixed probability p.

DynMix is a simulated oversampled network, with edges generated by two alternating fixed probability distribution functions. For our simulations, we used the Beta prime and the Gaussian distributions, with oversampling factor of $\alpha = 5$.

Reality Mining network consists of Bluetooth device interactions among 90 MIT students and faculty over a nine month period [1]. Participants are equipped with smart phones and an edge between two participants exists if a bluetooth connection is recorded. The authors in [16] argue the Reality Mining network is generated by a system in equilibrium. Therefore, the Reality Mining network can be viewed as representative of a network generated by an oversampled noisy stationary process.

Haggle Infocomm network consists of social interactions among attendees at an IEEE Infocomm conference [18]. There were 41 participants and the duration of the conference was 4 days. The interactions present in this network have periodic nature imposed by the regular conference breaks taking place at fixed times. This fixed periodicity makes the Haggle network another good example of a real network generated by a noisy stationary process.



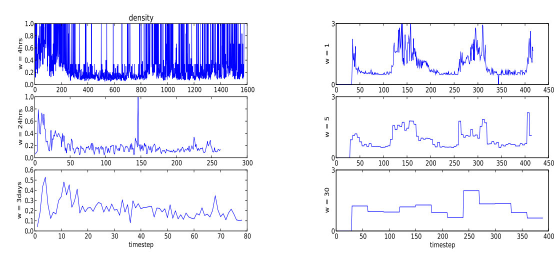
(a) $DynUR$ network
(b) $DynMix$ network
(c) Reality Mining network
(d) Haggle network

Figure 3. Variance of the density measure as a function of the window of aggregation for networks (a), (b), (c), and (d)

Figure 3 shows the plot of the variance of density as a function of the window of aggregation for each of the datasets mentioned above. The most immediate result illustrated in these plots is the fact that the variance function behaves distinctively different in the noisy network ($DynUR$), and the structured networks ($DynMix$, Haggle, and Reality Mining). The variance function is almost constant when computed over the $DynUR$ network (Figure 3(a)). On the contrary, in the case of the structured networks, the variance function stabilizes only for particular windows of aggregation and there is a visible trend. As illustrated in Figure 3(b), the variance of the density for the $DynMix$ network becomes stationary at window of aggregation 5 (and higher). This value corresponds to the oversampling factor used for simulating the $DynMix$ network. In the case of the Reality Mining and Haggle networks (Figure 3(a), (b)) there is a correspondence between the periodicity of the dataset (1 day for Reality Mining and 30 minutes for Haggle), and the times when variance approaches 0 or stabilizes.

Figure 4 displays the density function for the networks of Reality Mining and Haggle computed at three windows of aggregation: a very fine window of aggregation, an appropriate window of aggregation, and finally, a coarse window of aggregation. In the case of the Reality Mining dataset, aggregating interactions every 24 hours allows for the clear identification of the spring break event (represented by the high peak in Figure 4(a)) Similarly, periodicity of the Haggle network is better represented at window of aggregation 30 minutes, which corresponds to the length of the conference talks and the time between conference breaks when interactions were re-established.

The behavior of the average degree function is almost identical to the behavior of density across all the datasets analyzed above.



(a) Reality Mining
(b) Haggle

Figure 4. Density measure computed at three levels of aggregation: too fine (top), just right (middle), too coarse (bottom) for networks (a), and (b)

VII. Discussion/Conclusion

We have formalized and examined the process of aggregating a stream of edges into a time series of graphs. We have shown that for a temporal edge stream generated by an oversampled stationary process with added Gaussian noise, a linear function on the corresponding dynamic graph becomes stationary only for a particular windows of aggregation. This special window of aggregation represents the inherent temporal scale of the underlying processes in the stream of interactions. Besides being of theoretical interest, this is a sufficiently general and practically applicable result. Inference about long term trends or typical behavior is not useful if the underlying system is not stable. Alternatively, when the stationarity of a process changes it signals an important event or a change in the underlying nature of the process. Thus, the piece-wise stationary probabilistic streams of interactions is the main class of interest in analysis of dynamic complex systems. Our work here presents results for weakly stationary processes but can and should be extended to piece-wise stationary processes in the future.

