Design Science

Robust design of complex socio-technical systems against seasonal effects: a network motif-based approach

Yinshuang Xiao and Zhenghui Sha (ID)

Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, USA

Abstract

Seasonal effects can significantly impact the robustness of socio-technical systems (STS) to demand fluctuations. There is an increasing need to develop novel design approaches that can support capacity planning decisions for enhancing the robustness of STS against seasonal effects. This paper proposes a new network motif-based approach to supporting capacity planning in STS for an improved seasonal robustness. Network motifs are underlying nonrandom subgraphs within a complex network. In this approach, we introduce three motif-based metrics for system performance evaluation and capacity planning decision-making. The first one is the imbalance score of a motif (e.g., a local service network), the second one is the measurement of a motif's seasonal robustness, and the third one is a capacity planning decision criterion. Based on these three metrics, we validate that the sensitivity of STS performance against seasonal effects is highly correlated with the imbalanced capacity between service nodes in an STS. Correspondingly, we formulate a design optimisation problem to improve the robustness of STS by rebalancing the resources at critical service nodes. To demonstrate the utility of the approach, a case study on Divvy bike-sharing system in Chicago is conducted. With a focus on the size-3 motifs (a subgraph consisting three docked stations), we find that there is a significant correlation between the difference of the number of docks among the stations in a motif and the return/rental performance of such a motif against seasonal changes. Guided by this finding, our design approach can successfully balance out the number of docks between those stations that have caused the most severe seasonal perturbations. The results also imply that the network motifs can be an effective local structural representation in support of STS robust design. Our approach can be generally applied in other STS where the system performances are significantly impacted by seasonal changes, for example, supply chain networks, transportation systems and power grids.

Key words: socio-technical system, seasonal effect, systems robustness, network motif, bike-sharing system

1. Introduction

The notion of *socio-technical* is originally proposed by Trist & Bamforth ([1951]) in the context of labor studies with the idea that both knowledge accumulation and the improvement of work situations should be embraced in a research project. Over the years, there are two typical types of studies related to the socio-technical

concept that can be distinguished. The first type mainly focuses on improving the efficiency of project management (Keating *et al.* 2001; Hassannezhad *et al.* 2019). It highlights the complex interactions between the subjective perceptions of workers and the objective characteristics of work processes (Pan & Scarbrough 1998). In the second type, instead of only considering the operation-side actors (employees in firms, researchers, policymakers, etc.), a complete understanding of the human factors within the socio-technical methodology is achieved by taking the demand-side actors (end users, customers, special-interest groups, etc.) into account (Carayon 2006; Geels & Kemp 2007), but analysing complexity also concurrently increases because of this additional consideration.

When the socio-technical approach is applied to the system development discipline, the term *socio-technical system* (*STS*), as an extension of the socio-technical principle and conventional complex systems, is invited. Inheriting the integrative nature of socio-technical concepts, STS combines the social and technical features into the engineering design framework. On the one hand, considering this combination is beneficial because it makes the simulation models and design approaches more consistent to the real situation. On the other hand, it becomes challenging since complexity is introduced by the unpredictable features of the social aspects. This complexity runs through the three broad stages in the system engineering lifecycle: analysis, design and evaluation (Baxter & Sommerville 2011; ElMaraghy *et al.* 2012), and continues to affect the functionality of the entire system. It is also this complexity that raises a high requirement of the system robustness against various disturbances (Kalsi, Hacker, & Lewis 1999; Gribble 2001).

One typical disturbance that can impact numerous systems across domains is seasonal effects. Taken the bike-sharing system (BSS) as an example, as shown in Figure 1, seasonal effects not only require the systematic design of station distribution and the capacity of each station to fight against varied weather conditions, but could also generate demand fluctuation in different months that affects BSS's operational performance. Another example is the electric vehicles that, according to a recent study (Hao *et al.* 2020), the electricity consumption of electrical taxis in spring and fall is about 15.2 kWh per 10 km, which is 3.3% and 30% lower than that in summer and winter, respectively. Thus, seasonal effects request that the battery functionality is adaptive to varied climates. Similarly, other systems that can be affected by seasonal changes include power grids, agriculture systems (Ten Napel *et al.* 2011; Urruty, Tailliez-Lefebvre, & Huyghe 2016), ecosystems (Ahlström *et al.* 2012) and transportation systems (Sun, Wandelt, & Linke 2015; Markolf *et al.* 2019).

In this paper, we focus on the transportation-based STS where the concept of 'transportation' here is much broader than the conventional one. It can refer to the systems that transport passengers and freight such as the air transportation system, and also represents the exchange of information and energy transmission like the interconnected power grids (Blume 2017). One common feature of these systems is that they are composed of multiple connected single service subsystems, and each subsystem is designed with certain capacity constrained by resources. Here, the capacity in systems engineering represents the volume of products that a production system generates (Martnez-Costa *et al.* 2014) or the storage capability of physical systems, for example, computing system and warehouses of a supply chain network. It is the capacity planning decision that plays a critical role for the system
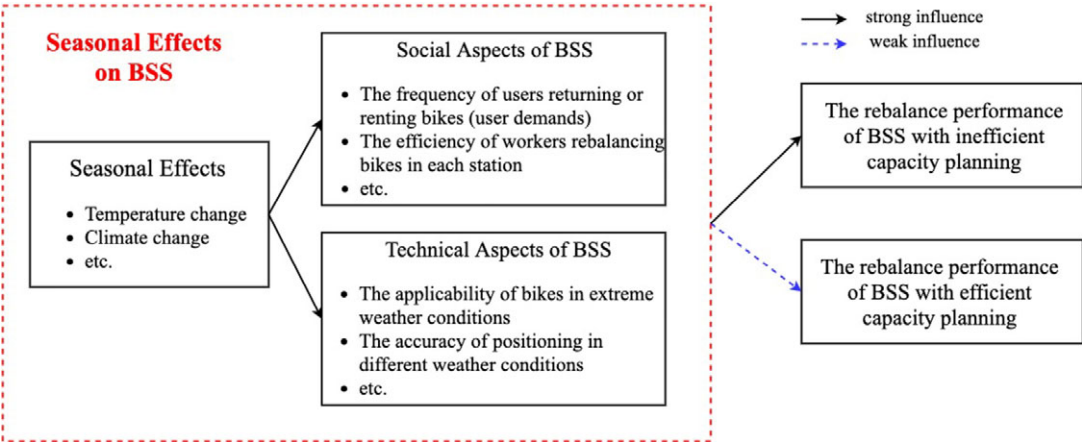
Figure 1. Seasonal effect on bike-sharing system (BSS) with efficient and inefficient capacity planning.

robustness enhancement of those STSs. Inefficient capacity planning can induce either underutilised resources or unfulfilled user demands, thus leading the system to be sensitive to the seasonal changes. Figure 1 explains how seasonal variations and capacity planning can affect BBS's performance in more detail. As shown in the figure, the significance of the influence, either negative or positive, is directly related to the number of docks served (i.e., the capacity planned) in each station of such a BSS.

To mitigate the seasonal fluctuation of the STS performance, we propose a new approach to improving the system robustness by optimising capacity planning decisions based on network motif theory. The present study is based on our prior work (Xiao & Sha 2020) that studies the features of local trip patterns and its correlations to the system-level performance of a BSS using network motifs. In this paper, we further develop a network motif-based framework to support STS robust design in light of seasonal effects. Still, using the bike sharing system as the application context, we show how the the number of docks of critical stations can be optimised to mitigate the seasonal influence on the system's rebalance performance. This design approach can be generally applied to many other transportation-based STSs where the system robustness against seasonal effects is a primary concern, such as air transportation systems and interconnected power grids.

The remainder of this paper is outlined as follows. Section 2 presents the technical background of the complex network and network motif. In Section 3, we introduce the proposed robust design approach in detail. In Section 4, the application of the approach in BSS is presented and the design problem in terms of capacity planning is formulated. In this section, we also assess the seasonal effect on the BSS's rebalance performance and discuss the results for capacity planning optimisation. At the end, Section 5 concludes the paper with closing thoughts and directions for future research.
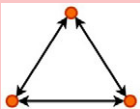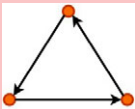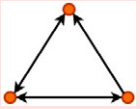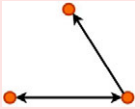
2. Technical background
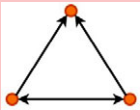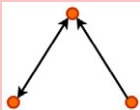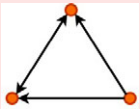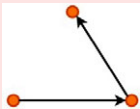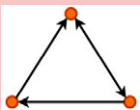
Complex network is a powerful representation for complex systems because of its ability in capturing the interconnectivity and relationship among the subsystems

and individual components. In complex system design and engineering, network science has exhibited its utility in various applications. For example, Sha *et al.* (Sha, Chaudhari, & Panchal 2019; Sha & Panchal 2013a, b; Sha & Panchal 2016) conducted a broad range of research on network-based engineering design of complex systems including autonomous system level Internet and the U.S. domestic air transportation system. Wang et al. performed a series of studies on applying stochastic network models (e.g., the Exponential Random Graph Model) to model the customer-product interactions in vehicle market systems (Wang *et al.* 2016; Wang *et al.* 2016; Fu *et al.* 2017; Bi *et al.* 2018; Sha *et al.* 2018; Wang *et al.* 2018; Sha *et al.* 2019; Cui *et al.* 2020). Regarding the network-based robustness analysis, Cats, Koppenol, & Warnier (2017) developed a robustness assessment model that can indicate the changes of network performance in different link capacity reductions. This model was successfully applied to public transport systems. In another study, Paparistodimou *et al.* (2020) proposed a network generator to support system architectures' robustness analysis in the initial design stages. Focusing on design process robustness, Piccolo, Lehmann, & Maier (2018) presented a bipartite network-based method to investigate the interplay between people and the design activities and its impact on the robustness of design progress. These studies validate the feasibility of using complex networks to research STS robustness. It is worth noting that a distinction exists between the concept of robustness in our study and that in the network science literature. The robustness assessments of complex networks focus on evaluating how the removal of nodes, especially the hub nodes, will impact network topologies, for example, a system's ability to react to failures of its components. However, the robustness defined in this study is investigating whether the nodes with limited cap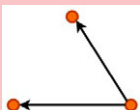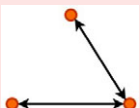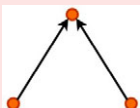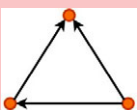acity can maintain their functions when imbalanced link information is transmitting among nodes. In other words, it indicates system's ability, relating to each single service component's capacity planning, to handle the demand fluctuation caused by seasonal effects.

