IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.DOI

Mining Graph-Fourier Transform Time Series for Anomaly Detection of Internet Traffic at Core and Metro Networks

MANUEL HERRERA[1], (Member, IEEE), YANIV PROSELKOV[1], MARCO PÉREZ-HERNÁNDEZ[1], (Member, IEEE), and AJITH K. PARLIKAD[1], (Member, IEEE)

[1]University of Cambridge, Department of Engineering, Institute for Manufacturing, 17 Charles Babbage Road, Cambridge, CB3 0FS, UK

Corresponding author: Manuel Herrera (e-mail: amh226@cam.ac.uk).

This research was funded by the EPSRC and BT Prosperity Partnership project: Next Generation Converged Digital Infrastructure, grant number EP/R004935/1.

ABSTRACT This paper proposes a framework to analyse traffic-data processes on a long-haul backbone infrastructure network providing internet services at a national level. This type of network requires low latency and fast speed, which means there is a large demand for research focusing on near real-time decision-making and resilience assessment. To this aim, this paper proposes two innovative, complementary procedures: a multi-view approach for the topology analysis of a backbone network at a static level and a time-series mining approach of the graph signal for modelling the traffic dynamics. The combined framework provides a deeper understanding of a backbone network than classical models, allowing for backbone network optimisation operations and management at near real-time. The applications of this methodology to the backbone infrastructure of one of the main internet service providers in the UK shows increased accuracy and computational efficiency on the detection of where and when anomalies and irregular patterns occur in the network signal.

INDEX TERMS Backbone network, Graph signal processing, Internet infrastructure, Network topology, Anomaly detection, Time series mining

I. INTRODUCTION

THE long-haul internet backbone network is a set of spatially distributed points of presence (PoPs) or interface points over a large area such as a country or world region. In the case of the long-haul backbone infrastructure, a PoP is located at internet exchange points and data centres and it comprises elements such as servers, routers, network switches, and multiplexers. PoPs connect each other by high speed fibre-optic cables for fast data transmissions and large bandwidths [1], [2]. The internet backbone provides internet access to lower capacity networks and is essential to the its final users situated across a nation or a geographical region. Therefore, the operation and performance of other critical infrastructures and governmental services (banking, road safety, digital healthcare, e.g.) that use internet data rely on the high performance and optimal resilience of the internet backbone [3], [4].

A nationwide internet backbone infrastructure comprises the so-called core and metro networks [5]. The core and metro nodes are PoPs comprising a set of multi-protocol label switching (MPLS) routers and switches. MPLS traffic (routing and label information) is carried between PoPs over fibre-optic cables; usually dense wavelength division multiplexing (DWDM) cables. The backbone network terminates on the metro nodes, where internet protocol (IP) based services are implemented. Therefore, metro nodes are the first location where IP traffic is routed and ready to provide further internet access to the users [6]. This paper comprises a topological and a signal processing analysis for the detection of anomalous network operation based on real-time data. In particular, the location of discords and anomalies at signals of nationwide core and metro infrastructure leading to a better network control and management. The challenge is to discover when and where those events occur through a quick identification process. To this end, this paper uses both static and network dynamics approaches.

- From a topological perspective, this paper uses complex networks as mathematical graphs representing actual

systems of interconnected elements. Complex networks help to decompose the backbone network into core and metro nodes/PoPs and to develop analysis further. Depending on the characteristics of their router stations, elements belonging to the core infrastructure are further divided into 2 types of PoPs: super and regional routers. In the literature, it is possible to find out works related to topology and safety and security within a complex networks perspective. This is the case in the work of Shatto and Cetinkaya [7] which focuses on the resilience assessment of backbone networks facing targeted attacks from a spectral graph theory approach. Durairajan *et al.* [8] present a comprehensive study of the United States backbone network topology along with its resilience assessment. The authors of [9] showed how network geometric properties may lead, under a certain re-normalisation based on network topologies, to efficient routing protocols. Other works focused on topological connectivity and data quality assessment of a nationwide internet infrastructure [10].

- From a network dynamics point of view, the paper proposes a time series analysis of the data traffic on the network within a graph signal processing (GSP) framework [11]. The most important part of the literature of telecommunication systems related to this approach focuses on traffic network analysis, monitoring and control. The main objectives of traffic monitoring and control are to maximise network performance and for routing optimisation [12]. In this regard, the work of Fang *et al.* [13] addressed monitoring for multiple traffic-related performance indicators. Taking into account certain network topological characteristics in addition to data traffic, Papadopoulos and Psounis [14] proposed a reduced version of the backbone network by discriminating between congested and uncongested links [15]. This approach supports network management by simplifying network analysis.

One of the main novelties of this paper is working with multiple network views, or network slices, to analyse how data traffic is distributed at each of them within the emergent framework of GSP [16]. A direct application of GSP comes from informing routing protocols of signal anomalies to then use alternative and backup paths for sending data packets. The first part of the paper presents a backbone network decomposition into 2 widely known topological structures: small-world and scale-free topologies [17]. The paper continues by introducing time series mining for the evolution analysis of the signal spectra on the network. Hence, the main contribution of this paper is the creation of a joint analysis framework taking into account both network structure and dynamics on networks. Time series mining [18], [19] of the signal spectra of the graph provides near real-time information about when and upon which nodes the network is exhibiting irregular patterns. Information about the most regular signal patterns on the network is also provided. Then, it is possible to use this information to support decision-making for network management and traffic control functions.

The following is an outline of the antecedents of this work. Coates *et al.* [20] developed an inverse problem framework to reconstruct an internet network signal. Blumenthal *et al.* [21] used wavelets methods to analyse optical packet switching networks, aiming to improve routing technologies. Lin and Hsueh [22] analysed spectral properties of optical communication systems. Shimizu *et al.* [23] investigated wavelength conversion for signal amplification at optical network nodes. Ji *et al.* [24] used signal processing to detect network traffic anomalies. To this aim, the authors work with statistical methods over the wavelet transform of the network signal. Poudereux *et al.* [25] designed a filter-bank system for signals in broadband communications. Pu *et al.* [26] worked in development of innovative materials and devices to enable optical networks with signal processing functionalities. The aforementioned signal processing methods and technologies fall short in providing a network perspective of the problem, so the spatial and structural components are analysed from a complex network approach [27], [28]. Instances of complex network approaches applied to communication networks are: spectral graph theory [29], epidemic/diffusion processes on networks [30], temporal networks [31], assortativity analysis [32], and percolation analysis [33], among others. Despite the extent of the literature related to signals and internet networks, there exists little research into strategies for combination of data mining methods and complex networks, as highlighted by Zanin *et al.* [34]. This paper aims to cover this gap.

The outline of the rest of the paper is as follows. Section II introduces the mesh topology of backbone networks. This section specifically discusses small-world and scale-free topologies as better candidates to represent the backbone infrastructure topology. While Section III introduces some necessary fundamentals of GSP; Section IV presents the main methodology proposed for mining time series of signal spectra evolving on the network. Section V presents the case-study of a backbone infrastructure providing internet services to the UK. This network topology is actually a combination of small-world and scale-free networks. GSP techniques are applied to the UK backbone using the global network but also using the network slices, which are substructures that divide up the network. The results show the fundamental advantages of this approach. Section VI closes the paper with the main conclusions arising from the work, along with an open discussion on current and future research in long-haul, internet backbone infrastructure.

II. INTERNET BACKBONE NETWORKS

The internet backbone network should be designed to support a variety of data transport services, some of which requiring high speed or ultra high speed channels to reach low latency for enabling fast communications. To achieve these requirements, network designers consider not only physical properties and capabilities of the links and network devices,

but also network topology to enable connections while minimising the number of hops within the network. The idea is that the fewer hops a data packet needs to go through before exiting the backbone, the faster it will reach its destination. This is the reason why the most common topologies adopted for the backbone infrastructure are highly dense, following a connectivity structure of "small-world" and "scale-free" topologies. Actually, a combination of both is widely used in the practical topology design of internet backbone networks [35]. After a discussion of the elements of a backbone network and their operation, this section presents a brief overview of the most essential internet network topologies. Small-world and scale-free topologies in complex networks are, specifically, introduced along with their implications for the backbone network performance and resilience [36].

A. INTRODUCTION TO BACKBONE NETWORKS

Backbone networks are networks of router stations and/or PoPs at a regional and national level (wide area networks - WAN - and metropolitan area networks - MAN) connected to other shorter-distance networks. The latter networks are connected to others at a more local level and so on, finally reaching the end-users. Ultimately, the topology of the global internet is a network of networks, where each of those sub-networks can be understood as an autonomous system (AS). From a coarser perspective, the backbone connecting ASs is named a transit-AS, making reference to the importance of routing data between multiple ASs and in contrast to the so called stub-ASs, which connect to only one other AS (corresponding to local area networks, LANs).

At the global and national level there are multiple internet service providers (ISPs) who own individual backbone networks and ASs [37]. All the backbones are physically connected to create the worldwide network giving access to a global internet routing table. This allows ISPs to efficiently deliver data traffic from source to destination through a hierarchy of regional and local providers. Data traffic along the network is enabled by the use of standard protocols, such as the Transport Control and Internet Protocols (TCP/IP).