In addition to showing the existence of a "natural" aggregation level of a stationary oversampled stream of interactions, we also showed that no such special aggregation level exists for a uniform random interaction stream. That is, a linear function computed over such a stream behaves identically at any window of aggregation. This allows us to separate uniform random interaction streams from those streams that have fixed and distinct natural temporal scales. This view can lead to an algorithmic approach that uses such measures and the notion of their stationarity at the "right window" to find that appropriate temporal scale. This is certainly the direction we would like to pursue in our future research.

This paper presents the first step in the direction of formalizing the problem of finding the inherent rhythms of a stream of interactions. We have rigorously shown what happens when there is a unique temporal scale to those interactions. A generalization of our result should include the possibility of multiple temporal scales and their possible change over time. The framework we have set up can be extended to include these more general problems while maintaining the principled approach to temporal analysis of interaction streams.

References

[1] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems," *Personal and Ubiquitous Computing*, vol. V10, no. 4, pp. 255–268, May 2006.

[2] Q. Gao, M. Li, M. B, and P. Vitnyi, "Applying mdl to learning best model granularity," 2000.

[3] S. Kotssiantis and D. Kanellopoulos, "Discretization techniques: A recent survey," in *GESTS International Transactions on Computer Science and Enginerting, 32(1)*, 2006, pp. 47–58.

[4] S. Papadimitriou, F. Li, G. Kollios, and P. S. Yu, "Time series compressibility and privacy," in *In VLDB*, 2007, pp. 459–470.

[5] M. H. Pesaran and A. Timmermann, "Model instability and choice of observation window," Department of Economics, UC San Diego, Economics Working Paper Series 99-19, Sep 1999.

[6] J. I. Adibi, "Enron email dataset," Downloaded from http://www.isi.edu/ adibi/Enron/Enron.htm.

[7] R. Sulo, T. Berger-Wolf, and R. Grossman, "Meaningful selection of temporal resolution for dynamic networks," in *Proc. 8th Workshop on Mining and Learning with Graphs*, ser. MLG '10. New York, NY, USA: ACM, 2010, pp. 127–136.

[8] C. Partridge, D. Cousins, A. W. Jackson, R. Krishnan, T. Saxena, and W. T. Strayer, "Using signal processing to analyze wireless data traffic," in *Proc. ACM workshop on Wireless Security*, 2002, pp. 67–76.

[9] A. Feldmann, A. C. Gilbert, W. Willinger, and T. Kurtz, "The changing nature of network traffic: Scaling phenomena," *Computer Communication Review*, vol. 28, pp. 5–29, 1998.

[10] T. Y. Berger-Wolf and J. Saia, "A framework for analysis of dynamic social networks," in *Proc. 12th ACM SIGKDD on Knowledge discovery and data mining*. New York, NY: ACM, 2006, pp. 523–528.

[11] D. Kempe, J. Kleinberg, and A. Kumar, "Connectivity and inference problems for temporal networks," *J. Comput. Syst. Sci.*, vol. 64, no. 4, pp. 820–842, 2002.

[12] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *Proc. 11th ACM SIGKDD on Knowledge discovery in data mining*. New York, NY: ACM, 2005, pp. 177–187.

[13] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," in *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*. New York, NY, USA: ACM, 2006, pp. 228–237.

[14] P. Desikan and J. Srivastava, "Mining Temporally Evolving Graphs," in *Proc. of WebKDD 2004*, 2004, pp. 22–25.

[15] C. Tantipathananandh, T. Berger-Wolf, and D. Kempe, "A framework for community identification in dynamic social networks," in *KDD '07: Proc. 13th ACM SIGKDD on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2007, pp. 717–726.

[16] A. P. Pentland, N. N. Eagle, and N. N. Eagle, "Machine perception and learning of complex social systems," in *Ph.D. Thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology*, 2005.

[17] J. Sun, C. Faloutsos, S. Papadimitriou, and P. S. Yu, "Graphscope: parameter-free mining of large time-evolving graphs," in *KDD '07: Proc. 13th ACM SIGKDD on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2007, pp. 687–696.

[18] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace cambridge/haggle/imote/infocom (v. 2006-01-31)," Jan 2006.