Furthermore, we would like to emphasise the equal importance of both global- and local-level robustness to a system's operational performance where they respectively indicate the capability of all the service nodes or local clusters of nodes to adapt to the seasonal effects. A better understanding of both of these two levels' robustness is helpful to guide stakeholders make a tradeoff between the entire system's performance and subsystems' functionalities. In this paper, since our focus is more on the robustness investigation of local-level service systems, *network motifs* – a fundamental local unit of a network (Wang *et al.* 2020) – is a natural adoption of analysing the local-level system performance.

Network motifs are underlying nonrandom subgraphs within the complex networks. Before named by Milo *et al.* (2002), network motifs experienced a long research period (Stone, Simberloff, & Artzy-Randrup 2019), which was originally considered as certain patterns statistically emerging in real-world networks instead of the same-sized random networks (Holland & Leinhardt 1974). Since then, motif research can be divided into two main subjects where the first one focuses on motif structure explanation (Alon 2007; Paranjape, Benson, & Leskovec 2017; Felmlee *et al.* 2018), and the second one is keen on motif mining algorithms (Kashtan *et al.* 2004; Wernicke & Rasche 2006; Choobdar, Ribeiro, & Silva 2012). A motif can be classified as directed or undirected and can also be categorised by the number of nodes it consists of. There are three commonly studied motifs in existing literature, including size-2 motifs (dyads), size-3 motifs (triads) and size-4 motifs (tetrads)

(Felmlee *et al.* 2018). As the simplest motifs, dyads are essential to the formation of higher-level motifs and the whole network. Triads, also called 'transitivity' motifs, greatly impact the growth of social networks. Tetrads are a newly research focus in recent years, and relevant research comes from a wide range of disciplines, such as biology, electronics and social analysis. Given that the triad is regarded as the foundation in social relationship and the most basic building block of many other complex networks, we decide to adopt size-3 motifs to study STS. Hence, in this paper, only size-3 directed motifs are considered. Their structures and IDs are shown in Table 1 (Rasche & Wernicke 2006).

Table 1. Size-3 directed motif list

ID	Structure	Adjacent matrix	ID	Structure	Adjacent matrix
238		0 1 1 1 0 1 1 1 0	140		0 1 0 0 0 1 1 0 0
174		0 1 0 1 0 1 1 1 0	14		0 0 0 0 0 1 1 1 0
46		0 0 0 1 0 1 1 1 0	164		0 1 0 1 0 0 1 0 0
166		0 1 0 1 0 0 1 1 0	12		0 0 0 0 0 1 1 0 0
102		0 0 1 1 0 0 1 1 0	6		0 0 0 0 0 0 1 1 0
78		0 0 1 0 0 1 1 1 0	36		0 0 0 1 0 0 1 0 0
38		0 0 0 1 0 0 1 1 0			

The motif IDs determined by Rasche & Wernicke (2006) consider each motif's adjacent matrix as a binary representation and transform the binary representation to a decimal number. For example, the binary representation of the decimal number 174 is 010101110, which is consistent with the adjacent matrix of motif 174. Regarding ordering the motifs in Table 1, from top to button and left to right, we rank them based on the number of their arrows from large to small.

There are two common statistics to assess the significance of a network motif in a complex network.

(i) Motif Z-score: Given a graph G and an n-size motif G', the frequency of G' in G is the number of times that G' appeared in G, which is denoted by $F_G(G')$. Then, considering an ensemble of random graphs corresponding to the null-model of G be $\Omega(G)$. $R(G)$ is a set that includes N randomised networks, all of which are from $\Omega(G)$. Accordingly, the Z-score is defined as

$$Z_G(G') = \frac{F_G(G') - u_R(G')}{\sigma_R(G')}, \tag{1}$$

where $u_R(G')$ and $\sigma_R(G')$ represent the mean and standard variation of the frequency in $R(G)$. In general, a higher Z-score indicates that G' is a more significant motif in G. Motifs in a larger network may more easily get a higher Z-score than that in a smaller network (Milo *et al.* 2002).

(ii) P-value: P-value indicates the probability of $F_r(G') > F_G(G')$, where $F_r(G')$ represents the frequency of G' in a random network $r \subset R(G)$. P-value can be obtained by

$$P_G(G') = \frac{1}{N} \sum_{j=1}^{N} \delta(F_r(G') > F_G(G')), \tag{2}$$

where N is the total number of random networks in $R(G)$. δ is the sign function that equals to 1 when $F_r(G') > F_G(G')$, and 0 otherwise. Normally, one motif is treated as a significant pattern if its P-value is smaller than a typical level of significance, normally 0.001, 0.01 or 0.05.

3. The research approach

In this section, we describe our proposed network motif-based approach in a stepwise framework, as shown in Figure 2.

Step 1: Identifying STS system capacity and seasonal effects

The main objective of this step is to formulate the seasonal robust design problem. It includes understanding the interconnections between different parts (as shown in the top-down structure) within an STS and identifying system capacity and seasonal effects. The in-depth understanding on the system helps lay down the foundation for the complex network construction in Step 2. During this step, data preprocessing is needed to organise the dataset by establishing the preprocessing tenets, for example, data preparation, cleaning, normalisation and transformation of data, and so on (Garca, Luengo, & Herrera 2015).

Step 2: Translating STS to complex network and mining network motifs

Based on the understanding of the target system and the robust design that needs to be addressed, the main tasks in Step 2 are to define and construct the complex network that best captures the STS structures as well as to mine the specific motif patterns in the established network. When building the complex network, we first
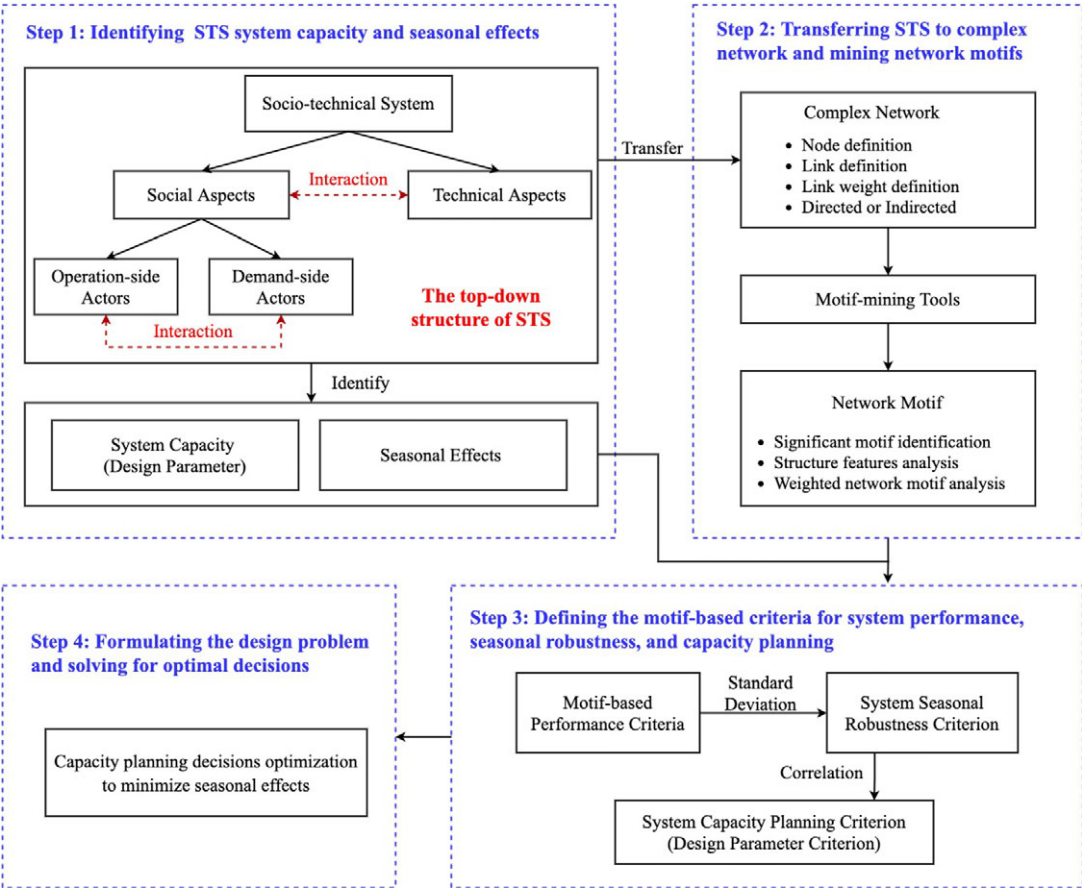
Figure 2. The framework for socio-technical systems (STS) robust design against seasonal effects by capacity planning decisions optimisation.

need to determine the node, node features, link, whether link carries weight or not, weight definition and whether the network is directed or undirected. Secondly, since seasonal data are always time-dependent, two general strategies of handing such a temporal dynamic trait are often used. The first strategy is to treat the year-round data as time-series data, and the second one is to create cross-sectional data at different time steps. Since seasonal information typically changes by month, we adopt the second strategy with the information aggregated from each month. For example, 1 year's dataset can be divided into 12 cross-sectional datasets, which forms 12 networks denoted as G_i ($i = 1,2,\ldots,12$).