The long-haul internet backbone network comprises the core and metro networks [38]. The key elements of the core and metro network and the mapping of these elements to the complex networks notions of links and nodes are described below as well as summarised in Fig. 1.

- Optical cables: These usually are the bandwidth-efficient fibre optical cables (network links). Thanks to wavelength division multiplexing (WDM) technology, fibre optical cables can carry data at high speeds through a synchronous optical networking (SONET) protocol. These advantages come with a high resistance to corrosion and low signal attenuation. Nevertheless, fibre optical cables have the downside of a high economic cost and are complex to install and maintain. In most of the cases, fibre optical cable is used in WANs while coaxial cable is used for LANs.

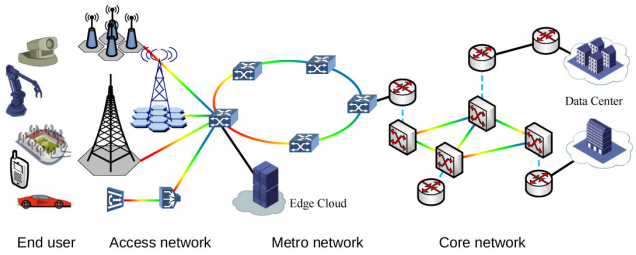

FIGURE 1. Basic elements and sub-systems of optical networks. Adapted from [39], under Creative Commons license.

- Core network: This is a cluster of ISP PoPs (network nodes), each built from large scale, high capacity routers with MPLS support for faster redirection of data using shortest paths labels rather than a conventional protocol addresses.
  - Inner core network: This is a fully meshed network of PoPs with a complete service range. Some of these PoPs have direct internet access via wired or satellite internet sources abroad. The inner core PoPs offer a full range of functionality across different types of traffic such as data, video and voice. These nodes are also named super routers.
  - Outer core network: PoPs of this network are densely connected to the inner core, while the connectivity between the same regional and metro nodes is only partially meshed. Some of the links at this network are one-directional cables (sideways links) and their function is just focused on traffic transit. Most of the outer core PoPS compress advanced infrastructure such as routers, switches and servers, necessary for network traffic. In this paper, nodes of this network are also called regional routers.
- Metro network: This is a collection of ISP PoPs comprising large scale routers which are responsible for the IP routing and LANs switching. Metro nodes are also gateways for voice, data and media on to the core network.

B. BACKBONE NETWORKS TOPOLOGIES

Traditionally, backbone networks have been built from other sub-networks designed using the basic topology types briefly presented in the following bullet points (also see Fig. 2).

- Bus: The nodes are connected to a single, central link cable.
- Ring: A ring topology is actually a bus topology in a loop. The signal is passed by the routers along the loop in one direction, until it reaches the destination.
- Star: The nodes are connected to a single central node called a hub. There are extended versions of the star topology where there are nodes which are also hubs for other nodes. A tree-like or hierarchical configuration is a particular case of an extended star topology.

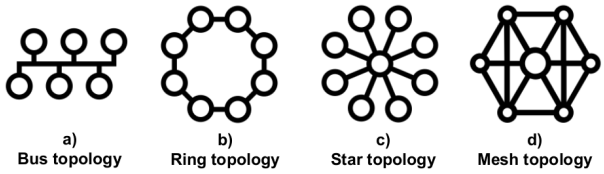

FIGURE 2. Basic backbone network topologies.

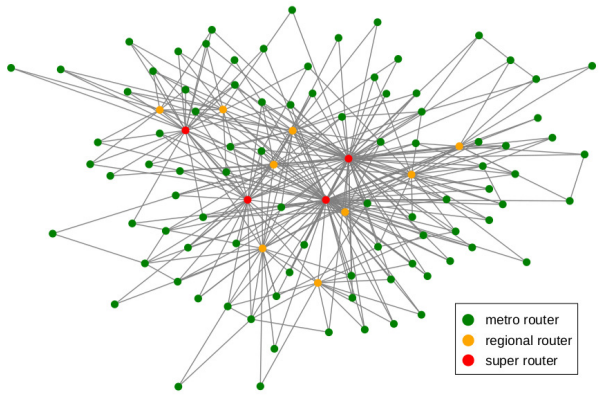

FIGURE 3. Mesh topology of the UK long-haul backbone network layout. Geographical information is withheld to preserve anonymity.

- Mesh (or lattice): The nodes can be fully connected to each other or partially connected. In the latter case, a proportion of the network nodes have point to point connectivity. However, richer connectivity is also allowed and one node can connect to more than one node.

Real-world topologies, as the showed in Fig. 3, often are hybrid topologies of the types mentioned above and can even follow a free combination of the systems. As a consequence, further topological classification and analysis should be drawn from complex network theory. The following subsections introduce the common topologies normally adopted in backbone networks design. Appendix A presents more formal definitions.

1) Small-world network

A network topology refers to how network nodes are organised and interconnected by network links. Extreme examples of network topologies are random networks and regular lattices, respectively having a completely random and a completely regular connectivity structure. A small-world topology is a topology in between of these two [40], [41]. This inherits the highly clustered features of regular lattices (large cluster coefficient, that is, a large number of triangles), while it also has short average path lengths (similar, in this sense, to random networks). Following the Watts–Strogatz model [42] for generation of graphs with small-world properties, the maximum path length for small-world topologies approaches $\lceil (n/2)/(k/2) \rceil = \lceil n/k \rceil$, where $n \gg k \gg \log(n) \gg 1$. The ceiling function, $\lceil x \rceil$, computes the smallest integer larger than or equal to x. Dynamics on networks with small-world topology have been associated an enhanced signal-propagation speed [43], a high computational power [44], and they are suitable for hosting synchronised processes [45]. These properties are the reasons for this network topology to be widely adopted for backbone network design.

Small-world network structure shows resilience to failures on account of its high inter-connectivity level, however there is a risk of quick propagation of network failure under conditions of heavy traffic load [46]. This risk can be considered as the payoff for the generally high speed of signal-propagation that so effectively suits backbone network requirements.

2) Scale-free network

The concept of the small-world can be extended to topologies that comprise a small number of high degree nodes, where the number of nodes with high degree decreases exponentially, following a power law [47]. That is, the probability that a certain node has a degree d is proportional to $(1/d)^{\alpha}$, where $\alpha > 1$ is the scaling component. Among the multiple methodologies to synthetically generate scale-free topologies, the most famous is the method of Barábasi-Albert, named after its authors [48]. Generally, any generation methodology has a common "preferential attachment" step. That is, if the generation process is sequential (adding a node and links at each step) there is a higher probability that the new network element is connected to the hubs of the network than to nodes of lower degree.

Scale-free networks topology shows a balance between keeping part of the high connectivity of small-world topology and increasing its resilience against failure propagation [49]. This latter feature however is only partially achieved, as shown in the literature by multiple percolation studies [50]. In this regard, the resilience to random failures is excellent while the resilience to targeted attacks (focused on the hubs that are part of any scale-free structure) is poorer. Still, even if a hub-node fails, the issue can be contained within areas of the network, avoiding the risk of complete black-outs of small-world topologies [51].

C. SIGNALS IN CORE AND METRO NETWORKS

This paper investigates graph signal processing on the long-haul, internet backbone infrastructure, focusing on the three bottom layers of the Open System Interconnection (OSI) model [52]. That is, the physical, data-link and network layers.

A long-haul backbone network is topologically characterised by a scale-free structure that embeds the core and metro networks. It is possible to define the extended inner-core (xIC) network as the network slice comprising the inner and regional networks. At a closer look, the xIC network (see Fig. 4) is, by design, a highly dense network with almost all the nodes connected to each other. Therefore, the core network can be represented by a small-world model. This is in contrast to the topology of the entire network which typically follows a scale-free model. One of the hypotheses sustained in this paper is that signal patterns are dependent

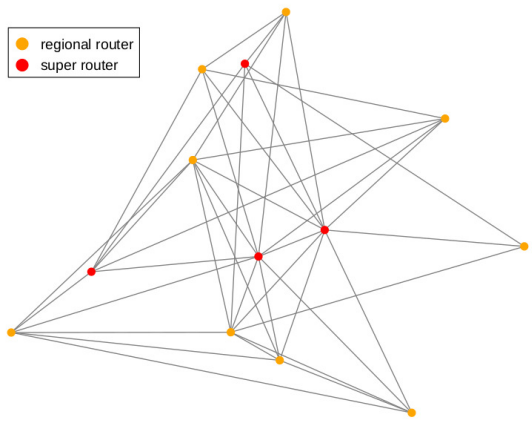

FIGURE 4. Small-world topology of the UK xIC network: inner core and regional PoPs. Geographical information is withheld to preserve anonymity.

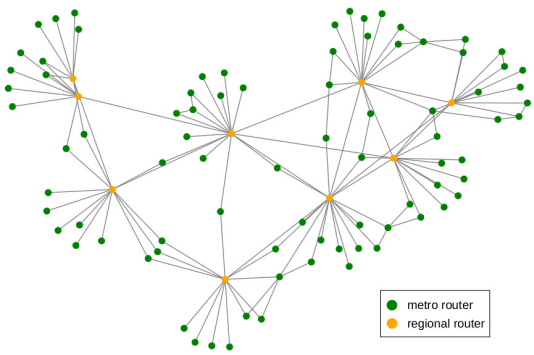

FIGURE 5. Scale-free topology of the UK xOC network: regional and metro PoPs. Geographical information is withheld to preserve anonymity.

on such models or topologies. In conducting graph signal processing analysis at finer scales of the network, it is of interest to find new patterns which normally may remain hidden in the entire network.