After the networks are constructed, motif mining tools like FANMOD (Rasche & Wernicke 2006) and Mfinder (Kashtan *et al.* 2002) are employed to enumerate motifs with a particular size in each network. The significance scores (i.e., *Z*-score and *P*-value) of each pattern can be obtained at the same time. It is worth noting that during the motif mining, link weights are not used and motif patterns are mined only based on link existence. Link weights can be added latter to the mined motif patterns for analysis if necessary. In our study, if a motif pattern is found significant in all the 12 networks, it is treated as a significant pattern throughout that entire year.

$$c = \sum_{j=1}^{a} w_{in\_j} - \sum_{j=1}^{b} w_{out\_j} = \begin{cases} > 0 & \text{In-biased node} \\ = 0 & \text{Balanced node} \\ < 0 & \text{Out-biased node} \end{cases}$$
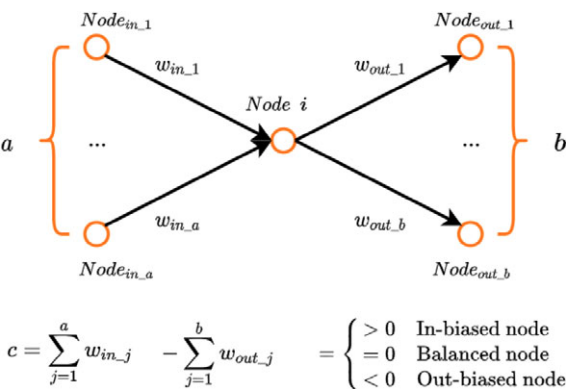
Figure 3. Categorising a node based on its balance performance.

Step 3: Defining the motif-based criteria for system performance, seasonal robustness and capacity planning

Before defining the criteria, we first introduce two node-level performance metrics for directed weighted networks. As shown in Figure 3, assuming a network G has T nodes and for any node $i \subset G$, there are a incoming weighted links and b outgoing weighted links. We define c as the difference between the sum of incoming link weights and the sum of outgoing link weights. Then, if $c = 0$, node i is defined as a *balanced node*; if $c < 0$, node i is considered to be *in-biased node* and if $c > 0$, node i is named *out-biased node*. This way we are able to quantify a node's balance performance in STS.

Next, we extend this node classification to network motifs. Supposed there is a size-3 motif comprising three nodes, *Node 1*, *Node 2* and *Node 3*. A fully connected motif structure is shown in Figure 4. If weight $= 0$ can be used to represent a nonexistent link (e.g., if $w_{12} = 0$, it means there is no link from *Node 1* to *Node 2*), then all the 13 size-3 directed motifs can be described with the following representation. According to the corresponding c values, *Nodes 1, 2 and 3* are divided into three sets: $I(m)$, $O(n)$ and $B(l)$, where m, n and l represent the number of in-biased nodes, out-biased nodes and balanced nodes in each set, and $m + n + l = 3$ holds. Moreover, the relationship among the three c values follows:

$$c_1 + c_2 + c_3 = 0. \tag{3}$$
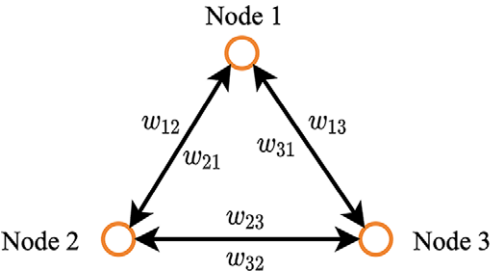
- For Node 1: $c_1 = (w_{21} + w_{31}) - (w_{12} + w_{13})$
- For Node 2: $c_2 = (w_{12} + w_{32}) - (w_{21} + w_{23})$
- For Node 3: $c_3 = (w_{13} + w_{23}) - (w_{31} + w_{32})$

Figure 4. A general motif structure.

Based on these definitions, two metrics, α and β, are created (see Equations (4) and (5)) to grade every single motif in which α depicts the out-biased score and β indicates the in-biased score.

$$\alpha = \frac{1}{n}\sum |c_O|, \tag{4}$$

$$\beta = \frac{1}{m}\sum |c_I|, \tag{5}$$

where c_o or c_I indicates the nodes' c values that falling into the set $O(n)$ or $I(m)$. The higher the scores are, the less balanced is the motif. Based on Equation (3), it can be proved that a linear relation exists between α and β (see Appendix A). The advantages of adopting these two metrics reflect in two aspects. First, they are good indicators of the local-level system performance and can help designers locate the worst performed subpatterns. Second, not only can these two criteria capture the link weights, but they also integrate the topological characteristics of specific motifs.

At the system level, assuming a complex network G consists of K number of motif g (g is the motif ID in Table 1), two motif-based system performance criteria can be obtained below. Similarly, $\overline{\alpha}_g$ and $\overline{\beta}_g$ hold a linear relationship, and a higher value indicates a worse balance performance.

$$\text{Out-biased score}: \quad \overline{\alpha}_g = \frac{1}{K}\sum_{j=1}^{K} \alpha_{g,j}, \tag{6}$$

$$\text{In-biased score}: \quad \overline{\beta}_g = \frac{1}{K}\sum_{j=1}^{K} \beta_{g,j}. \tag{7}$$

Based on the motif-based system performance criteria, the *seasonal robustness criterion*, as a quantitative representation of the seasonal effect, is defined as the standard deviation of the year-round in- or out-biased score of a motif[1]. For example, according to Step 2, we can get 12 monthly networks G_i ($i = 1, 2, \ldots, 12$) and the yearly significant motif patterns. For each significant motif, its aggregated in- or out-biased score over the 12 consecutive months can be calculated, and the resulting standard deviation from the 12 months, therefore, indicates the system robustness against seasonal changes.

Finally, we define the *capacity planning criterion* based on the capacity (v) of each service node in a network G. We denote the average capacity difference of a motif as

$$d = \frac{|v_1 - v_2| + |v_1 - v_3| + |v_2 - v_3|}{3}, \tag{8}$$

where v_i ($i = 1, 2, 3$) is the capacity of each node i in a size-3 motif. Correspondingly, in the network G consisting of K motif g, the average capacity difference of motif g is

---

[1] Because of the linear relationship, the year-round distributions of $\overline{\alpha}_g$ and $\overline{\beta}_g$ should have a consistent trend, and only the amplitudes are different.

Table 2. The interpretations of the metric in different applications

Motif-based metrics	Interpretation	Application examples
Imbalance score	The *motif-based imbalance score* is proposed to quantitively describe the local service networks' rebalance performance, that is, the difference between the in-flow and out-flow information/traffic of a local system. This metric can quantitively evaluate the local service networks' performance of STSs.	• In the interconnected power grid, the *motif-based imbalance score* is the average difference of transmitting-in and -out power within a local service power grid. • In BSS, the *motif-based imbalance score* is the average difference of numbers of rental and return bikes of a local-level service system (e.g., a system including three service stations).
Seasonal robustness criterion	The *motif-based seasonal robustness criterion* is the standard deviation of the year-round imbalance score. It is a quantitative representation of the seasonal effect where a larger value indicates a local service system is more sensitive to the seasonal demand fluctuation.	• In the interconnected power grid, the *motif-based seasonal robustness criterion* represents how the average difference of transmitting-in and -out power of a local service power grid fluctuates along with seasonal changes. • In BSS, the *motif-based seasonal robustness criterion* represents the variation of the average difference of the rental and return bikes within a local service system along with seasonal changes.
Capacity planning criterion	The *motif-based capacity planning criterion* describes the average capacity difference in a local service network. It is an efficient indicator of whether the local service system's resource distribution is balanced or not.	• In the interconnected power grid, the *motif-based capacity planning criterion* is the average difference of the maximise energy storage ability within a local power grid. • In BSS, the *motif-based capacity planning criterion* is the average difference of the dock numbers within a local service system.

Abbreviations: BSS, bike-sharing system; STS, socio-technical systems.

$$\overline{d}_g = \frac{1}{K} \sum_{j=1}^{K} d_{g,j}. \tag{9}$$

To provide more insights into how the three motif-based metrics be utilised and extended to different systems, a quick overview of the metric interpretations along with their application examples are summarised in Table 2

Step 4: Formulating the design problem and solving for optimal decisions

This step's objectives are twofold: (i) investigate the correlation between seasonal effects (represented by the seasonal robustness criterion) and the capacity planning

criterion, as identified in Step 3 and (ii) formulate the design problem and solve it to obtain optimal decisions for improving the system's robustness against seasonal disturbance. It would be ideal that the factors influencing the system's robustness are known from existing domain knowledge, so such factors will be formulated into the design problem as the decision variable to be optimised. Otherwise, correlation analysis and/or causal inference need to be applied to identify the key design variables.

4. Case study

In this study, the Chicago Bike Share program, Divvy Bikes, is selected to demonstrate the proposed approach. The Divvy Bikes' data is publicly achievable (Divvy_Bike 2020), and the data from 2014 to 2017 are adopted due to the availability of capacity information (i.e., the number of docks) at each station. Figure 5 shows the station distribution of Divvy Bikes in the third and fourth quarters of 2017 and the number of stations in each year. In this study, we aim to mitigate the sensitivity of the system's rebalance performance to seasonal effects.

4.1. Data preprocessing

The station and trip data packages contain information like station geographic coordinates, the number of docks, trip start and end station IDs, trip time and duration, and user basic information (e.g., gender and birth year). We follow four steps to process the raw data of each year. The final data frame consists of 12 monthly trip datasets, each of which has three columns, including start station ID, end station ID, and the reoccurring frequency of each unique trip.

(i) Basic trip information extraction. The essential data, such as the trip start and end station IDs, the number of docks, and start and end times are extracted from the raw dataset.

System Name: Divvy Bike

Locale: Chicago

Date of Operation Began: June 28, 2013

Year	Number of Station
2014	300
2015	475
2016	581
2017	585
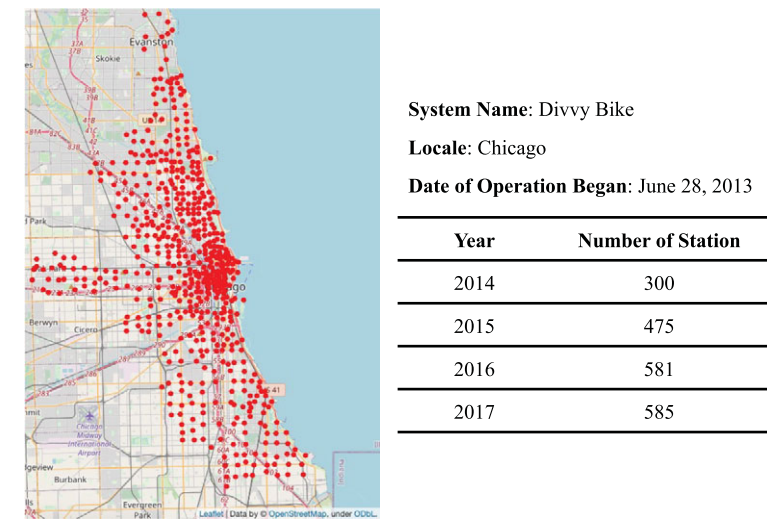
Figure 5. Divvy Bike system information.

   (ii) Data cleaning. We delete those trips with missing data and the the testing stations (e.g., station 512 is a station for testing purpose only) along with their associated trips.

  (iii) Monthly trip network data preparation. In this step, we split trips by month based on their starting time.

  (iv) Trip reoccurring frequency calculation. We count the number of times that a trip between a pair of stations occurs in each month.

4.2. Trip network building and motif mining

Based on the monthly trip datasets, the monthly trip networks are constructed. In each network, stations are represented as nodes, and a trip between two stations is defined as a link and its reoccurring frequency in a month is the link weight. Since the trip from Station A to Station B is different from the trip from B to A, the resulting trip network is a directed network.

    To focus on the network that captures the most significant traffics, we delete those links that have less occurred trips, such as those links with just one-time transit. The threshold for such a link removal process is set as the minimum mean (u) of the link weights among the 12 monthly trip networks in the interested years:

$$u = Min\left(u_{ij}\right), i = \text{index of years}, \ j = 1,2,\dots,12, \tag{10}$$

where u_{ij} is the link weight mean of the jth network in year i. For example, from 2014 to 2017, $u = 3.03$. Then, all the links with weights lower than 3.03 are removed from the network. Figure 6 illustrates the link weight distribution of Divvy Bikes in July 2017. It reveals that the statistical features of link weights will not be altered by removing the links with weights below the threshold. Figure 7 shows the visualisation of the reduced trip network.

    After obtaining the weighted directed trip networks, their binary counterparts (i.e., the same network without link weights) are used for motif mining, which
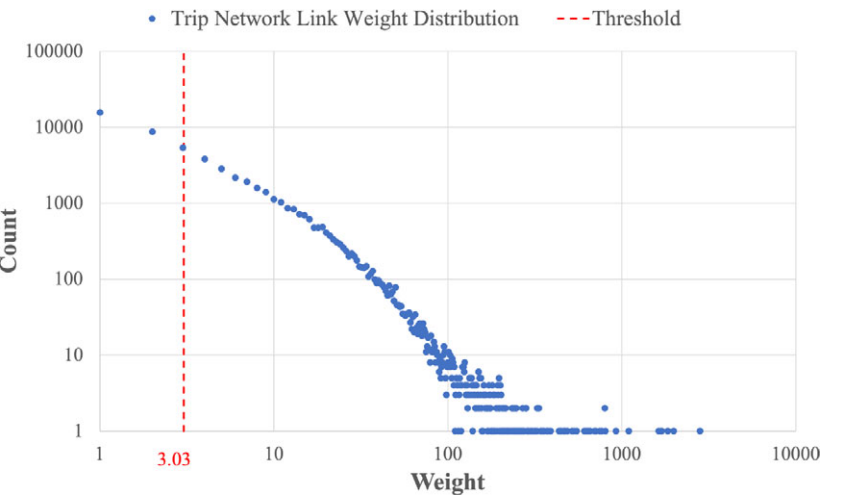
Figure 6. Weight distribution of Divvy Bike trip network (Jul, 2017, total edges: 57,225).
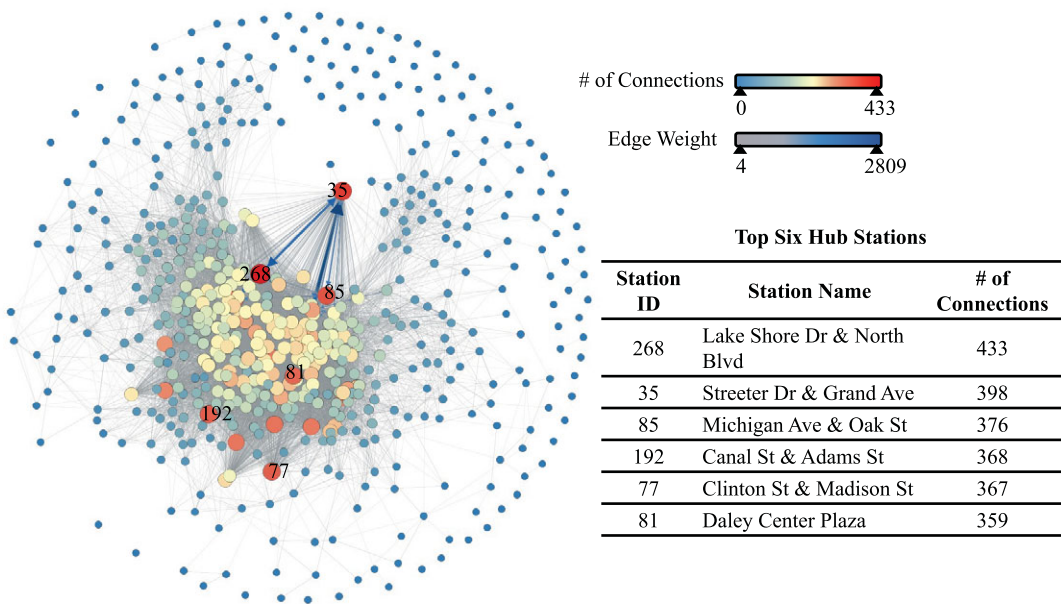
Figure 7. A visualisation of Divvy Bike trip network after removing the links with less occurred trips (Jul, 2017, total edges: 27,415).

reports the motif structures, *Z*-scores, *P*-values, and the adjacent matrix list of all existing motifs. In this study, the motif mining tool FANMOD is adopted. Table 3 shows the 13 size-3 directed motif IDs in every month of 2017, ranked from top to bottom based on their *Z*-scores from high to low. The bold IDs are insignificant motifs under the level of significance 0.001.

According to the results, we observe that the motifs with high transitivity[2] are more likely to be significant and ranked higher in the trip network. This is also the reason that motif 78 is always ranked lowest in all networks. The similar phenomenon is also observed in the years from 2014 to 2016, as shown in Appendix B. In the following analysis, only the significant motif patterns in over 2 years, including motifs 238, 102, 174, 166, 38, 46 and 140, are considered.

4.3. Identifying BSS design parameters and seasonal effect

In our prior work (Xiao & Sha 2020), it is found that seasonal changes can influence the average distances of trip motifs. For example, users tend to ride a longer distance in warmer seasons. Moreover, seasonal changes can impact the traffic of local networks, which is a critical factor to the system rebalance performance. As to the design parameter, based on the correlation analysis (see Table 4), it is found that the number of docks of each station plays an important role in the rebalance problem because it directly relates to the availability of bikes that a user can rent or return in a station.

---

[2] A triad involving nodes *i*, *j* and *k* is transitive if whenever *i* connects to *j* and *j* connects to *k* then *i* connects to *k* (Wasserman, Faust, *et al.* 1994). A digraph has a high transitivity if most triads it contains is transitive.

Table 3. Divvy Bike motif Z-score ranks of each month in 2017

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
238	238	238	46	46	166	46	46	46	46	238	238
102	102	102	102	102	102	166	166	166	102	102	102
46	46	46	166	166	238	102	102	102	166	166	166
174	166	166	238	238	38	238	238	238	238	46	46
166	174	174	38	38	140	38	38	38	38	174	174
38	38	38	174	174	46	140	140	140	174	38	38
140	140	140	140	140	12	12	174	174	140	140	140
12	12	12	12	12	174	174	12	12	12	12	12
6	6	6	6	14	14	14	14	6	6	6	36
36	36	36	14	6	164	164	6	14	14	36	6
164	14	14	164	164	6	6	164	164	164	14	164
14	164	164	36	36	36	36	36	36	36	164	14
78	78	78	78	78	78	78	78	78	78	78	78

Table 4. Divvy Bike yearly correlation coefficient between seasonal effect and motif dock differences

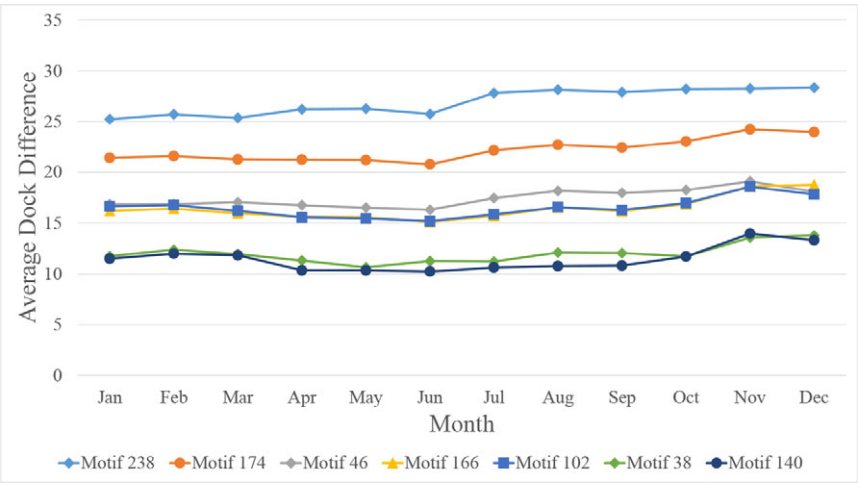| Year	2014	2015	2016	2017
Correlation coefficient	0.848	0.921	0.914	0.922

Figure 8. Divvy Bike yearly motif dock difference curves (2017).