A second network view of the backbone network is through the extended outer-core (xOC) network, that is, the network slice containing regional and metro networks (see Fig. 5). This network slice plays a key role in the regular data traffic that reaches the metro nodes and exits the backbone to other ISP systems that finally reach the end-user. Metro nodes preferentially connect to the regional network rather than to the inner core, since the costs associated with a metro inner-core network are higher.

III. GRAPH SIGNAL PROCESSING: A BRIEF INTRODUCTION

GSP comprises a set of techniques for analysis of signals defined on an irregular, non-Euclidean domain. This is the case of graphs, and specifically of the internet networks as it is the case of the research presented herein. In general, GSP techniques involve weighted graphs. Nodes are weighted by signals and links by similarity between the nodes they connect. The analyses within a GSP framework usually encompass a range of graph spectral methods, where the graph Laplacian matrix plays a pivotal role. Among the aims of GSP, highlighted are filtering, manipulation and extracting signal patterns out of graphs.

A. LAPLACIAN MATRIX

Consider a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of n vertices (also called nodes if referring to a network) and m edges (links). A vertex set is defined as $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$, and the edges between them as $\mathcal{E} = \{e_1, e_2, \ldots, e_m\}$. A graph, then, is normally defined by its adjacency matrix, $\boldsymbol{A} = [a_{ij}]$. Here $a_{ij} \in \{0, 1\}$, such that when $(i, j) \in \mathcal{E}$, that is, i and j share an edge, $a_{ij} = 1$, but if not, $a_{ij} = 0$.

The Laplacian matrix [53] has the ability to capture important properties of graphs and it is key part of the essentials in GSP theory and applications. For instance, one of the interpretations of a Laplacian matrix is about its ability to show in which direction and how smoothly any "energy" or information may diffuse over a graph [54]. Then, Laplacian matrices have been used in the literature to investigate how a signal propagates in a complex network. The Laplacian is a real, symmetric matrix defined as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{D} = [d_{ij}]$, and d_{ii} is the degree of node i, and $\boldsymbol{D}$ is 0 everywhere off the leading diagonal. In the case of weighted networks, it is possible to define the Laplacian matrix by $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{W}$, by straightforwardly replacing $\boldsymbol{A}$ with the weighted adjacency matrix, $\boldsymbol{W} = [w_{ij}]$, where $w_{ij} \geq 0$ and zero when $(i, j) \notin \mathcal{E}$.

The Laplacian matrix spectrum is defined by its eigenvalues and their multiplicities. That is, the eigenvalues $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ corresponding to the set of orthonormal eigenvectors $\boldsymbol{U} = \{\boldsymbol{u}_i\}_{i=1}^{n}$ define $\boldsymbol{L}$. The eigenvalues can be considered arranged in a matrix $\boldsymbol{\Lambda} = \text{diag}[\lambda_1, \ldots, \lambda_n]$. Properties of structural connectivity and energy on networks can be analysed by using the Laplacian spectrum [29].

B. GRAPH FOURIER TRANSFORM

The eigenvalues of the Laplacian operator are related to the signal frequencies on the graph. This suggests a parallelism between the classical Fourier transform on a regular domain and its equivalent, performed on graphs. Therefore, it is possible to define the graph-Fourier transform (GFT) in similar terms as in the definition in the regular domain. This is by considering a signal, $\boldsymbol{s} \in \mathbb{R}^n$, on graph, $\mathcal{G}$, represented by eigenvalues of the Laplacian matrix associated with $\mathcal{G}$. As a result a GFT is a mapping of the graph signal frequencies (or a function of such signals), as expressed in (1).

$$\hat{\boldsymbol{s}} = \boldsymbol{U}^{-1}\boldsymbol{s}. \tag{1}$$

A signal can be reconstructed back from the GFT using the inverse of the transform. Out of the scope of other GSP operations [55], graph filters [56] can be represented by the matrix $\boldsymbol{H}$, $\boldsymbol{H} = h(\boldsymbol{A})$, and can be, for instance, a polynomial combination of the matrix $\boldsymbol{A}$ (among other possible filters). The relation of (2) is valid only by considering regular

conditions such as $\boldsymbol{A}$ having a complete set of eigenvectors and the existence of $\boldsymbol{U}^{-1}$.

$$\boldsymbol{H} = h(\boldsymbol{A}) = h\left(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1}\right) = \boldsymbol{U}h(\boldsymbol{\Lambda})\boldsymbol{U}^{-1}, \quad (2)$$

where $h(\boldsymbol{\Lambda}) = \text{diag}[h(\lambda_1), \ldots, h(\lambda_n)]$.

The outcome, $\boldsymbol{s}_{out}$, of filtering a given signal, $\boldsymbol{s}_{in}$, follows the equations in (3), written in relation to the GFT, $\hat{\boldsymbol{s}}$.

$$\boldsymbol{s}_{out} = \boldsymbol{H}\boldsymbol{s}_{in} = \boldsymbol{U}h(\boldsymbol{\Lambda})\boldsymbol{U}^{-1}\boldsymbol{s}_{in} = \boldsymbol{U}h(\boldsymbol{\Lambda})\hat{\boldsymbol{s}}_{in}. \quad (3)$$

IV. MINING GRAPH-FOURIER TRANSFORM TIME SERIES

The study of complex networks is commonly addressed from a topological, static point of view [57]. However, those analyses should be complemented by the study of the flow passing, if any exists, through such a network. In this regard, time series (TS) mining methods [58] play a key role in facilitating operational monitoring of the network for each of the previously analysed sub-structures. TS mining methods are able to extract common signal-patterns or motifs as well as especially irregular signal-patterns from the traffic flow, e.g. discords or anomalies. This Section develops a novel methodology to deal with a GFT-TS corresponding to the evolution of the observed signal on the network.

A. METHODOLOGY

This paper aims to extract patterns from the signal of data traffic evolving in time in order to detect anomalous patterns. To this end, the methodology combines into one single process, different methods borrowed from research areas such as network science, graph signal processing, and TS mining.

The overall process, first, takes an internet backbone network from multiple-views (network slices) of expected well defined topologies. As a result, the backbone is viewed through a combination of its parts. The work focuses on the xIC network, normally associated with a small-world topology, and the xOC network, which often presents a scale-free topology. As discussed in Section II and also through the rest of the paper, network topologies play a role on how a signal spreads over the network as well as to inherently have important characteristics for data traffic control and network operation and management. Therefore, each of the 2 network slices, xIC and xOC, has a separate GFT associated with the signal values. Since such network signals (data traffic) evolve over time, the process finally gets a GFT-TS, one per each network slice, xIC and xOC. Fig. 6 presents a flowchart with the main steps proposed for the overall GFT-TS process. Further details are in Algorithm 1 splitting the process into the following phases: initialisation; GFT-TS creation; extraction of GFT patterns related to the time in which they happened; computation of influence nodes or nodes having higher participation (if any) in those patterns; and the output phase.

Algorithm 1 Algorithm for mining GFT-TS at xIC and xOC networks

Input: Network topology and graph signal TS
Output: Graph signal TS motifs and discords at the time (and nodes) in which the patterns occur

INITIALISATION :
Parse a backbone network infrastructure
K = number of eigenvectors in the GSP spectra

GFT-TS creation:
for $t = 1$ to T **do**
  Compute $GFT(t)$ % *vector* $K \times 1$
  GFT.append($GFT(t)$) % *end in a vector* $KT \times 1$
end for

TIME and GFT of significant patterns:
$\boldsymbol{t}_d, \boldsymbol{t}_m$ = matrix-profile(GFT, K)
% $\boldsymbol{t}_d$, $\boldsymbol{t}_m$ *times of discords and motifs, respectively*
for t_p in $\boldsymbol{t}_d$, $\boldsymbol{t}_m$ **do**
  $d_1(t_p)$ = distance($GFT(t_p, t_p + K), \overline{GFT}$)
  if $d_1(t_p) > 1.645\sigma_{GFT}$ **then**
    $time$.append(t_p) % *times of significant patterns*
    for $k = 2$ to $K + 1$ **do**
      if $|GFT(t_p)_k - \overline{GFT}_k| > 1.645\sigma_{GFT(t_p)}$ **then**
        $candidate$.append($GFT(t_p)_k$)
      end if
    end for
  end if
end for

NODE of significant patterns:
for t in $time$, c in $candidate$ **do**
  $\boldsymbol{u}_c[t] \leftarrow \boldsymbol{U}[t](\cdot, c)$; $\bar{\boldsymbol{u}}_c \leftarrow \overline{\boldsymbol{U}}(\cdot, c)$
  for $n = 1$ to N **do**
    if $|u_c[t](n) - \bar{\boldsymbol{u}}_c| > 1.645\sigma_{u(c)}$ **then**
      $nodes$.append(n) % *influential nodes*
    end if
  end for
end for

OUTPUT:
return $nodes, time$

Further considerations to take into account in the mining of GFT-TS are:

- The output of a GFT is a vector, as shown in Section III, since this is the signal mapped into the eigenvector matrix of the Laplacian of the network graph.
- The eigenvector matrix of $\boldsymbol{L}$, $\boldsymbol{U}$, has a first column with all the elements of value 1. That is, $(u_{11}, \ldots, u_{1n})^T = \mathbb{1}$. This means that the first element of the GFT is the value of the signal itself and it can be out of the range of the rest of values in the GFT output. Therefore, the current analysis omits the GFT value associated with the

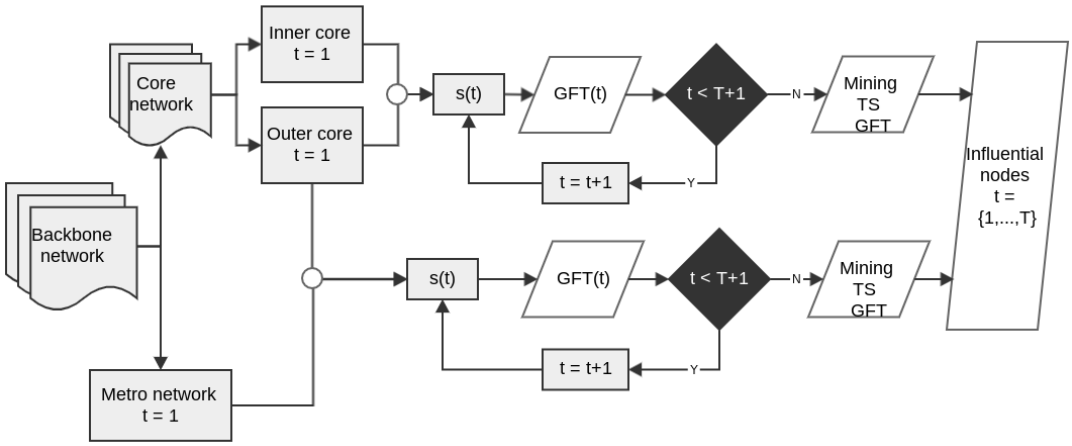

FIGURE 6. Outline of the procedure for mining GFT-TS

first eigenvector.

- There is a need to take a decision about the use of any filter over the GFT. The filter at node i, $\hat{s}(\lambda_i)$, is related to the neighbourhood of such a node. Hence, it is directly related to how the signal locally spreads at each node of the network structure. This is an important point because it aids to explain how the signal is locally distributed at each node which may enhance the knowledge that a GFT captures about the network connectivity and signal.
- GFT at each time is a vector with as many components as eigenvalues. Working with so-called GFT bandwidth means to select only the eigenvectors associated with eigenvalues below a certain threshold. An alternative to this approach is to directly work with a fixed number of the top K eigenvalues, which is the approach taken herein. Note that the term 'top' is with respect to the inverse order of the eigenvalues, where the top K eigenvalues are the K eigenvalues of lowest value. This is because the top eigenvectors are the most meaningful eigenvectors explaining the structure of the network and the corresponding mapping of the signal, while the rest of the spectrum is associated with noise and contains little new information. The advantage of working just with a set of eigenvalues and eigenvectors is the dimensional reduction of the problem with a minimum lost in the information about the problem.

Note that the output of a GFT for a given signal has a vector dimension at each time unit of the TS. Among the multiple possibilities that can be adapted the best to vector TS analysis, the popular ones are based on joint models of multiple TS [59], [60], TS clustering [61], symbolic representation [62], and matrix profile [63]. Out of these methods, matrix profile is chosen in this paper due to its scalability and computational efficiency, while being also robust and parameter-free [64].

B. MATRIX PROFILE

A matrix profile is a TS of distances such that at each timestamp it records the distance of a sub-sequence of length m to its nearest neighbour in the original TS. Matrix profile matches perfectly to the objective of the analysis herein, since it generally addresses TS analysis based on small local sub-sequences. This is the case in sequences of hours, words (as a sequence of letters) or GFT (sub-)sequences, as in the current case. Details on how matrix profile constructs the measure of distance, based on where it is possible to extract patterns, are given in Appendix B.

Matrix profile is based on comparison of distances between sub-sequences within a TS. The lowest distances between sub-sequences are the most common patterns, which often have (almost) repeated values. These sub-sequences are called TS motifs [65]. The highest distances correspond to discords and represent the most unusual sub-sequences in a TS [66]. If such distances to other sub-sequences are large enough, a discord becomes an anomaly. Note that matrix profile is parameter-free and requires less training data than any deep-learning based algorithm, providing better or at least comparable results [67], [68]. Then, matrix-profile works well with relatively small data sets, while it is also efficient to deal with extremely large time series.

Computing a matrix profile with a brute force algorithm takes $O(T^2m)$, as it is the number of pairs of distance comparisons between sub-sequences of length m in a TS of length T. In most of the cases this is a not feasible solution. However, there are algorithms such as the scalable TS ordered-search matrix profile (STOMP), that takes just $O(T^2)$ to provide an exact solution [69] or the scalable TS anytime matrix profile (STAMP), that takes only $O(\log T)$ to provide an accurate approximation [70]. The efficiency of the STOMP and STAMP algorithms is key for this paper since there is a need to work with a large series of concatenated GFT vectors.

In GFT-TS mining, the extracted motifs and discords via matrix profile are vectors of components directly related to the eigenvalues $\boldsymbol{\lambda} = (\lambda_2, \ldots, \lambda_{K+1})$ of the graph spectrum and the signal on the network. An analysis on the deviation with respect to the mean of each of the vector components provides information about the set of eigenvalues influencing the pattern. A typical rule to get those influential eigenvalues is to select them if the deviation from the mean is greater than 1.645 times the standard deviation, σ, of all the measures. The number 1.645 is selected since this is the 95% percentile of the standard normal distribution, often used as a criterion for outlier detection [71]. So, deviations from the mean larger than 1.645σ will capture the top 10% eigenvalues furthest from the mean, called the influential eigenvalues. The set of influential eigenvalues point match influential eigenvectors. Since an eigenvector is a combination of weights and nodes, it is possible to apply the same criterion than before to finally extract a set of influential network nodes. The output is, therefore, a set of influential network nodes at each of the times associated to motifs and/or discords in the GFT-TS mining.

C. EXPERIMENTAL STUDY

The experimental study lies in the generation of a benchmark of networks having similar topologies to those that core and metro networks usually follow, that is, small-world and scale-free topologies. On each of these networks, we simulate data traffic, any-to-any, mimicking the backbone network dynamics.

The network simulation is an adaptation based on the "anx" Python package proposed by [72]. This is developed in Python 3, mainly using the "SimPy" and "NetworkX" [73] libraries, respectively for discrete event simulation and to define the network topology. Network nodes generate packets destined for other nodes. Such packets travel from source to destination via shortest paths. The packet generation regime follows a Poisson process by default, but it can be addressed differently. Nodes have internal structure, and implement packet queue, packet sink, and packet forwarding, among other features within a minimal structure. Networks packets that travel between the nodes are a Python dictionary carrying any information intended. Nodes and links at the simulator have the properties summarised in Table 1.

TABLE 1. Main elements and properties of the network simulator. Nodes are named {'n0', 'n1', 'n2' } depending on the router station characteristics (in this case: super, regional, metro). Links are defined by connecting node types: 'ni-nj', $i, j = 0, 1, 2$.

Node ni : definition	Link ni-nj : definition
ni : {	ni-nj : {
node_pkt_rate	link_transm_delay
node_proc_delay	link_capacity
node_queue_check	}
node_queue_cutoff	
}	

The network simulator has been adapted to be able to include disruption and degradation events on the fly, as well as recovery of such events after a certain time. This is key for creating a benchmark of networks having controlled disruptions and to verify the performance of anomaly detection using the proposed GFT-TS framework. The GFT-TS based on matrix profile is, in parallel, compared to the use of the matrix profile algorithm at each of the network nodes, which means over the vertex domain.

The performance of the methodologies has been tested over 10 small-world and 10 scale-free synthetic network topologies with the same number of nodes as those associated with the UK backbone network (Section V). The disruption at each test has been introduced randomly in time as well as with respect to the node(s) affected. Such a disruption is based on an increase of the "node_proc_delay" from 0.013 to a uniformly distributed value, defined by $\mathcal{U}(1.8, 2.2)$ at randomly selected node/s during a time of 10 time units (out of the 720 time units of each simulation). The experiment was repeated 10 times per network. The comparisons have been computed regarding the sensitivity and specificity. Sensitivity or true positive rate, TPR, measures the ratio of anomalies classified as such by the method. Specificity or true negative rate, TNR, measures the ratio of regular patterns classified as such. The results are shown in Table 2.

TABLE 2. Benchmark of synthetic networks by topology. Average results for GFT-TS mining and comparison to vertex-domain analysis.

Method	Topology	TPR [%]	TNR [%]
GFT-TS	Small-world	80	99
	Scale-free	85	99
Vertex TS	Small-world	80	98
	Scale-free	80	97

Table 2 shows the benefit of mining the GFT-TS in its performance results.