Figure 8 shows the average dock difference of those significant motifs in the 12 months of 2017 following Equations (8) and (9). It is observed that the average dock difference curves from top to bottom correspond to the rank of transitivity of

the motifs from high to low. For example, motif 238, the pattern with the highest transitivity in the trip networks, has the largest dock difference in the entire year of 2017, while motifs 140 and 38 have the smallest differences. While the causation needs to be further investigated, one possible reason for such a correlation could be that the stations with large capacities are more likely located in high-demand areas, thus more users will return or rent bikes. So, they are hub stations and will connect to many other stations. Meanwhile, there is a majority of stations within the system having low capacities. Therefore, these stations and hubs form a large proportion of motifs with high transitivity and large capacity differences.

4.4. Trip motifs performance and robustness analysis

In this section, by following Step 3 in Section 3, we calculate the local BSS rental and return performance scores, which correspond to the motif-based in- and out-biased values. In a BSS, a higher rental or return score indicates that a serious rebalance issue could occur in a trip motif. Figure 9 shows the rebalance performance scores of the seven significant trip motifs.

As indicated in both Figure 9a,b, the trip motif's rebalance performance is potentially related to the motif structure. Taken motifs 46, 166 and 140 as examples, both motifs 46 and 166 have apparent unbalanced structures where *Node 1* only has in- or out-arrows (see Figure 10). This leads them to be vulnerable to the return and rental problems. In contrast, the number of in- and out-arrows of all nodes in motif 140 are the same, so motif 140 is expected to have a low rebalance performance score. However, there are also a few exceptions. For example, motif 238 has a balanced structure but still experiences return and rental problems in several months from April to November. These abnormal fluctuations remind us of the potential seasonal effects, so we use the standard deviation of the return/rental performance scores in a year to quantify such fluctuations, as shown in the second row of Table 5. A larger deviation means that a trip motif is more sensitive to seasonal changes.

4.5. Design problem formulation

To confirm the targeted design variable, we firstly conducted a correlation analysis between the system robustness and the average capacity difference. Since the robustness score is measured based on yearly data, the mean of the capacity differences of every significant motif during the entire year is calculated, as listed in the third row of Table 5. Based on this table, the correlation coefficient between the system robustness and the capacity planning criterion in 2017 can be obtained, and similarly, for the data from 2014 to 2016.

The results are summarised in Table 4 and reveal a significantly high correlation between the capacity difference and the system's robustness. In other words, if a trip motif has a large average capacity difference, its rebalance performance would be more sensitive to seasonal changes. This observation has led to our design objective – to optimise the capacity of the stations in the motifs that are most influential to the system's robustness. To this end, we split the task into two subtasks: (i) identify the stations that need to be optimised for their number of docks and (ii) plan the capacity, that is, the number of docks for those stations, either by adding docks or removing docks, to minimise the average dock difference.
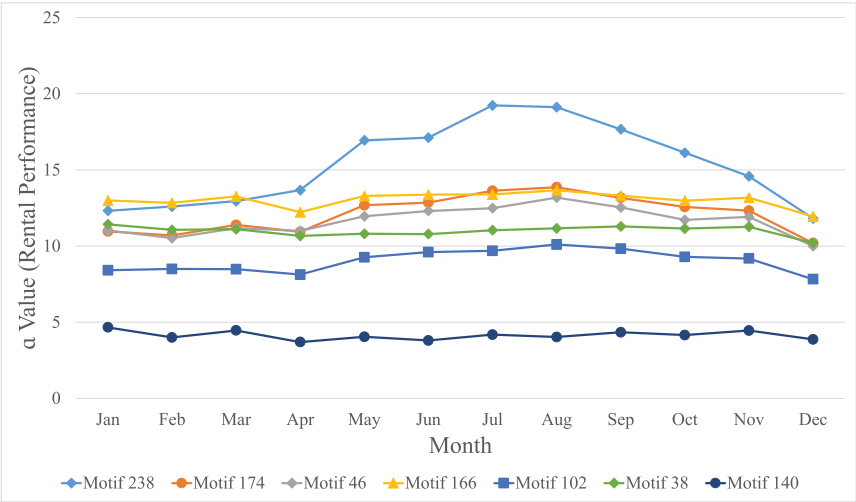
(a) Rental performance
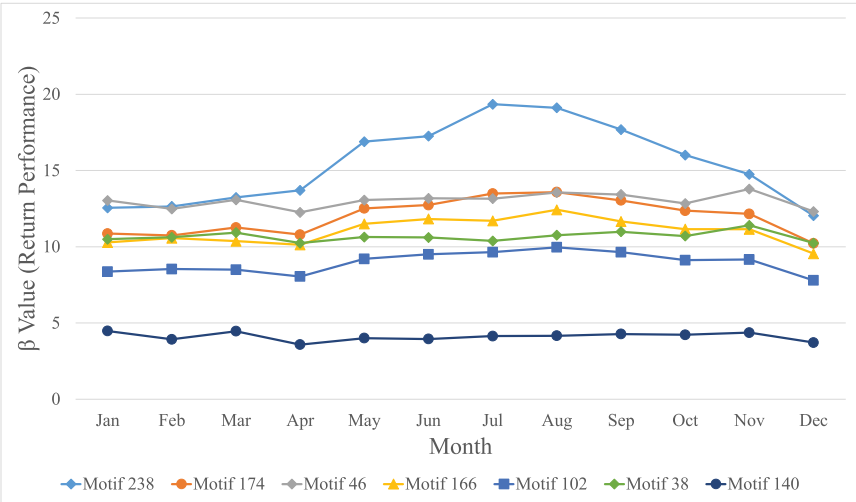
(b) Return performance

Figure 9. Divvy Bike yearly motif rebalance performance (2017).

In the first subtask, the motif pattern that is the most sensitive to seasonal changes is chosen (assuming its ID is g_{season}). Then, we determine the objective motifs with the largest dock differences every month and identify the station IDs that construct those motifs. Based on the number of times those identified stations appear in the objective motifs in each month, two decision rules[3] are used to decide which stations' capacity needs to be optimised.

---

[3] These two rules can be extended. For example, instead of choosing the most frequently appearing stations, the most and the second most frequently appearing stations can be chosen in both rules to achieve more deduction in capacity difference.
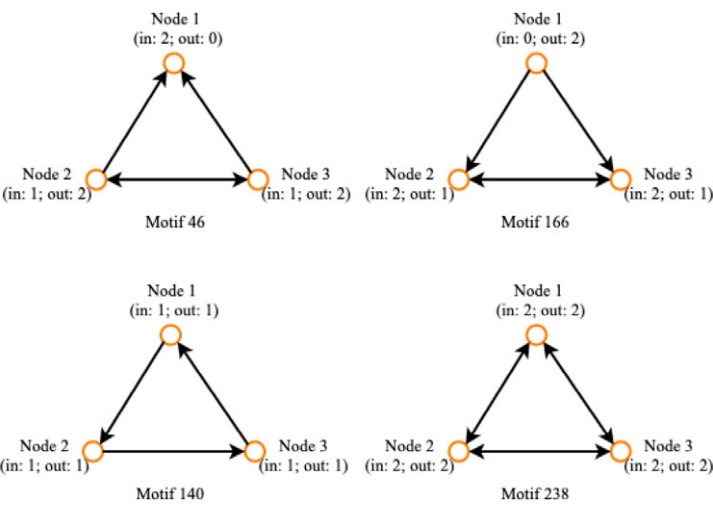
Figure 10. Trip motif structure analysis.

Table 5. Divvy Bike seasonal robustness criteria and capacity planning criteria of significant trip motifs (2017)

Motif ID	238	174	46	166	102	38	140
Seasonal robustness score (the standard deviation of β value)	2.622	1.160	0.476	0.847	0.694	0.331	0.282
Capacity planning criterion	26.931	22.177	17.458	16.467	16.497	11.974	11.452

(i) The first rule is that among all 12 months, if a station appears in the most number of months, then it will be regarded as the critical station and its capacity will be taken into account for optimisation. From the first rule, we will identify a set of critical stations, S_1.

(ii) In the second rule, the stations appearing most frequently in each month are chosen as critical stations, and the corresponding station set is defined as S_2. Finally, we define all the critical stations being represented as $S = S_1 \cup S_2$.