In comparison to the analysis at the vertex domain (1 time series per node). GFT-TS shown similar or superior performance than vertex domain analyses, as well as a naturally lower computational burden and faster detection speed. This is since GFT-TS are based on having 1 vector (sequence of GFT at each eigenvalue) at each time unit it is possible to work with a detection window in matrix profile equal in length to the vector length per time-unit. This is to avoid any delay in the decision-making, likely necessary when working with individual points per time unit. The TPR of the Table 2 experiments follows (4),

$$\text{TPR} = \frac{\#\text{true positive}}{\#\text{anomalies}} = \frac{\#\text{true positive}}{\#\text{true positive} \ + \ \#\text{false negative}}; \tag{4}$$

while the TNR is the result of (5),

$$\text{TNR} = \frac{\#\text{true negative}}{\#\text{not anomalies}} = \frac{\#\text{true negative}}{\#\text{true negative} + \#\text{false positive}}. \tag{5}$$

Where '#true positive' is the number of classified anomalies that actually are anomalies; '#true negative' is the number of classified non-anomalies that actually are not. In contrast, '#false positive' is the number of classified anomalies that are not anomalies; '#false negative' is the number of not classified anomalies that actually are anomalies.

In both analyses - TPR and TNR - GFT-TS shows better results than those related to the vertex domain time series. This especially occurs for the specificity (TNR) results on the tested scale-free network topologies. In addition, the dimension of the problem for GFT-TS is smaller than the vertex domain approach. Although both approaches use matrix profile, GFT-TS is based on an analysis of a single time series of length $K \times T$, where K is the number of the top eigenvalues considered for the computations. The vertex-domain procedure runs matrix profile on the time series signal observed at each node, and so runs n many times, each for a time series of length T. Therefore, the larger the network the larger the difference between K and n, where $K \leq n$ in all cases.

V. THE UK BACKBONE NETWORK

The case study presented in this paper is the long-haul backbone network of one of the main ISPs in the UK. The network comprises 103 router PoPs (nodes) and 309 fibre-optic cables (links). Among the router-PoPs there are: 4 super-router PoPs, 9 regional-router PoPs, and 90 metro-router PoPs. The nodes are coded with their identification number (sorted from 0 to $n-1$) in the database where they are stored and a letter corresponding the type of router they represent: 's' (super), 'r' (regional) and 'm' (metro). Fig. 3 shows the network layout of this infrastructure.

A first approach to further analyse the network is to decompose it into 2 slices: one corresponding to the xIC (super and regional PoPs) and another corresponding to the xOC (regional and metro PoPs). Table 3 shows topological characteristics of these sub-networks comparing the xIC to a Watts–Strogatz random model network, and the xOC to a Barabási–Albert random model network, where each random network has the same dimensions as the network it is being compared to (13 and 92 nodes, respectively). The topological measures computed in Table 3 are the values of the average node degree, d_G; average shortest path length, l_G; network efficiency, E_G; and the clustering coefficient, C_G.

TABLE 3. Comparison of network slice topologies in the backbone network and against reference synthetic random graphs.

Network	Param.	d_G	l_G	E_G	C_G
xIC	–	6.31	1.47	0.76	0.66
xOC	–	2.85	3.23	0.35	0.24
Watts-Strogatz	$n = 13, p = 0.5$	6.31	1.47	0.76	0.52
Barabási-Albert	$n = 92, m = 1$	1.98	4.33	0.27	0.00

The column 'Param.' in Table 3 describes the parameters used to generate the random networks: n is the number of nodes, p is the probability of connecting one node to another for Watts-Strogatz models, and m is the number of links to connect from a new node to existing nodes for Barabási-Albert models. Table 3's results confirm that the xIC network has a small-world topology (results almost identical to the Watts-Strogatz generated model), while the xOC network can be considered to have a scale-free topology (results closer to those seen for the Barabási-Albert generated model). Results are discussed later in this paper as the analyses for the xIC and xOC networks are handled separately. The topological analysis is complemented with data resulting from the signal passing through the core network graph. This is 24h of data registered every 2 minutes, creating a TS with a total signal length of 720 time units for each network node. The data collected is from 14:00 of January 14th 2020 to 13:58 of January 15th 2020.

A. EXTENDED INNER CORE NETWORK

Out of the network represented in Fig. 3, the extended inner core (xIC) network comprises of super and regional PoPs, inner core and regional, respectively. The study of the xIC topology is of interest given that in normal conditions, internet data traffic in the inner core only addresses other inner core and regional PoPs. Fig. 4 represents this xIC network, showing a small-world topology, highly resilient both to random failures and targeted attacks.

A GFT is computed at each time step of the 720 time units. A filter to count the signal values (data traffic) of the immediate neighbourhood at each node has been used. As discussed in Section III, the related GFT operations only consider the top 10 eigenvectors of $\boldsymbol{U}$. That is: $\{\lambda_2, \ldots \lambda_{11}\}$, since the first eigenvalue not analysed. Then, GFT is a sequence of 10 values per each time step. Each sequence is concatenated to another in order, up to the 720 time units. Fig. 7 shows the results of matrix profile analysis on the GFT-TS for the inner core network.

The top block of Table 4 summarises the results found from the analysis shown in Fig. 7. The samples, or time stamps, related to discords are converted into day time, since each time step is 2 minutes and the TS started at 14:00. For the xIC network, the top discords mostly occur during early morning, between 5am and 7am. The fifth of the top discords happened near midnight. An explanation to that may lie in the overall process of connection and disconnection to/from internet by the users. The nodes having more influence on these discord events are also identified for 3 out of the 5 discords. There are cases in which any node could be categorised as influential in the discord. That means that the network had a global abnormal behaviour with no specific node responsible for that (marked as '–' in Table 4). In the xIC network, the nodes having a significant participation in the discords found (so-called influential nodes, following Algorithm 1), are one single node at each time and, in all cases, the node is classified as a super-node.

Following the process explained in the Section IV, it is possible to ascertain specific nodes, if any, associated with

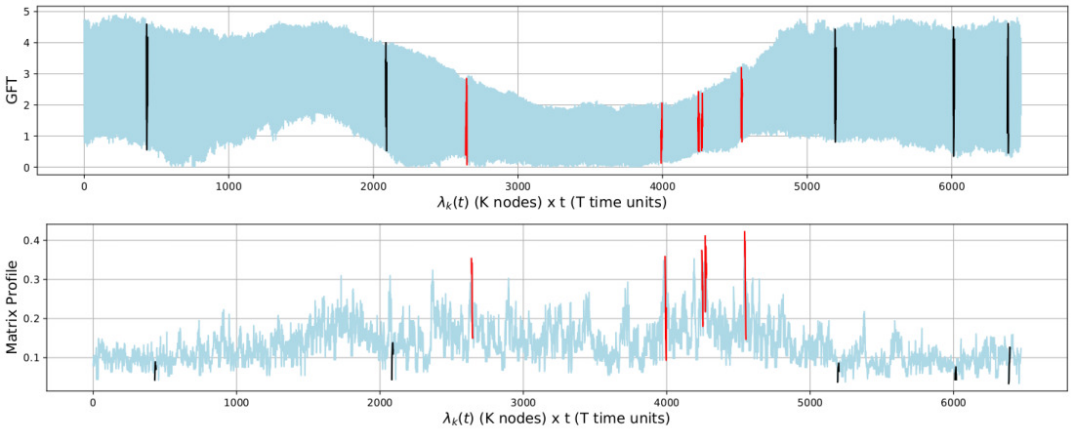

FIGURE 7. xIC network: GFT-TS (on top) and its associated matrix profile (at bottom). Top 5 discords marked in red. Top 5 motifs marked in black.

the discords detected. The first step in this task is to compute the deviation from the k-trimmed GFT-TS mean.

Computing the deviation from the k-trimmed GFT mean is key for the analysis as we can filter which eigenvalues at the time of the network (matrix profile) discord have the biggest impact on it. Then, out of the eigenvector matrix at the time of the discord, the corresponding columns to the highlighted eigenvectors are selected. Each eigenvector comprises a combination of nodes, out of which it is possible to get the node, or set of nodes, with the highest weight in such a combination and hence more strongly related to the discord than all other nodes.

B. EXTENDED OUTER-CORE NETWORK

Just by removing the super nodes from the entire network, the system dimensions are significantly reduced and new patterns may come out, creating the extended outer-core (xOC) network. The new topology is of interest because PoPs remain belonging to the regional network ISPs and most of the traffic data providing services to end-users circulate between them in regular conditions. This traffic is also shown through the connections from the regional to metro nodes.

Fig. 5 provides a visualisation of the entire network without super nodes (xOC network). The remaining structure clearly follow a scale-free topology. In contrast to the small-world which shows a high resilience facing any disruption scenario, scale-free networks are only resilient to random failures but vulnerable to targeted attacks. This is the case of any regional router failure in Fig. 5 leading to disconnected networks.

The bottom block of Table 4 summarises the results found from the analysis shown in Fig. 8. The analysis for the xOC network repeats the process done for the xIC network. For the xOC network, the top discords mostly occur near midnight, approximately between 11pm and 12:30am. Similar to the above explanation for xIC, a possible explanation for the discords in xOC seems to be the process of disconnection from internet by the users. However, certain programmed maintenance activities may have an influence in detecting such discords at this time of the day. The nodes having more influence on these discord events are also identified for 3 out of the 5 discords. In all the cases, such influential nodes belong to the metro network.

C. DISCUSSION

The bottom block of Table 4 summarises the results found from the analysis that Fig. 8 shows. The explanation for these results is similar to the one given for the xIC network case. For the xOC network, the top discords occur around midnight. An explanation to that may lie in the many users disconnecting from the internet at around this time. The nodes with more influence on these discord events are also identified for 3 out of the 5 discords. In those cases, as is the case for the xIC network, more than 1 single network node has been identified as influential.

Overall, Table 4 shows how the patterns in the xIC are not usually related to a particular node but to the all the nodes in the network. This is in line with how the small-world topology of the xIC makes to have a stronger relationship between its nodes than in the xOC's case. The time of discords is another difference between xIC and xOC networks. Most of the top discords found for xIC are in the early morning (interpretable as the time in which internet traffic starts a day cycle), while all of the top xOC discords occur at night (internet traffic ends a day cycle). The xOC discords are better localised in specific nodes than those in xIC. Such discords are in metro network nodes. A plausible explanation is that the detected discords are actually representing maintenance actions carried out not at these specific nodes but at parts of the infrastructure closer to the end-user such as digital subscriber line access multiplexers or optical line terminations.

The topological position of regional routers (between inner core and metro networks) means that they are more vulnerable than super routers to failure. This is due to their connectivity to metro routers in the extended outer-core network

TABLE 4. Summary of the top 5 discords and motifs in the GFT-TS of xIC and xOC networks.

Network	Top discords Time	Top discords Influence nodes	Top motifs Time	Top motifs Influence nodes
xIC	06:50	0s	13:40	–
	05:50	6s	09:14	–
	05:44	6s	15:36	–
	04:48	–	12:16	–
	23:48	–	21:44	–
xOC	00:26	4m, 44m, 77m, 90m	09:16	16m, 55m, 62m, 67r, 69m, 91m
	22:58	7m, 25m, 30m, 54m	10:24	16m, 55m, 62m, 67r, 69m, 91m
	00:08	4m, 44m, 77m, 90m	08:58	5m, 27m, 53m, 68m, 72m, 85m
	00:04	–	15:06	–
	00:00	–	08:02	–

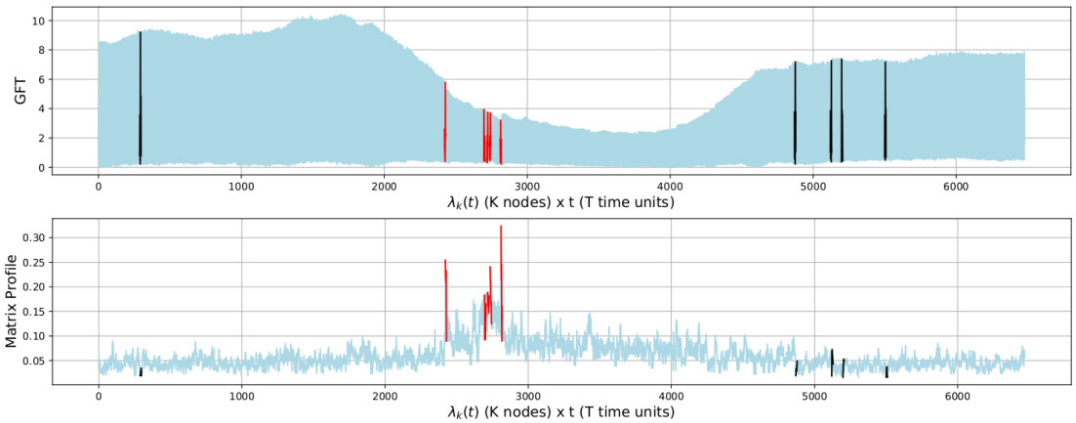

FIGURE 8. xOC network: GFT-TS (on top) and its associated matrix profile (at bottom). Top 5 discords marked in red. Top 5 motifs marked in black.

that makes regional routers to play the role of node-hubs in a scale-free network. Still, the topological vulnerability of regional routers does not mean a higher likelihood of failure. Actually, as mentioned above, the detected anomalies and discords are actually located at the metro network.

The outcome of the GFT-TS proposed herein is useful for a network operator since it provides additional knowledge for internet backbone network operation and management. The results point out which nodes at which time present the most irregular patterns or anomalies. This may be useful to implement rerouting protocols to optimise network traffic by knowing when and where they are necessary. On the other hand, with the information about motifs, the network operator can implement certain adjustments, inspections and other operations related to scheduled maintenance under controlled conditions. On top of this, given the computational efficiency of matrix profile methods such as STAMP and the dimension reduction of the proposed GFT-TS framework, it will be possible to carry-out near real-time operations, for instance by detecting anomalies within the frequency of data sampling (2 minutes in this case) in which data traffic changes are registered.

VI. CONCLUSIONS

This paper proposes a holistic framework analysis comprising concepts of complex networks, graph signal processing, and time series (TS) mining. The aim is to extract regular and irregular signal patterns that occur on the network structure. This leads to better network control and management, providing a deeper understanding of a given network's topology and its dynamics. The proposed framework analysis was demonstrated in the study of a long-haul, internet backbone network. The relevance of this paper is due to the high requirements placed on communication speed and resilience in these backbone networks, for which a comprehensive multi-view approach enables rapid detection of patterns and anomalies that are normally more diffuse when analysed from a single perspective.

One important finding of this work is about the topological importance of the regional routers. This is since they are densely connected to super router stations (small-world topology), so communication between super and regional routers will show high performance and resilience, since any super router will have the straightforward backup of other super routers, along with regional routers. These subnetworks encompass the named extended inner-core network. However, it is observed that most of internet traffic

discords and anomalies are in metro routers. This is likely since they provide a gateway from the core network to smaller area networks and are, therefore, closer to an end-user infrastructure than the rest of the backbone network.

The novelties proposed in this paper are highlighted below:

- A multi-view topological approach to facilitate the extraction of knowledge out of a network. Importantly, this is a network slice, not a partition. That means that a set of network nodes is allowed to appear in more than 1 view or slice of the problem, although the analysis may be done from a different perspective.
- A TS mining approach for graph-Fourier transform (GFT) sequences. By construction, each of these sequences contains information on the network topology and the signal on the network at each time step. The provision of a methodology to extract information on the evolution of the GFT brings the benefits of considering all the information available to any further decision-making support. Both, regular and irregular patterns (motifs and discords, respectively) have been extracted to inform network managers and network traffic engineers.
- An automatic, simple way to inverse the solution obtained at a spectral domain to a vertex domain has shown to be able to infer which specific network nodes play an influential role in the network spectrum.

Overall, this paper contributes with a novel methodology for the analysis of complex networks dynamics. This methodology is demonstrated with a case study within the UK's internet backbone network. However, today, the world is highly interconnected by the internet of things and the use of cyber-physical systems. The methodology presented herein has the capability to make the difference in its application to the analysis of the plethora of networks governing infrastructures, companies, and social systems.

Future work will explore further integrative models for TS and complex network analysis. In addition, the authors are looking forward to bring specific anomaly detection methods to the GFT-TS. Future research on GFT-TS can be also approached within a interpretable machine learning framework due to the complete knowledge of the network state and signal process/es over time.

.

APPENDIX A BASIC DEFINITIONS IN COMPLEX NETWORK TOPOLOGY

From a mathematical point of view, it is possible to define a vertex (node) set as $\mathcal{V} = \{v_0, v_1, \ldots, v_{N-1}\}$, and the edges (links) between them as $\mathcal{E} = \{e_0, e_1, \ldots, e_{M-1}\}$. Both elements sets, vertices and edges, make a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of n vertices and m edges. This notation helps to introduce a number of graph theoretical concepts used to define common topological features of small-world and scale-free topologies [74]–[78].

- Average node degree (d_G): This is the average number of links connected to each node in a network. It is a measure of the relative importance of each node in a network, from the point of view of their connectivity. Average node degree is computed following Equation (6).

$$d_G = \frac{1}{N}\sum_{j=1}^{n} d_j \tag{6}$$

where n is the number of nodes of the network, G, and d_j is the degree of node j.

- Average shortest path length (l_G): This is the average number of links along the shortest paths for all possible pairs of nodes in a network. It is a measure of the efficiency on how information flows on a network. The average shortest path length, l_G, is computed following Equation (7).

$$l_G = \frac{1}{n(n-1)} \cdot \sum_{i \neq j} \delta(i,j) \tag{7}$$

where $\delta(i,j)$ is the shortest path length between a node i and another node j.

- Network efficiency (E_G): From a global point of view, this is the average of how difficult is to share information between all possible pairs of nodes in a network, as Equation (8) states.

$$E_G = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{\delta(i,j)} \tag{8}$$

At a local level, the efficiency of a certain node is a measure of how well the information is exchanged on its neighbourhood after the removal of such a node from the network.