In the second subtask, assuming the significant motif set is M, including m different types of motifs, we identify the significant motifs (from M) in which a critical station s ($s \in S$) appears, and put the same type of motifs with the ID $g \in M$ in one set, $M_{s,g}$. Next, we define the decision variable x_s as the number of docks that station s need to add ($x_s > 0$) or remove ($x_s < 0$). Then, the updated average capacity difference of the motif g, $d_{s,g,j}$, can be calculated by following Equation (11), where s_1, s_2 and s_3 represent three stations' IDs in a motif. Depending on whether s_1, s_2 and s_3 belong to the critical station set S or not, $d_{s,g,j}$ is calculated differently.

$$
\begin{cases}
\begin{aligned}
d_{s,g,j}(x_{s_1}) = \frac{1}{3}[&|(v_{s_1} + x_{s_1}) - v_2| \\
&+ |(v_{s_1} + x_{s_1}) - v_3| + |v_2 - v_3|]
\end{aligned} & \text{if } s_1 \in S \\
\begin{aligned}
d_{s,g,j}(x_{s_1}, x_{s_2}) = \frac{1}{3}[&|(v_{s_1} + x_{s_1}) - (v_{s_2} + x_{s_2})| \\
&+ |(v_{s_1} + x_{s_1}) - v_3| + |(v_{s_2} + x_{s_2}) - v_3|]
\end{aligned} & \text{if } s_1, s_2 \in S \\
\begin{aligned}
d_{s,g,j}(x_{s_1}, x_{s_2}, x_{s_3}) = \frac{1}{3}[&|(v_{s_1} + x_{s_1}) - (v_{s_2} + x_{s_2})| \\
&+ |(v_{s_1} + x_{s_1}) - (v_{s_3} + x_{s_3})| \\
&+ |(v_{s_2} + x_{s_2}) - (v_{s_3} + x_{s_3})|]
\end{aligned} & \text{if } s_1, s_2, s_3 \in S
\end{cases}
\tag{11}
$$

Finally, the updated average dock difference for motifs in set $M_{s,g}$ can be obtained

$$
\overline{d}_{s,g} = \frac{1}{m_{s,g}} \sum_{j=1}^{m_{s,g}} d_{s,g,j},
\tag{12}
$$

where $m_{s,g}$ is the number of motifs in $M_{s,g}$.

Since the objective is to minimise the average dock difference of those identified trip motifs, a multiobjective optimisation is formulated in Equation (13):

$$
\begin{cases}
\min \overline{d}_{s_1,g_1} = \min \frac{1}{m_{s_1,g_1}} \sum_{j=1}^{m_{s_1,g_1}} d_{s_1,g_1,j} \\
\dots \\
\min \overline{d}_{s_1,g_m} = \min \frac{1}{m_{s_1,g_m}} \sum_{j=1}^{m_{s_1,g_m}} d_{s_1,g_m,j} \\
\dots \\
\min \overline{d}_{s_l,g_1} = \min \frac{1}{m_{s_l,g_1}} \sum_{j=1}^{m_{s_l,g_1}} d_{s_l,g_1,j} \\
\dots \\
\min \overline{d}_{s_l,g_m} = \min \frac{1}{m_{s_l,g_m}} \sum_{j=1}^{m_{s_l,g_m}} d_{s_l,g_m,j}
\end{cases}
\tag{13}
$$

$$
\text{S.T. } x_{s_1} \geq -v_{s_1}, \dots, x_{s_l} \geq -v_{s_l} \text{ and } x_{s_1}, \dots, x_{s_l} \in \mathbf{Z},
$$

where $s_1, \dots, s_l \in S$, $g_1, \dots, g_m \in M$ and $v_{s_1}, \dots, v_{s_l}$ are the original dock numbers of station $s_1, \dots, s_l$. $\mathbf{Z}$ denotes Integer. In Equation (13), all the relevant motifs in M, even if they are not g_{season}, are considered. This is because while we are changing the number of docks for those stations in motif g_{season}, there is a possibility that the average dock difference in the other motifs which include the stations of g_{season} increases too.

To solve this optimisation problem, we adopt the weighting method (Miettinen 2012) to transform the multiobjective optimisation problem to a single-objective one in Equation (14). Suppose all objective functions in Equation (13) are equally important, and $\sum_{i=1}^{m \times l} q_i = 1$, then $q_i = q = \frac{1}{m \times l}$ ($i = 1, \dots, m \times l$). Equation (14) is a typical nonlinear integer optimisation problem, and the genetic algorithm, *ga* function in MATLAB (2020) is applied to solve this problem.

$$\min D(x_{s_1}, \ldots, x_{s_l}) = \min \left[q \left(\overline{d}_{s_1, g_1} + \overline{d}_{s_1, g_m} + \ldots + \overline{d}_{s_l, g_1} + \overline{d}_{s_l, g_m} \right) \right], \quad (14)$$

$$\text{S.T. } x_{s_1} \geq -v_{s_1}, \ldots, x_{s_l} \geq -v_{s_l} \text{ and } x_{s_1}, \ldots, x_{s_l} \in Z,$$

In our case study, based on Figure 9, we identify $M = \{238, 174, 46, 166, 102, 38, 140\}$ and motif 238 is the target motif we need to focus on because it is the most sensitive one in light of seasonal changes. Table 6 lists all the critical stations which form motif 238 and yield the largest dock difference. From Table 6, we can observe that, station 3, the most frequently appeared station (9 months out of 12), should be considered as a critical station, that is, $S_1 = \{3\}$. Regarding the stations that appear most in each month, taking March as an example, we identified 15 critical motif 238s, and station 35 and 172 are the most frequently appeared stations in all of the 15 critical motifs. Thus, they are considered as critical stations. Similarly, another four critical stations are identified, thus $S_2 = \{3, 35, 45, 97, 172, 263\}$. By combining these two sets, we obtain the final critical station set $S = S_1 \cup S_2 = \{3, 35, 45, 97, 172, 263\}$.

By solving the optimisation problem in Equation (14), we obtain the optimal capacity planning decision for the decision variables $x_{s_1}, \ldots, x_{s_6}$. The results are shown in Table 7, along with the original and updated number of docks. To verify if the redesigned capacity can effectively decrease the average dock difference of the significant trip motifs or not, we recalculate the trip motifs' mean values of the updated number of docks in a year, as shown in Table 8. By comparing the updated dock differences with the original ones, it is found that the decreases are achieved for all significant motifs, and the dock difference of motif 238 is decreased by 4.6%. With such a decrease, the enhance of the system robustness against seasonal effects is expected to be achieved effectively.

5. Conclusion

It is the uncertainty and complicated interactions within an STS that make the system vulnerable to various perturbations. The occurrence of certain perturbations can significantly influence the STS performance, and the seasonal effect is a common one because it directly impacts human behaviour in STS. In this study, we develop a new design framework for improving STS robustness against seasonal changes based on the network motif theory. Using the concepts of motif, we created three metrics for system performance evaluation and capacity planning decision-making. The first one is the imbalance score of a motif (e.g., a local service network), the second one is the measurement of a motif's seasonal robustness, and the third one is a design parameter-based capacity planning decision criterion. We apply our developed approach to a real-world STS, Divvy Bikes, a Chicago Bike Share program, to improve the system's rebalance performance and its robustness against seasonal changes. The results from this study show that our approach can effectively reduce the average dock differences among the stations of critical trip motifs (i.e., local trip networks), thereby improving the system's robustness.

The main contributions of this paper are summarised in three aspects: (i) We introduce a network motif-based approach for guiding the STS robust design, emphasising optimising system capacity planning to weaken the impact from demand fluctuations caused by seasonal changes. (ii) We propose a set of motif-based criteria to

Design Science

Table 6. Station list of constructing the motif 238s with the largest dock difference values

Jan Station ID	Jan ncy	Feb Station ID	Feb ncy	Mar Station ID	Mar ncy	Apr Station ID	Apr ncy	May Station ID	May ncy	Jun Station ID	Jun ncy	Jul Station ID	Jul ncy	Aug Station ID	Aug ncy	Sep Station ID	Sep ncy	Oct Station ID	Oct ncy	Nov Station ID	Nov ncy	Dec Station ID	Dec ncy
35	5	35	20	35	15	97	82	97	138	97	255	3	25	3	10	3	5	3	26	3	26	3	6
172	5	172	19	172	15	35	75	172	95	172	149	42	17	263	10	263	5	172	26	24	18	45	6
59	2	97	11	25	1	288	23	195	43	35	134	263	8	338	1	150	1	268	1	97	8	97	3
97	2	45	7	321	1	24	22	24	29	192	88	623	2	287	1	76	1	264	1	59	3	59	2
25	1	150	4	301	1	199	21	199	29	24	45	338	2	255	1	62	1	211	1	90	3	37	1
45	1	76	2	286	1	45	17	45	28	199	44	237	2	237	1	4	1	199	1	255	2	90	1
194	1	255	2	283	1	338	17	288	25	45	39	150	2	150	1			196	1	195	2	255	1
176	1	4	1	211	1	150	11	173	21	288	37	62	2	76	1			194	1	174	2	192	1
338	1	6	1	199	1	192	11	192	20	338	36	370	1	62	1			192	1	91	2	168	1
126	1	24	1	176	1	146	8	338	19	173	35	178	1	6	1			181	1	52	2	91	1
99	1	25	1	173	1	76	6	150	15	112	29	147	1	4	1			177	1	43	2	43	1
		26	1	161	1	90	6	146	11	370	23	120	1					173	1	35	2		
		291	1	145	1	3	5	370	10	150	23	91	1					161	1	26	2		
		289	1	141	1	43	5	76	9	146	20	81	1					142	1	370	2		
		286	1	99	1	49	5	90	8	3	10	72	1					99	1	338	2		
		268	1	77	1	85	5	6	7	76	10	59	1					91	1	341	1		
		211	1	47	1	284	5	26	7	85	10	45	1					90	1	284	1		
		199	1			255	5	43	7	90	10	255	1					85	1	264	1		
		181	1					52	7	2	9	76	1					76	1	194	1		
		177	1					268	7	6	9	4	1					52	1	47	1		
		173	1					2	6	43	9	41	1					49	1	44	1		
		168	1					3	6	52	9	33	1					38	1	41	1		
		99	1					85	6	91	9	5	1					37	1	25	1		
		341	1					99	6	255	9							36	1	623	1		
		164	1					284	6	26	8							35	1	4	1		
		145	1					255	6	36	8							26	1				
		141	1					4	5	100	8							25	1				
		126	1					62	5	268	8							24	1				
		110	1					142	5	264	8												
		94	1					145	5	181	8												
		47	1					341	5	4	7												
		90	1					309	5	37	7												
										44	7												
										99	7												
										419	7												
										211	7												
										196	7												
										195	7												
										321	7												
										33	6												
										59	6												
										62	6												
										75	6												
										145	6												
										341	6												
										284	6												
										177	6												
										623	5												

☐ : Among twelve months, stations that appear in the most number of months

☐ : Stations that appear most in each month

Due to the space limitation, stations with appearing frequencies less than 5 in April, May and June are ignored, and this ignorance has no effect on critical station identification.