- Clustering coefficient (C_G): This index measures the degree in which nodes in a network tend to cluster together. There are multiple ways to approach this index. The global clustering coefficient is a measure based on the proportion of patterns in the network called triplets (closed triangles). The local clustering coefficient of a node computes how close such a node is to be part of a clique. A formal definition of clustering coefficient for the node i is given by Equation (9),

$$C_G(i) = \frac{\tau(i)}{\tau_{max}(i)} = \frac{\tau(i)}{\binom{d_i}{2}} = \frac{2\tau(i)}{d_i d_i - 1)} \tag{9}$$

where $\tau(i)$ is the number of triangles containing node i and d_i is the degree of the node i. Equation (10) gives the expression for the clustering coefficient of the entire network.

$$C_G = \frac{\sum_{i:s_i>1} C_G(i)}{N_{d>1}} \tag{10}$$

APPENDIX B DISTANCE DEFINITION FOR MATRIX PROFILE

Matrix profile is built upon an Euclidean distance metric. Given 2 TS: $x = \{x_1, \ldots, x_n\}$ and $y = \{y_1, \ldots, y_n\}$; their z-normalised Euclidean distance is defined by (11),

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (\hat{x}_i - \hat{y}_i)^2}, \tag{11}$$

where $\hat{x}_i = \frac{x_i - \mu_x}{\sigma_x}$ and $\hat{y}_i = \frac{y_i - \mu_y}{\sigma_y}$.

The Pearson's correlation coefficient, expressed in (12), is directly related to the Euclidean distance.

$$\begin{aligned} corr(x,y) &= \frac{(E(x)-\mu_x)(E(y)-\mu_y)}{\sigma_x \sigma_y} \\ &= \frac{\sum_{i=1}^{m} x_i y_i - m\mu_x\mu_y}{m\sigma_x\sigma_y}. \end{aligned} \tag{12}$$

Pearson's correlation coefficient is therefore defined by the summary statistics: $\sum_{i=1}^{m} x_i y_i$, $\sum_{i=1}^{m} x_i$, $\sum_{i=1}^{m} y_i$, $\sum_{i=1}^{m} x_i^2$, and $\sum_{i=1}^{m} y_i^2$. The coefficient can also be expressed, from a different approach, with respect to the Euclidean distance: $d(\hat{x},\hat{y}) = \sqrt{2m(1 - corr(x,y))}$. This expression brings sensible differences to the Euclidean distance, such as [79]:

- Pearson's correlation coefficient does not fulfil the triangular inequality, while Euclidean distance does.
- There is an inverse relationship between Pearson's correlation coefficient and Euclidean distance. Therefore, one strategy to maximise the correlation coefficient is to minimise the Euclidean distance.
- Pearson's correlation coefficient is bounded between -1 and 1, while z-normalised Euclidean distance moves in a domain between 0 and a positive number dependent on m.

Taking into account the aforementioned considerations, (13) expresses the distance metric that is proposed in matrix profile,

$$d(\hat{x},\hat{y}) = \sqrt{2m\left(1 - \frac{\sum_{i=1}^{m} x_i y_i - m\mu_x\mu_y}{m\sigma_x\sigma_y}\right)}, \tag{13}$$

subject to every sub-sequence is z-normalised at all datasets, which has been shown to improve the accuracy and performance of the overall process [80].

REFERENCES

[1] D Johnson and T Gilfedder. Evolution of optical core networks. BT Technology Journal, 25(3-4):57–64, 2007.

[2] Andjelka Kelic, Michael Mitchell, and Donald Shirah. Modeling the internet. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2017.

[3] Arun K. Majumdar. Chapter 9 - Potential applications for broadband global internet connectivity. In Arun K. Majumdar, editor, Optical Wireless Communications for Broadband Global Internet Connectivity, pages 259 – 268. Elsevier, 2019.

[4] Nathaniel Evans and William Horsthemke. Regional critical infrastructure. In Cyber Resilience of Systems and Networks, pages 355–380. Springer, 2019.

[5] Alexandros Stavdas. Core and metro networks. John Wiley & Sons, 2010.

[6] Prabhat Kumar and MP Singh. IP based services. In 2012 1st International Conference on Recent Advances in Information Technology (RAIT), pages 214–219. IEEE, 2012.

[7] Tristan A Shatto and Egemen K Cetinkaya. Spectral analysis of backbone networks against targeted attacks. In DRCN 2017-Design of Reliable Communication Networks; 13th International Conference, pages 1–8. VDE, 2017.

[8] Ramakrishnan Durairajan, Paul Barford, Joel Sommers, and Walter Willinger. Intertubes: A study of the us long-haul fiber-optic infrastructure. In Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, pages 565–578, 2015.

[9] Guillermo García-Pérez, Marián Boguñá, and M Ángeles Serrano. Multiscale unfolding of real networks by geometric renormalization. Nature Physics, 14(6):583–589, 2018.

[10] Christopher R Myers, David C Vollum, and Evan B Goldstein. Mobile Systems IV. Technical report, Worcester Polytechnic Institute, 2020.

[11] Ireneusz Jabłoński. Graph signal processing in applications to sensor networks, smart grids, and smart cities. IEEE Sensors Journal, 17(23):7659–7666, 2017.

[12] Deepshikha Bhargava and Sonali Vyas. Pervasive Computing: A Networking Perspective And Future Directions. Springer, 2019.

[13] Weiwei Fang, Zishuo Wang, Jaime LIoret, Daqiang Zhang, and Zhijun Yang. Optimising data placement and traffic routing for energy saving in backbone networks. Transactions on Emerging Telecommunications Technologies, 25(9):914–925, 2014.

[14] Fragkiskos Papadopoulos and Konstantinos Psounis. Efficient identification of uncongested internet links for topology downscaling. ACM SIGCOMM Computer Communication Review, 37(5):39–52, 2007.

[15] Changqiao Xu, Jia Zhao, and Gabriel-Miro Muntean. Congestion control design for multipath transport protocols: A survey. IEEE Communications Surveys & Tutorials, 18(4):2948–2969, 2016.

[16] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst. Graph signal processing: Overview, challenges, and applications. Proceedings of the IEEE, 106(5):808–828, 2018.

[17] Xiao Fan Wang and Guanrong Chen. Complex networks: small-world, scale-free and beyond. IEEE Circuits and Systems magazine, 3(1):6–20, 2003.

[18] Amin Fakhrazari and Hamid Vakilzadian. A survey on time series data mining. In 2017 IEEE International Conference on Electro Information Technology (EIT), pages 476–481. IEEE, 2017.

[19] Yan Zhu, Shaghayegh Gharghabi, Diego Furtado Silva, Hoang Anh Dau, Chin-Chia Michael Yeh, Nader Shakibay Senobari, Abdulaziz Almaslukh, Kaveh Kamgar, Zachary Zimmerman, Gareth Funning, Abdullah Mueen, and Eamonn Keogh. The Swiss army knife of time series data mining: ten useful things you can do with the matrix profile and ten lines of code. Data Mining and Knowledge Discovery, 34(4):949–979, 2020.

[20] Mark Coates, Alfred Hero, Robert Nowak, and Bin Yu. Internet tomography. IEEE Signal Processing magazine, 19(3):47–65, 2002.

[21] Daniel J Blumenthal, John E Bowers, Lavanya Rau, Hsu-Feng Chou, Suresh Rangarajan, Wei Wang, and KN Poulsen. Optical signal processing for optical packet switching networks. IEEE Communications magazine, 41(2):S23–S29, 2003.

[22] Gong-Ru Lin and Pai-Shen Hsueh. Spectral characteristics of a regenerative semiconductor optical amplifier mutually injection locked with a fabry-perot laser diode. Applied Optics, 43(1):153–159, 2004.

[23] Katsuhiro Shimizu, Abdullah Al Amin, Shunsuke Tanaka, Mitsuru Takenaka, Toshiharu Miyahara, Tatsuo Hatta, Kuniaki Motoshima, Motoshi Ono, Yuki Kondo, Naoki Sugimoto, et al. High-speed gain control for EDWA in optical burst switching node employing wavelength converters. In Optical Amplifiers and Their Applications, page OTuA4. Optical Society of America, 2006.

[24] Soo-Yeon Ji, Seonho Choi, and Dong Hyun Jeong. Designing an internet traffic predictive model by applying a signal processing method. Journal of Network and Systems Management, 23(4):998–1015, 2015.

[25] Pablo Poudereux, Álvaro Hernández, Raúl Mateos, Freddy A Pinto-Benel, and Fernando Cruz-Roldán. Design of a filter bank multi-carrier system for broadband power line communications. Signal Processing, 128:57–67, 2016.

[26] Minhao Pu, Hao Hu, Luisa Ottaviano, Elizaveta Semenova, Dragana Vukovic, Leif Katsuo Oxenløwe, and Kresten Yvind. Ultra-efficient and broadband nonlinear algaas-on-insulator chip for low-power optical signal processing. Laser & Photonics Reviews, 12(12):1800111, 2018.

[27] Reuven Cohen and Shlomo Havlin. Complex networks: structure, robustness and function. Cambridge University Press, 2010.

[28] Piet Van Mieghem. Performance analysis of complex networks and systems. Cambridge University Press, 2014.

[29] Piet Van Mieghem. Graph spectra for complex networks. Cambridge University Press, 2010.

[30] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. Reviews of Modern Physics, 87(3):925, 2015.