Table 7. The calculating results of Equation (14).

Station ID	3	35	45	97	172	263
Original dock number	55	47	15	55	11	11
Added/deleted dock number[a]	−11	−11	8	−32	12	12
Updated dock number	44	36	23	23	23	23

[a]Correspond to $x_{s_1}, \ldots, x_{s_6}$.

Table 8. Divvy Bike yearly mean values of significant motif dock differences, before update versus after update (2017)

Motif ID	238	174	46	166	102	38	140
The mean of the motif dock difference before update	26.931	22.177	17.458	16.467	16.497	11.974	11.452
The mean of the motif dock difference after update	25.679	21.199	16.758	15.810	15.832	11.557	11.044
The percentage of decrease (%)	4.6	4.4	4.0	4.0	4.0	3.5	3.6

help evaluate system's performance and the impact of seasonal effects on it. (iii) High correlation between the seasonal effects and the average dock difference of motifs is discovered in BSS, from which a multiobjective design problem is formulated to aid capacity planning decisions for improved system robustness.

There are two limitations in this study. First, in the STS robustness analysis, only seasonal effects are considered. However, in reality, it is common that several types of disturbances, such as the explicit interaction of BSS with other public transportation systems and varying population growth in different areas, could co-exist and influence a system's performance in a more unpredictable and dynamic manner. Second, in the robust design, it is expected to have a predictive model for the trip network so that after docks are added or removed at the critical stations, the resulting imbalance scores can be updated and the local trip networks' performance can be re-evaluated to further verify the effectiveness of the design solutions. Taking this work as a starting point, we would like to develop a dynamic approach, in conjunction with the temporal motif concept, to support STS short-term robustness analysis. In this case, more temporal uncertainties, such as varying demands at different periods of a day or system self-rebalancing strategy (e.g., bike-sharing company utilising trucks to rebalance the number of bikes in different stations), will be considered. Furthermore, we would also like to pursue a more comprehensive framework to guide the robust design of complex STS by taking into account more influence disturbances.

References

Ahlström, A., **Schurgers, G.**, **Arneth, A.** & **Smith, B.** 2012 Robustness and uncertainty in terrestrial ecosystem carbon response to cmip5 climate change projections. *Environmental Research Letters* **7** (4), 044008; doi:10.1088/1748-9326/7/4/044008.

Alon, U. 2007. Network motifs: theory and experimental approaches. *Nature Reviews Genetics* **8** (6), 450–461; doi:10.1038/nrg2102.

Baxter, G. & **Sommerville, I.** 2011. Socio-technical systems: From design methods to systems engineering. *Interacting with Computers* **23** (1), 4–17.

Bi, Y., **Xie, J.**, **Sha, Z.**, **Wang, M.**, **Fu, Y.** & **Chen, W.** 2018. Modeling spatiotemporal heterogeneity of customer preferences in engineering design. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 51753)*, p. V02AT03A050. American Society of Mechanical Engineers; doi:10.1115/DETC2018-86245.

Blume, S. W. 2017. Interconnected power systems, in *Electric Power System Basics for the Nonelectrical Professional*. Hoboken, NJ, USA: John Wiley & Sons, inc., pp. 165–186; doi:10.1002/9781119180227.ch8.

Carayon, P. 2006. Human factors of complex sociotechnical systems. *Applied Ergonomics* **37** (4), 525–535; doi:10.1016/j.apergo.2006.04.011.

Cats, O., **Koppenol, G.-J.** & **Warnier, M.** 2017. Robustness assessment of link capacity reduction for complex networks: Application for public transport systems. *Reliability Engineering & System Safety* **167**, 544–553. doi:10.1016/j.ress.2017.07.009.

Choobdar, S., **Ribeiro, P.** & **Silva, F.** 2012. Motif mining in weighted networks. In *2012 IEEE 12th International Conference on Data Mining Workshops*, pp. 210–217. IEEE; doi:10.1109/ICDMW.2012.111.

Cui, Y., **Ahmed, F.**, **Sha, Z.**, **Wang, L.**, **Fu, Y.** & **Chen, W.** 2020. A weighted network modeling approach for analyzing product competition. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 84003)*, p. V11AT11A036. American Society of Mechanical Engineers; doi:10.1115/DETC2020-22591.

Divvy_Bike 2020. Divvy system data (downloadable on February 21st 2020), https://www.divvybikes.com/system-data

ElMaraghy, W., **ElMaraghy, H.**, **Tomiyama, T.** & **Monostori, L.** 2012. Complexity in engineering design and manufacturing. *CIRP Annals* **61** (2), 793–814; doi:10.1016/j.cirp.2012.05.001.

Felmlee, D., **McMillan, C.**, **Towsley, D.** & **Whitaker, R.** 2018. Social network motifs: A comparison of building blocks across multiple social networks. In *Annual Meetings of the American Sociological Association, Philadelphia, US.*

Fu, J. S., **Sha, Z.**, **Huang, Y.**, **Wang, M.**, **Fu, Y.** & **Chen, W.** 2017. Two-stage modeling of customer choice preferences in engineering design using bipartite network analysis. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 58127)*, p. V02AT03A039. American Society of Mechanical Engineers; doi:10.1115/DETC2017-68099.

Garca, S., **Luengo, J.** & **Herrera, F.** 2015. *Data Preprocessing in Data Mining (Vol. 72)*. Springer.

Geels, F. W. & **Kemp, R.** 2007. Dynamics in socio-technical systems: Typology of change processes and contrasting case studies. *Technology in Society* **29** (4), 441–455; doi:10.1016/j.techsoc.2007.08.009.

Gribble, S. D. 2001. Robustness in complex systems. In *Proceedings Eighth Workshop on Hot Topics in Operating Systems*, pp. 21–26. Springer; doi:10.1109/HOTOS.2001.990056.

Hao, X., **Wang, H.**, **Lin, Z.** & **Ouyang, M.** 2020. Seasonal effects on electric vehicle energy consumption and driving range: A case study on personal, taxi, and ridesharing vehicles. *Journal of Cleaner Production* **249**, 119403; doi:10.1016/j.jclepro.2019.119403.

Hassannezhad, M., **Cantamessa, M.**, **Montagna, F.** & **Clarkson, P. J.** 2019. Managing sociotechnical complexity in engineering design projects. *Journal of Mechanical Design* **141** (8), 1; doi:10.1115/1.4042614.

Holland, P. W. & **Leinhardt, S.** 1974. The statistical analysis of local structure in social networks. *National Bureau of Economic Research, Inc, NBER Working Papers* 6; doi: 10.3386/w0044.

Kalsi, M., **Hacker, K.** & **Lewis, K.** 1999. A comprehensive robust design approach for decision trade-offs in complex systems design. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 19715)*, pp. 1343–1354. American Society of Mechanical Engineers; doi:10.1115/1.1334596.

Kashtan, N., **Itzkovitz, S.**, **Milo, R.** & **Alon, U.** 2002. Mfinder tool guide, Department of Molecular Cell Biology and Computer Science and Applied Math., Weizmann Inst. of Science, Rehovot Israel, technical report.

Kashtan, N., **Itzkovitz, S.**, **Milo, R.** & **Alon, U.** 2004. Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics* **20** (11), 1746–1758; doi:10.1093/bioinformatics/bth163.

Keating, C. B., **Fernandez, A. A.**, **Jacobs, D. A.** & **Kauffmann, P.** 2001. A methodology for analysis of complex sociotechnical processes. *Business Process Management Journal* **7** (1), 33–50; doi:10.1108/14637150110383926.

Markolf, S. A., **Hoehne, C.**, **Fraser, A.**, **Chester, M. V.** & **Underwood, B. S.** 2019. Transportation resilience to climate change and extreme weather events–beyond risk and robustness. *Transport Policy* **74**, 174–186; doi:10.1016/j.tranpol.2018.11.003.

Martnez-Costa, C., **Mas-Machuca, M.**, **Benedito, E.** & **Corominas, A.** 2014. A review of mathematical programming models for strategic capacity planning in manufacturing. *International Journal of Production Economics* **153**, 66–85; doi:10.1016/j.ijpe.2014.03.011.

MATLAB 2020. Matlab genetic algorithm (downloadable on February 21st 2020), https://www.mathworks.com/help/gads/ga.html#d122e41247

Miettinen, K. 2012. *Nonlinear Multiobjective Optimization (Vol. 12)*. Springer Science & Business Media.

Milo, R., **Shen-Orr, S.**, **Itzkovitz, S.**, **Kashtan, N.**, **Chklovskii, D.** & **Alon, U.** 2002. Network motifs: Simple building blocks of complex networks. *Science* **298** (5594), 824–827; doi: 10.1126/science.298.5594.824.

Pan, S. L. & **Scarbrough, H.** 1998. A socio-technical view of knowledge sharing at buckman laboratories. *Journal of Knowledge Management* **2** (1), 55–66; doi:10.1108/EUM0000000004607.

Paparistodimou, G., **Duffy, A.**, **Whitfield, R. I.**, **Knight, P.** & **Robb, M.** 2020. A network tool to analyse and improve robustness of system architectures. *Design Science* **6**, e8; doi: 10.1017/dsj.2020.6.

Paranjape, A., **Benson, A.** & **Leskovec, J.** 2017. Motifs in temporal networks. In *Proceedings of the International Conference on Web Search and Data Mining*, pp. 601–610; doi: 10.1145/3018661.3018731.

Piccolo, S. A., **Lehmann, S.** & **Maier, A.** 2018. Design process robustness: A bipartite network analysis reveals the central importance of people. *Design Science* **4**, e1; doi: 10.1017/dsj.2017.32.

Rasche, F. & **Wernicke, S.** 2006. Fanmod fast network motif detection–manual. *Bioinformatics* **22** (9), 1152–1153.

Sha, Z., **Bi, Y.**, **Wang, M.**, **Stathopoulos, A.**, **Contractor, N.**, **Fu, Y.** & **Chen, W.** 2019. Comparing utility-based and network-based approaches in modeling customer preferences for engineering design. In *Proceedings of the Design Society: International Conference on Engineering Design (Vol. 1)*, pp. 3831–3840. Cambridge University Press; doi:10.1017/dsi.2019.390.

Sha, Z., **Chaudhari, A. M.** & **Panchal, J. H.** 2019. Modeling participation behaviors in design crowdsourcing using a bipartite network-based approach. *Journal of Computing and Information Science in Engineering* **19** (3), 031010. doi:10.1115/1.4042639.

Sha, Z., **Huang, Y.**, **Fu, J. S.**, **Wang, M.**, **Fu, Y.**, **Contractor, N.** & **Chen, W.** 2018. A network-based approach to modeling and predicting product coconsideration relations. *Complexity* **2018**, 2753638; doi:10.1155/2018/2753638.

Sha, Z. & **Panchal, J. H.** 2013a. Estimating the node-level behaviors in complex networks from structural datasets. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 55867)*, p. V02BT02A003. American Society of Mechanical Engineers; doi:10.1115/DETC2013-12063.

Sha, Z. & **Panchal, J. H.** 2013b. Towards the design of complex evolving networks with high robustness and resilience. *Procedia Computer Science* **16**, 522–531; doi:10.1016/j.procs.2013.01.055.

Sha, Z. & **Panchal, J. H.** 2016. A degree-based decision-centric model for complex networked systems. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 50084)*, p. V01BT02A016. American Society of Mechanical Engineers; doi:10.1115/DETC2016-60036.

Stone, L., **Simberloff, D.** & **Artzy-Randrup, Y.** 2019. Network motifs and their origins. *PLoS Computational Biology* **15** (4), e1006749; doi:10.1371/journal.pcbi.1006749.

Sun, X., **Wandelt, S.** & **Linke, F.** 2015. Temporal evolution analysis of the european air transportation system: air navigation route network and airport network. *Transportmetrica B: Transport Dynamics* **3** (2), 153–168. doi:10.1080/21680566.2014.960504.

Ten Napel, J., **Van der Veen, A.**, **Oosting, S.** & **Koerkamp, P. G.** 2011. A conceptual approach to design livestock production systems for robustness to enhance sustainability. *Livestock Science* **139** (1–2), 150–160; doi:10.1016/j.livsci.2011.03.007.

Trist, E. L. & **Bamforth, K. W.** 1951. Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human Relations* **4** (1), 3–38; doi:10.1177/001872675100400101.

Urruty, N., **Tailliez-Lefebvre, D.** & **Huyghe, C.** 2016. Stability, robustness, vulnerability and resilience of agricultural systems. A review. *Agronomy for Sustainable Development* **36** (1), 15. doi:10.1007/s13593-015-0347-5.

Wang, M., **Chen, W.**, **Huang, Y.**, **Contractor, N. S.** & **Fu, Y.** 2016. Modeling customer preferences using multidimensional network analysis in engineering design. *Design Science* **2**, e11; doi:https://doi.org/10.1017/dsj.2016.11.

Wang, M., **Sha, Z.**, **Huang, Y.**, **Contractor, N.**, **Fu, Y.** & **Chen, W.** 2016. Forecasting technological impacts on customers' co-consideration behaviors: A data-driven network analysis approach. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 50107)*, p. V02AT03A040. American Society of Mechanical Engineers; doi:10.1115/DETC2016-60015.

Wang, M., **Sha, Z.**, **Huang, Y.**, **Contractor, N.**, **Fu, Y.** & **Chen, W.** (2018), 'Predicting product co-consideration and market competitions for technology-driven product design: a network-based approach', *Design Science* **4**. DOI: 10.1017/dsj.2018.4.

Wang, T., **Peng, J.**, **Peng, Q.**, **Wang, Y.** & **Chen, J.** 2020. Fsm: Fast and scalable network motif discovery for exploring higher-order network organizations. *Methods* **173**, 83–93; doi:10.1016/j.ymeth.2019.07.008.

Wasserman, S. & **Faust, K.** 1994. Social network analysis: Methods and applications. Cambridge University Press; doi: 10.1017/CBO9780511815478.

Wernicke, S. & **Rasche, F.** 2006. Fanmod: A tool for fast network motif detection. *Bioinformatics* **22** (9), 1152–1153; doi:10.1093/bioinformatics/btl038.

Design Science

Xiao, Y. & **Sha, Z.** 2020. Towards engineering complex socio-technical systems using network motifs: A case study on bike-sharing systems. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 84003)*, p. V11AT11A045. American Society of Mechanical Engineers; doi: 10.1115/DETC2020-22631.

Appendix A. Validating the linear relationship between α and β

Based on Equation (3), all the possible calculation of α and β, depending on whether c values are larger than 0 or not, are enumerated as follows:

(i) If $c_1 > 0$, $c_2 < 0$, and $c_3 < 0$:

$$\beta = c_1, \quad \alpha = \frac{1}{2}|c_2 + c_3|, \tag{A1}$$

$$\text{So,} \quad \beta = 2\alpha. \tag{A2}$$

Similar relationship can be achieved when $c_2 > 0$, $c_1 < 0$ and $c_3 < 0$ or $c_3 > 0$, $c_1 < 0$ and $c_2 < 0$.

(ii) If $c_1 > 0$, $c_2 > 0$, and $c_3 < 0$:

$$\beta = \frac{1}{2}(c_1 + c_2), \quad \alpha = |c_3|, \tag{A3}$$

$$\text{So,} \quad \beta = \frac{1}{2}\alpha. \tag{A4}$$

Similar relationship can be achieved when $c_1 > 0$, $c_3 > 0$ and $c_2 < 0$ or $c_2 > 0$, $c_3 > 0$ and $c_1 < 0$.

(iii) If $c_1 = 0$, $c_2 > 0(< 0)$ and $c_3 < 0(> 0)$:

$$\beta = c_2, \quad \alpha = |c_3|, \tag{A5}$$

$$\text{So,} \quad \beta = \alpha. \tag{A6}$$

Similar relationship can be achieved when $c_1 > 0(< 0)$, $c_2 = 0$ and $c_3 < 0(> 0)$ or $c_1 > 0(< 0)$, $c_2 < 0(> 0)$ and $c_3 = 0$.

(iv) If $c_1 = 0$, $c_2 = 0$ and $c_3 = 0$:

$$\beta = \alpha = 0. \tag{A7}$$

Therefore, the linear relationship between α and β is validated.

Appendix B. Divvy Bike motif *Z*-score ranks in 2014–2016

From Table B1 to Table B3, the same trend described in Section 4.2 is further verified that the motifs with higher transitivity are more likely to be significant. This is also the reason that motif 238 and 46 are always ranked highest while motif 78 lowest in almost all networks.

Table B1. Divvy Bike motif *Z*-score ranks of each month in 2014

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
238	238	238	46	46	46	46	46	166	46	238	238
102	174	46	238	166	166	166	166	46	102	166	102
174	46	102	102	102	102	238	102	238	166	102	46
38	102	166	166	238	238	102	238	102	238	46	166
46	38	174	38	38	38	14	164	38	38	174	174
166	166	38	174	12	164	164	14	12	14	38	38
140	140	140	140	164	14	38	12	14	12	12	140
36	12	12	12	14	12	140	38	164	164	140	12
12	36	6	14	140	140	12	140	140	140	36	36
6	6	36	164	174	174	6	174	174	174	6	6
164	164	164	6	6	6	174	6	36	6	164	164
14	14	14	36	36	36	36	36	6	36	14	14
78	78	78	78	78	78	78	78	78	78	78	78

Table B2. Divvy Bike motif *Z*-score ranks of each month in 2015

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
238	238	238	238	46	46	46	46	46	46	102	238
102	174	166	166	166	166	166	166	166	166	166	102
166	102	102	46	102	102	102	102	102	102	238	46
174	46	46	102	238	238	238	238	238	238	46	166
46	166	174	38	38	38	38	38	38	38	174	174
38	38	38	174	140	140	140	140	140	174	38	38
140	140	140	140	174	12	12	12	174	140	140	140
12	12	12	12	12	174	174	174	12	12	12	12
36	36	36	164	164	164	164	164	164	14	6	6
6	6	6	14	14	14	14	14	14	6	36	36
14	164	164	36	6	6	6	6	6	164	164	164
164	14	14	6	36	36	36	36	36	36	14	14
78	78	78	78	78	78	78	78	78	78	78	78

Table B3. Divvy Bike motif Z-score ranks of each month in 2016

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
238	238	102	102	46	46	46	46	46	46	46	238
102	102	46	46	166	166	102	102	102	102	102	102
174	166	238	166	102	102	166	166	166	166	166	166
166	46	166	238	238	238	238	238	238	238	238	174
46	174	38	38	38	38	38	38	38	38	38	46
38	38	174	174	140	12	140	174	140	174	174	38
140	140	140	140	174	140	174	140	174	140	140	140
12	12	12	12	12	174	12	12	12	12	12	12
36	36	6	6	14	14	164	6	14	36	6	36
6	6	36	164	164	164	14	14	6	14	36	6
14	164	14	36	6	6	6	164	164	6	14	14
164	14	164	14	36	36	36	36	36	164	164	164
78	78	78	78	78	78	78	78	78	78	78	78