[31] Dane Taylor, Sean A Myers, Aaron Clauset, Mason A Porter, and Peter J Mucha. Eigenvector-based centrality measures for temporal networks. Multiscale Modeling & Simulation, 15(1):537–574, 2017.
[32] Rogier Noldus and Piet Van Mieghem. Assortativity in complex networks. Journal of Complex Networks, 3(4):507–542, 2015.
[33] Jianxi Gao, Sergey V Buldyrev, H Eugene Stanley, Xiaoming Xu, and Shlomo Havlin. Percolation of a general network of networks. Physical Review E, 88(6):062816, 2013.
[34] Massimiliano Zanin, David Papo, Pedro A Sousa, Ernestina Menasalvas, Andrea Nicchi, Elaine Kubik, and Stefano Boccaletti. Combining complex networks and data mining: why and how. Physics Reports, 635:1–44, 2016.
[35] Guanhua Yan, Stephan Eidenbenz, Sunil Thulasidasan, Pallab Datta, and Venkatesh Ramaswamy. Criticality analysis of Internet infrastructure. Computer Networks, 54(7):1169–1182, 2010.
[36] David Hutchison and James PG Sterbenz. Architecture and design for resilient networked systems. Computer Communications, 131:13–21, 2018.
[37] Mark Winther. Tier 1 ISPs: What they are and why they are important. IDC White Paper, pages 1–13, 2006.
[38] Andrew D Ellis, David Cotter, S Ibrahim, Ruwan Weerasuriya, Chi Wai Chow, Juerg Leuthold, Wolfgang Freude, Stylianos Sygletos, Philipp Vorreau, R Bonk, et al. Optical interconnection of core and metro networks. Journal of Optical Networking, 7(11):928–935, 2008.
[39] Yue Zong, Chuan Feng, Yingying Guan, Yejun Liu, and Lei Guo. Virtual network embedding for multi-domain heterogeneous converged optical networks: Issues and challenges. Sensors, 20(9):2655, 2020.
[40] Luıs A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. Proceedings of the National Academy of Sciences, 97(21):11149–11152, 2000.
[41] Vito Latora, Vincenzo Nicosia, and Giovanni Russo. Small-World Networks, page 107–150. Cambridge University Press, 2017.
[42] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. Nature, 393(6684):440–442, 1998.
[43] Yuji Kawai, Jihoon Park, and Minoru Asada. A small-world topology enhances the echo state property and signal propagation in reservoir computing. Neural Networks, 112:15–23, 2019.
[44] Wei Cheng, Jiguo Yu, Feng Zhao, and Xiuzhen Cheng. Ssdnet: Small-world super-dense device-to-device wireless networks. IEEE Network, 32(1):186–192, 2017.
[45] Shuang Liu and Qingyun Wang. Outer synchronization of small-world networks by a second-order sliding mode controller. Nonlinear Dynamics, 89(3):1817–1826, 2017.
[46] Xi Zhang, Choujun Zhan, and K Tse Chi. Modeling the dynamics of cascading failures in power systems. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 7(2):192–204, 2017.
[47] Hidefumi Sawai. Exploring a new small-world network for real-world applications. In International Conference on Networked Digital Technologies, pages 90–101. Springer, 2012.
[48] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. Science, 286(5439):509–512, 1999.
[49] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. Physical Review Letters, 86(14):3200, 2001.
[50] Di Zhou, Jianxi Gao, H Eugene Stanley, and Shlomo Havlin. Percolation of partially interdependent scale-free networks. Physical Review E, 87(5):052812, 2013.
[51] Hao-Ran Liu, Yan-Long Hu, Rong-Rong Yin, and Yu-Jing Deng. Cascading failure model of scale-free topology for avoiding node failure. Neurocomputing, 260:443–448, 2017.
[52] Yadong Li, Danlan Li, Wenqiang Cui, and Rui Zhang. Research based on osi model. In 2011 IEEE 3rd International Conference on Communication Software and Networks, pages 554–557. IEEE, 2011.
[53] Bojan Mohar, Y Alavi, G Chartrand, and OR Oellermann. The laplacian spectrum of graphs. Graph Theory, Combinatorics, and Applications, 2(871-898):12, 1991.
[54] Aden Forrow, Francis G Woodhouse, and Jörn Dunkel. Functional control of network dynamics using designed laplacian spectra. Physical Review X, 8(4):041043, 2018.
[55] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst. Vertex-frequency analysis on graphs. Applied and Computational Harmonic Analysis, 40(2):260–291, 2016.
[56] Ljubiša Stanković, Miloš Daković, and Ervin Sejdić. Introduction to graph signal processing. In Vertex-Frequency Analysis of Graph Signals, pages 3–108. Springer, 2019.
[57] Almerima Jamakovic and Steve Uhlig. On the relationships between topological measures in real-world networks. Networks & Heterogeneous Media, 3(2):345, 2008.
[58] Sahar Torkamani and Volker Lohweg. Survey on time series motif discovery. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7(2):e1199, 2017.
[59] Ross S Bowden and Brenton R Clarke. A single series representation of multiple independent ARMA processes. Journal of Time Series Analysis, 33(2):304–311, 2012.
[60] Emily B Fox, Michael C Hughes, Erik B Sudderth, Michael I Jordan, et al. Joint modeling of multiple time series via the beta process with application to motion capture segmentation. The Annals of Applied Statistics, 8(3):1281–1313, 2014.
[61] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. Time-series clustering–a decade review. Information Systems, 53:16–38, 2015.
[62] Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi. Experiencing SAX: a novel symbolic representation of time series. Data Mining and Knowledge Discovery, 15(2):107–144, 2007.
[63] Chin-Chia Michael Yeh, Nickolas Kavantzas, and Eamonn Keogh. Matrix profile vi: Meaningful multidimensional motif discovery. In 2017 IEEE 17th International Conference on Data Mining (ICDM), pages 565–574. IEEE, 2017.
[64] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Zachary Zimmerman, Diego Furtado Silva, Abdullah Mueen, and Eamonn Keogh. Time series joins, motifs, discords and shapelets: a unifying view that exploits the matrix profile. Data Mining and Knowledge Discovery, 32(1):83–123, 2018.
[65] Pranav Patel, Eamonn Keogh, Jessica Lin, and Stefano Lonardi. Mining motifs in massive time series databases. In 2002 IEEE 2nd International Conference on Data Mining (ICDM), pages 370–377. IEEE, 2002.
[66] Eamonn Keogh, Jessica Lin, and Ada Fu. Hot SAX: Efficiently finding the most unusual time series subsequence. In 2005 IEEE 5th International Conference on Data Mining (ICDM), page 8. IEEE, 2005.
[67] Chin-Chia Michael Yeh, Nickolas Kavantzas, and Eamonn Keogh. Matrix profile iv: using weakly labeled time series to predict outcomes. Proceedings of the VLDB Endowment, 10(12):1802–1812, 2017.
[68] Simon Duque Anton, Lia Ahrens, Daniel Fraunholz, and Hans Dieter Schotten. Time is of the essence: Machine learning-based intrusion detection in industrial time series data. In 2018 IEEE International Conference on Data Mining Workshops (ICDMW), pages 1–6. IEEE, 2018.
[69] Yan Zhu, Zachary Zimmerman, Nader Shakibay Senobari, Chin-Chia Michael Yeh, Gareth Funning, Abdullah Mueen, Philip Brisk, and Eamonn Keogh. Matrix profile ii: Exploiting a novel algorithm and gpus to break the one hundred million barrier for time series motifs and joins. In 2016 IEEE 16th International Conference on Data Mining (ICDM), pages 739–748. IEEE, 2016.
[70] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn Keogh. Matrix profile i: all pairs similarity joins for time series: a unifying view that includes motifs, discords and shapelets. In 2016 IEEE 16th international conference on data mining (ICDM), pages 1317–1322. IEEE, 2016.
[71] Luis Torgo and Rita Ribeiro. Predicting outliers. In European Conference on Principles of Data Mining and Knowledge Discovery, pages 447–458. Springer, 2003.
[72] Vladimir Likic and Kamran Shafi. Battlespace mobile/ad hoc communication networks: Performance, vulnerability and resilience. In Data and Decision Sciences in Action, pages 303–314. Springer, 2018.
[73] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. In Gäel Varoquaux, Travis Vaught, and Jarrod Millman, editors, 7th Python in Science Conference (SciPy2008), pages 11–15, Pasadena, CA USA, 2008.
[74] Mark EJ Newman. The structure and function of complex networks. SIAM Review, 45(2):167–256, 2003.
[75] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. Physics Reports, 424(4-5):175–308, 2006.
[76] Anna Nagurney and Qiang Qiang. A network efficiency measure with application to critical infrastructure networks. Journal of Global Optimization, 40(1-3):261–275, 2008.
[77] Albert-László Barabási et al. Network science. Cambridge University Press, 2016.

[78] Manuel Herrera, Marco Pérez-Hernández, Ajith K. Parlikad, and Joaquín Izquierdo. Multi-agent systems and complex networks: Review and applications in systems engineering. Processes, 8(3):312, 2020.
[79] Abdullah Mueen, Suman Nath, and Jie Liu. Fast approximate correlation for massive time-series data. In Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, pages 171–182, 2010.
[80] Reza Akbarinia and Bertrand Cloez. Efficient matrix profile computation using different distance functions. arXiv preprint arXiv:1901.05708, pages 1–11, 2019.